

Technische Universität Berlin
Fakultät I Geisteswissenschaften
Institut für Sprache und Kommunikation

**Entwicklung einer Multimedia-Lehranwendung
zur Psychoakustik: Auditive Wahrnehmung
und die Rolle kritischer Bandbreiten**

**Dokumentation des Stands der Wissenschaft
Implementierung der Lehreinheit**

Magisterarbeit im Fachgebiet Kommunikationswissenschaft

Betreut durch:
Prof. Dr. Klaus Hobohm (Gastprofessor)

Vorgelegt von:
Sebastian Kauer
Fehrbelliner Straße 39
10119 Berlin
Matr.-Nr. 171395

Inhaltsverzeichnis

1	Einleitung	7
1.1	Ziel dieser Arbeit	7
1.2	Zur Vorgehensweise	8
1.3	Zu Notation & Terminologie	9
Teil A: Auditive Wahrnehmung und kritische Bandbreiten		
2	Begriffe und Grundlagen	10
2.1	Einordnung des Themengebietes	10
2.2	Psychometrie und Methoden der Psychoakustik	10
2.3	Physikalische Größen und Hörempfindungsgrößen	12
2.4	Akustische Filter.....	14
3	Gehöranatomie und -physiologie	15
3.1	Das auditive System	15
3.2	Das Außenohr.....	16
3.3	Das Mittelohr	16
3.4	Das Innenohr	17
3.4.1	Die Hörschnecke (Cochlea).....	17
3.4.2	Basilarmembran, Cortisches Organ und Haarzellen.....	18
3.5	Stimulationsmechanismus des Gehörsinns	21
4	Die Frequenzanalyse durch das Gehör	23
4.1	Die Frequenzselektivität der Basilarmembran	23
4.1.1	Das Wanderwellen-Prinzip der Basilarmembranschwingung	23
4.1.2	Die Funktion der äußeren Haarzellen	26
4.2	Die Frequenzselektivität der Hörnervenfasern.....	27
4.3	Phasenkopplung der neuronalen Entladungen	28
4.4	Frequenzcodierung	29
4.4.1	Örtliche Codierung	30
4.4.2	Zeitliche Codierung	30
4.5	Die auditive Vorverarbeitung durch die Cochlea.....	31
5	Auditive Filter und kritische Bandbreiten	32
5.1	Historische Entwicklung	32
5.1.1	Entdeckung des kritische Bandbreiten-Phänomens	32
5.1.2	Energiespektrum-Modell und Critical Ratio.....	33
5.1.3	Bark-Skala, Frequenzgruppen und Kanalmodell	34
5.1.4	Unsicherheit über physiologische Zusammenhänge	36
5.1.5	Neuausrichtung der Filtertheorie und neue Bandbreitenwerte	37
5.1.6	Zur Definition des Begriffs „kritische Bandbreite“.....	38
5.2	Maskierung und neuronale Erregungsmuster	39
5.2.1	Maskierung und das Frequenzauflösungsvermögen des Gehörs	39
5.2.2	Psychophysische Maskierungsmuster.....	40

5.2.3	Berechnung neuronaler Erregungsmuster.....	41
5.2.4	Maskierung als Effekt der Überlagerung neuronaler Aktivität.....	42
5.2.5	Modelle der Lautheit	43
5.3	Eigenschaften der auditiven Filter.....	44
5.3.1	Die Form der auditiven Filter.....	44
5.3.1.1	Psychophysische Frequenz-Tuningkurven.....	44
5.3.1.2	Die Notched-Noise Methode.....	46
5.3.1.3	Asymmetrie und Pegelabhängigkeit der Filterform.....	47
5.3.2	Die Bandbreite der auditiven Filter.....	47
5.3.3	Nicht-simultane Maskierung und auditive Filter	50
6	Kritische Bandbreiten als empirisches Phänomen.....	51
6.1	Lautheit komplexer Schallsignale.....	51
6.2	Hörschwelle komplexer Signale	52
6.3	Maskierungshörschwelle und bandbegrenztes Rauschen.....	54
6.4	Two-Tone Masking.....	55
6.5	Phasen- und Modulationseffekte.....	55
6.6	Rauhigkeit und Schwankungsstärke	56
6.7	Unterscheidbarkeit von Teilschwingungen komplexer Signale	58
7	Filterübergreifende und zeitliche Verarbeitungsprozesse	59
7.1	Profilanalyse	59
7.2	Comodulation Masking Release	60
7.3	Modulation Discrimination Interference.....	61
7.4	Auditive Objektwahrnehmung	62
8	Exkurs: Anwendungen des Filtermodells.....	63
8.1	Allgemeine Anwendungsbereiche.....	63
8.2	Mathematische Modelle	63
8.3	Hörgeräte und Cochlea-Implantate	65
8.4	Perzeptive Digitale Audiocodierung	65
9	Abschlussbetrachtung Teil A.....	66
9.1	Anmerkungen zur Literatur.....	66
9.2	Der heutige Stellenwert des Filtermodells.....	67

Teil B: Entwicklung der Multimedia-Anwendung

10	Grundlagen Multimedia-Lehranwendungen.....	69
10.1	Warum Lehren und Lernen mit Software?	69
10.2	Lernparadigmen und Vermittlungskonzepte von Lernsoftware.....	69
10.3	Zum Vorgehen bei Anwendungsplanung und -entwicklung.....	72
11	Anwendungsplanung	73
11.1	Prüfung der Anwendbarkeit einer Multimedia-Lösung	73
11.2	Analyse existierender Anwendungen.....	74
11.3	Zielformulierung und technische Realisierbarkeit	75

12	Anwendungsentwurf und Umsetzung	77
12.1	Inhalt und inhaltliche Struktur.....	77
12.2	Benutzeroberfläche: Layout und Navigation	78
12.3	Die Hilfe-Funktion.....	80
12.4	Hörbeispiele und Hörversuche.....	81
12.4.1	Die verwendeten Audiosamples	81
12.4.2	Zu Entwicklung und Test der Hörversuche	82
12.4.3	Die Hörbeispiele.....	83
12.4.4	Hörversuch: Rauhigkeit.....	84
12.4.5	Hörversuch: Frequenzabhängigkeit der Mithörschwelle	84
12.4.6	Hörversuch: Effektive Maskierungsbandbreite	86
12.5	Einbindung ins WWW	88
13	Abschlussbetrachtung Teil B	88
14	Fazit.....	89
15	Literatur	91
16	Anhang	93
16.1	Designvorgaben: Bildschirmmaske.....	93
16.2	Designvorgaben: Text / Typographie	93
16.3	Designvorgaben: Navigationselemente	94
16.4	Designvorgaben: Abbildungen	94
16.5	Designvorgaben: Text / Inhalt und Stil	94
16.6	Programmstruktur der Hörversuche.....	95
16.6.1	Rauhigkeit.....	95
16.6.2	Effektive Maskierungsbandbreite.....	96
16.6.3	Frequenzabhängigkeit der Mithörschwelle	98
16.7	Inhalte, Verzeichnisse und Dateinamen der Anwendung	99
17	Anlage: CD-ROM.....	102

Abbildungen

Abbildung 1: Grundbegriffe zur Beschreibung akustischer Filter.....	14
Abbildung 2: Peripheres auditives System	15
Abbildung 3: Gehörknöchelchen und Schwingungsübertragung im Mittelohr	16
Abbildung 4: Längsschnitt durch die teilweise „entrollte“ Cochlea.....	18
Abbildung 5: Längsschnitt durch die vollständig „entrollte“ Cochlea	18
Abbildung 6: Längsschnitt der Cochlea, Breite der Basilarmembran	19
Abbildung 7: Cochlea und Cortisches Organ.....	19
Abbildung 8: Schematische Darstellung innerer und äußerer Haarzellen.....	20
Abbildung 9: Anordnung der Haarzellen und Stereozilien	20
Abbildung 10: Auslenkungsmechanismus der Basilarmembran.....	22
Abbildung 11: Auslenkung der Stereozilien	22
Abbildung 12: Augenblickszustand der Wanderwelle auf der Basilarmembran	24
Abbildung 13: Frequenz-Positions-Zuordnung auf der Basilarmembran.....	24
Abbildung 14: Hüllkurven der Basilarmembranschwingung	25
Abbildung 15: Basilarmembranschwingung an verschiedenen Positionen	25
Abbildung 16: Neuronale Frequenz-Tuningkurven (FTC).....	28
Abbildung 17: Abhängigkeit der kritischen Bandbreite von der Frequenz.....	35
Abbildung 18: Neuronale Erregungsmuster.....	40
Abbildung 19: Maskierungsmuster (Masked Audiograms)	41
Abbildung 20: Berechnung des Erregungsmusters für ein 1 kHz - Sinussignal	42
Abbildung 21: Überlagerung neuronaler Erregungsmuster	43
Abbildung 22: Grundstruktur von Modellen der Lautheit	44
Abbildung 23: Psychophysische Frequenz-Tuningkurven.....	45
Abbildung 24: Typische Form eines auditiven Filters	46
Abbildung 25: Abhängigkeit der Filterform vom Schallpegel	47
Abbildung 26: Übersicht über kritische Bandbreiten nach verschiedenen Studien	48
Abbildung 27: Bark-Skala, „Cambridge“-ERBs und 1/3-Oktav-Filter.....	49
Abbildung 28: PTCs für simultane und nicht-simultane Maskierung	50
Abbildung 29: Lautheit eines Bandpass-Rauschens in Abhängigkeit von der Bandbreite.....	52
Abbildung 30: Absolute Hörschwelle für einen Komplex von Sinustönen	53
Abbildung 31: Maskierungshörschwelle und bandbegrenztes Rauschen	54
Abbildung 32: Two-Tone-Masking	55
Abbildung 33: Existenzbereich von Rauigkeit und Schwankungsempfindung	57
Abbildung 34: Unterscheidbarkeit von Teilschwingungen als Maß für kritische Bandbreiten ..	58
Abbildung 35: Mithörschwellen bei unmoduliertem und comoduliertem Rauschen	61
Abbildung 36: Ausschnitt aus der inhaltlichen Struktur der Anwendung	78
Abbildung 37: Die Benutzeroberfläche	79
Abbildung 38: Inhalt der Hilfe-Funktion.....	80
Abbildung 39: Aufruf der Lexikon-Funktion aus dem Text.....	81

Abbildung 40: Einbettung des Hörbeispiels in die Seitendarstellung	83
Abbildung 41: Hörversuch Rauigkeit.....	84
Abbildung 42: Hörversuch zur Frequenzabhängigkeit der Mithörschwelle.....	85
Abbildung 43: Auswertung des Versuchs Frequenzabhängigkeit der Mithörschwelle	86
Abbildung 44: Hörversuch effektive Maskierungsbandbreite.....	87
Abbildung 45: Auswertung des Hörversuchs zur effektiven Maskierungsbandbreite.....	87
Abbildung 46: Link zu externer WWW-Ressource.....	88

Tabellen

Tabelle 1: Lernsoftware.....	71
Tabelle 2: Übersicht über existierende Lehranwendungen	74
Tabelle 3: Schwachpunkte existierender Anwendungen	75

Formeln

Formel 1: Schalldruckpegel L [dB]	13
Formel 2: Verhältnislautheit N [sone].....	13
Formel 3: Äquivalente rechteckige Filterbandbreite (ERB).....	48
Formel 4: Bandbreite-Funktion für Cambridge-ERB.....	49
Formel 5: roex(ρ) - Filter	64

1 Einleitung

1.1 Ziel dieser Arbeit

Der Gegenstand der vorliegenden Arbeit, die Rolle *kritischer Bandbreiten* für die menschliche Hörwahrnehmung, ist von mehreren Besonderheiten geprägt. Die Erforschung des Gehörsinns liegt im Schnittpunkt mehrerer wissenschaftlicher Disziplinen. Der eigentliche Gegenstand ist der direkten Beobachtung überwiegend entzogen: Hörwahrnehmungen sind Teil des individuellen Bewusstseins der Wahrnehmenden und nur über Befragungen und psychologische Tests zugänglich. Physiologische Untersuchungen wiederum stoßen schnell an Grenzen, da die beteiligten Sinnesorgane und physiologischen Vorgänge sehr fein, komplex und zudem verletzlich sind. Diese Umstände haben dazu beigetragen, dass noch viele Zusammenhänge der menschlichen Hörwahrnehmung ungeklärt sind. Das in den 30er Jahren des 20. Jahrhunderts zum ersten Mal beschriebene Phänomen der kritischen Bandbreiten war der Ausgangspunkt für die Entwicklung des heutigen Verständnisses der Filtereigenschaften des peripheren auditiven Systems. Diese sind unverzichtbarer Bestandteil heutiger Modelle der Hörwahrnehmung. Die Kenntnis der Filtereigenschaften wird für viele technische Anwendungen immer wichtiger: Verfahren der digitalen Audiocodierung, Hörhilfen wie Cochlea-Implantate und automatische Sprachverarbeitung bedienen sich filterbasierter Modelle des Gehörs. Doch die unterschiedlichen Forschungsperspektiven der beteiligten Disziplinen und die rasante Entwicklung der Erkenntnisse zur auditiven Wahrnehmung seit den 1930er Jahren führten dazu, dass das Thema überwiegend nur in sehr speziellen Artikeln der Fachzeitschriften behandelt wird. Sekundär- und Überblicksliteratur ist rar und weist nicht selten Lücken und Widersprüche auf, was das Verständnis besonders schwer macht.

Das Ziel dieser Arbeit besteht darin, mit einem systematischen Überblick über den Stand der Forschung zu kritischen Bandbreiten dem angesprochenen Mangel zu begegnen. Die Ergebnisse bilden die Grundlage einer Multimedia-Lehranwendung, die nicht nur dazu dienen soll, die Ergebnisse zugänglich zu machen. Die Möglichkeiten multimedialer Präsentation erscheinen gerade angesichts der behandelten physiologischen Mechanismen, der Effekte der Hörwahrnehmung und typischen Hörversuche besonders geeignet, um die Inhalte besonders anschaulich und verständlich zu gestalten.

1.2 Zur Vorgehensweise

Die zwei Hauptteile dieser Arbeit stehen in engem Zusammenhang, obwohl Teil A für sich inhaltlich abgeschlossen ist. Da dieser Teil jedoch die Grundlage für die inhaltliche Ausgestaltung der Lehranwendung ist, wurden für Teil A bereits im Voraus einige inhaltliche Vorgaben getroffen. Dabei wurde von folgenden Voraussetzungen ausgegangen:

Im Mittelpunkt dieser Arbeit stehen das Phänomen der kritischen Bandbreiten und der Theorie auditiven Filterung. Dabei handelt es sich um ein wissenschaftliches Spezialthema, so dass der Anwendungsbereich einer Lehranwendung von vorneherein beschränkt ist. Angestrebt wird der Einsatz an der Universität durch Studierende in Fachbereichen, die das Thema berühren - z.B. Kommunikationswissenschaften mit naturwissenschaftlicher Ausrichtung, Akustik, Nachrichtentechnik, Psychologie und Medizin. Ein grundlegendes Verständnis für die wissenschaftliche Beschäftigung mit dem Gegenstand kann daher vorausgesetzt werden. Für das Verständnis des Themas sind jedoch neben weiteren Aspekten Grundlagenkenntnisse der Gehörphysiologie Voraussetzung, die über die Grundlagen der genannten Fachbereiche hinausgehen. Des Weiteren sind die zum Thema gehörenden Begriffe vielschichtig und werden z.T. widersprüchlich gebraucht. Die Existenzberechtigung für Multimediale Lehranwendungen leitet sich aber unter anderem daraus ab, dass sie eine *selbständige* Einarbeitung in bestimmte Lehrgebiete ermöglichen (vgl. Abschnitt 10.1). Eines ihrer wichtigsten Merkmale ist die selbsterklärende Verständlichkeit.

Aus diesen Gründen beinhaltet Teil A einen einführenden Teil zur Gehörphysiologie und einen Überblick zur historischen Entwicklung der Forschung und Theoriebildung. Die genannten Vorbedingungen haben auch Einschränkungen zur Folge: Übersichtlichkeit und Verständlichkeit werden im Zweifelsfall Detailreichtum vorgezogen. Daher wird z.B. die Entwicklung im Bereich mathematischer Modelle der auditiven Wahrnehmung nicht in dem Maße berücksichtigt, wie sie die aktuelle Literatur beherrscht. In Teil B werden die gewählte Form der multimedialen Lehranwendung zum Thema theoretisch begründet, Planung und Entwicklung skizziert und die entstandene Anwendung vorgestellt.

1.3 Zu Notation & Terminologie

Die verwendete Literatur lag zum überwiegenden Teil auf Englisch vor. Wo möglich, wurden für Fachbegriffe die in der Literatur gebräuchlichen deutschen Übersetzungen oder eindeutig erscheinende eigene Übersetzungen verwendet, um den Text lesbarer zu gestalten. In Zweifelsfällen wird der englische Begriff in Klammern ergänzt, um einen Vergleich mit der Originalliteratur zu erleichtern. Englische Begriffe und Abkürzungen englischer Begriffe werden nur dann verwendet, wenn sie in der deutschen Literatur überwiegend auch gebraucht werden.

Teil A: Auditive Wahrnehmung und kritische Bandbreiten

2 Begriffe und Grundlagen

2.1 Einordnung des Themengebietes

Der Begriff *kritische Bandbreiten* bezeichnet ein Phänomen der Hörwahrnehmung. In vielen Zusammenhängen ändert sich die Hörwahrnehmung, je nachdem ob die Bandbreite eines Schallsignals über oder unter einem kritischen Wert liegt. Kritische Bandbreiten beruhen auf den *Filtereigenschaften* der Gehörphysiologie. Diese sind die Grundlage für die Theorie der auditiven Filterung, der zufolge das Gehör Schall frequenzabhängig in seine Komponenten zerlegt und in getrennten Kanälen verarbeitet.¹ Damit berührt das Thema Bereiche der Akustik (physikalische Beschreibung von Schallereignissen), der Biologie (Anatomie, Physiologie des Gehörs), Neurologie (Entstehen und Verarbeitung von Nervensignalen) und Psychologie (Interpretation von Wahrnehmungen).

2.2 Psychometrie und Methoden der Psychoakustik

Die Psychoakustik widmet sich den Zusammenhängen zwischen akustischen Reizen und den dadurch hervorgerufenen Hörwahrnehmungen. Sie überbrückt damit gewissermaßen "zwei Welten": die der Physik und die der Wahrnehmung bzw. Sinnesempfindung.² Die in dieser Arbeit beschriebenen psychoakustischen Versuche beruhen auf Messverfahren der *klassischen Psychophysik*, die von G.T. Fechner (1801 -1887) begründet wurde. Es wird davon ausgegangen, dass eine Empfindung durch den auslösenden Reiz gemessen werden kann, indem die - direkt messbare - physikalische Reizstärke mit einer Skala von Empfindungsgrößen in Bezug gebracht wird. Diese Beziehung wird Reiz-Empfindungsfunktion genannt.

Um die Zusammenhänge zwischen Stimuli und Hörempfindungen systematisch zu untersuchen, müssen Hörempfindungen in geeigneter Form gemessen und ausgewertet werden. Die Messung von Hörempfindungen besteht immer in einer Art der Befragung von Versuchspersonen. Praktisch werden psychoakustische Messungen als Laborversuche unter möglichst genau spezifizierten Bedingungen durchgeführt. Dabei werden alle als nicht relevant betrachteten Einflüsse (vor allem Umgebungsgeräusche) möglichst ausgeschlossen. Die meisten psychoakustischen Experimente beschränken sich daher auf eindeutig definierte, einfach zu beschreibende

¹ Vgl. Abschnitt 5.1 zur historische Entwicklung

² Vgl. Terhardt 1998, S. 2

akustische Reize wie z.B. Sinustöne und (bandbegrenzt)es Rauschen. Die wichtigsten Messmethoden der Psychoakustik sind die Bestimmung von Reizschwellen und Unterschiedsschwellen. Die Reizschwelle liegt bei der geringsten Reizgröße, die gerade eben eine Empfindung auslöst (z.B. absolute Hörschwelle). Die Unterschiedschwelle kennzeichnet die Änderung des Reizes gegenüber einem Ausgangsreiz, bei der gerade eben ein Unterschied in der Empfindung wahrnehmbar ist. Diese kleinsten ebenmerklichen Unterschiede (just notable differences, JNDs) bilden die Einheiten der Empfindungsskala. Die Befragung der Versuchspersonen beschränkt sich in der Regel darauf, dass diese aufgefordert werden, eine Entscheidung zwischen oder eine Auswahl unter vorgegebenen Aussagen über die dargebotenen Stimuli zu treffen (z. B. hörbar/nicht hörbar, gleich laut/doppelt/dreifach so laut, etc.). Die Aussage wird meist über Schalter mitgeteilt.

Schwellen zwischen wahrnehmbar und nicht wahrnehmbar sind nicht eindeutig. Zu ihrer Festlegung werden statistische Grenzen gezogen. So wird z.B. die absolute Hörschwelle oftmals bei demjenigen Pegel festgelegt, bei dem eine Versuchsperson einen Testton in 50% der Fälle wahrnimmt und in 50% der Fälle nicht. Die Leistungen einzelner Probanden hängen von Aufmerksamkeit, Ermüdungsprozessen, Übung und nicht zuletzt von der Einstellung und den Erwartungen an den Versuch ab. Bei den Leistungen einzelner Probanden spielen individuell unterschiedliche Erfahrungen und Sinnesleistungen eine Rolle. Daher ist sowohl bei der Planung, als auch bei der Durchführung und der Interpretation der Ergebnisse der Versuche große Sorgfalt angebracht. Um repräsentative Ergebnisse zu erhalten, müssen die Experimente mit einer repräsentativen Stichprobe von Teilnehmern durchgeführt werden.³

Da eine direkte Beobachtung des Gegenstandes "Hörwahrnehmung" nicht möglich ist, wird das menschliche Gehör in psychoakustischen Versuchen oftmals als "Black Box" behandelt. Jedoch wäre es kaum möglich, die beobachteten, ungeheuer komplexen Zusammenhänge zwischen den Eigenschaften des "Input" (Schallsignale) und des "Output" (Hörwahrnehmungen) im Fall des Gehörs auf diese Weise systematisch zu erklären. Für das Verständnis des Gehörs ist es unerlässlich, Erkenntnisse der Psychoakustik mit Kenntnissen der "Signalverarbeitung", d.h. der physiologischen Verarbeitungsmechanismen des Gehörs zu vergleichen.

Um gültige Aussagen über die Funktionsweise des Gehörs treffen zu können, muss sich die Psychoakustik immer auch damit beschäftigen, ob die gemessenen Effekte der Hörwahrnehmung durch die physikalischen Eigenschaften der Schallsignale verursacht werden, oder durch die Einflüsse von Verarbeitungsprozessen durch die Sinnesorgane oder durch zentrale Interpretationsprozesse (im Gehirn).

³ Vgl. Schmidt (2000), S. 219; Hellbrück (1993), S. 177 ff.; Terhardt (1997), S. 11 ff.

Es ist offensichtlich, dass die genannten Laborbedingungen und üblichen Teststimuli nicht den natürlichen Hörgewohnheiten entsprechen. Die allermeisten natürlichen Schallsignale sind weitaus komplexer. In einer natürlichen Hörsituation überlagern sich zudem fast immer komplexe Schallanteile von verschiedenartigsten Quellen. Daher stellt sich stets die Frage, ob bzw. wie sich die im Laborbefunde auf die natürliche Hörwahrnehmung übertragen lassen.⁴ Darüber hinaus liegt (vor allem früheren) Interpretationen von psychoakustischen Experimenten ein Ansatz zugrunde, der als "mikroskopischer" bzw. "bottom-up" Ansatz bezeichnet wird und dessen Anwendbarkeit in vielen Bereichen der Hörwahrnehmung zweifelhaft ist. Dabei wird angenommen, dass sich eine gegebene Wahrnehmung auf irgendeine Art und Weise aus mehreren Empfindungsgrößen zusammensetzt. Daraus wird gefolgert, dass die Schallwahrnehmung insgesamt untersucht werden kann, indem man einzelne Aspekte der Wahrnehmung isoliert betrachtet und ihre Abhängigkeit von der Variation einzelner oder weniger Reizparameter testet. Dieser Ansatz wird dem komplexen Phänomen der Hörwahrnehmung in vielen Fällen erkennbar nicht gerecht (vgl. Abschnitt 9).⁵

2.3 Physikalische Größen und Hörempfindungsgrößen

Im Folgenden spielen eine Reihe von Größen entscheidende Rollen, deren Definition und deren Bezüge zueinander nicht ganz einfach sind. Damit sind vor allem die logarithmische Skala des Schalldruckpegels und die Empfindungsgröße Lautheit gemeint, die neben anderen kurz dargestellt werden. Nicht behandelt werden Grundbegriffe der Akustik wie die Begriffe Schall, Schwingung, Frequenz, Spektrum etc., deren Verständnis vorausgesetzt wird.

Schalldruck und Schalldruckpegel: Der Schalldruck p kennzeichnet die Schallenergie an einem bestimmten Ort innerhalb des Schallfeldes. Die Einheit ist Pascal (Pa). Das Gehör kann enorme Druckunterschiede verarbeiten. Zur besseren Handhabung dieses *Dynamikumfangs* benutzt man daher die logarithmische dB (*Dezibel*) – Skala des *Schalldruckpegels*. Der Schalldruckpegel ist das Verhältnis des herrschenden Schalldrucks zu einem Bezugsschalldruck p_0 . Dies ist der Schalldruck, der im empfindlichsten Frequenzbereich des Gehörs gerade eben eine Empfindung auslöst. Üblich ist der Bezugsschalldruck $p_0 = 20 \mu\text{Pa}$ ($2 \cdot 10^{-5}$ Pa). Der Schalldruckpegel L errechnet sich folgendermaßen:

⁴ Vgl. Luce (1993), S. 196

⁵ Vgl. Hellbrück (1993), S. 180

$$L = 20 \times \log\left(\frac{p}{p_0}\right) \quad [1]$$

L = absoluter Schalldruckpegel [dB]
 p = Schalldruck [Pa]
 p₀ = Bezugsschalldruck 20 μPa (2*10⁻⁵ Pa)

Formel 1: Schalldruckpegel L [dB]

Die Einheit ist dB (Dezibel). Um deutlich zu machen, dass die Skala mit dem genannten Bezugsschalldruck von 20 μPa benutzt wird, wird der Pegel mit dB SPL gekennzeichnet (SPL = Sound Pressure Level, Schalldruckpegel).

Schalleistung und Schallintensität: Zur Beschreibung von Schallquellen werden üblicherweise die Schalleistung und Schallintensität benutzt. Die Schalleistung P [Watt] ist die von einer Schallquelle abgegebene Energie. Die Schallintensität I [W/m²] kennzeichnet die Schallenergie pro Fläche (senkrecht zur Ausbreitungsrichtung der Schallwellen) in einem Schallfeld.

Lautheit bzw. empfundene Lautstärke: Die *physikalische* Größe Schallpegel unterscheidet sich von der subjektiv empfundenen Lautstärke des Schalls. Diese wird als Lautheit bezeichnet (engl. *Loudness*) und ist eine *psychoakustische* Größe. Eine oft genutzte Skala der Lautheit ist die Skala der Verhältnislautheit N nach S.S. Stevens. Ihre Einheit ist sone. Die Lautheit von 1 sone wurde definiert als die empfundene Lautstärke eines Sinustons mit einer Frequenz von 1 kHz bei einem Schallpegel von 40 dB SPL. Die Verhältnislautheit N folgt einem Potenzgesetz (Formel 1). Daraus ergibt sich, dass eine Erhöhung des Schallpegels um 10 dB zu einer Verdoppelung der Lautheit führt.⁶

$$N = k \left(\frac{p}{p_0}\right)^{0,6} \quad [2]$$

N = Verhältnislautheit [sone]
 p = Schalldruck [Pa]
 p₀ = Bezugsschalldruck 20 μPa (2*10⁻⁵ Pa)
 k = Konstante, abhängig von Versuchsbedingungen

Formel 2: Verhältnislautheit N [sone]

⁶ Vgl. Hellbrück (1993), S. 45 ff.

2.4 Akustische Filter

Unter Filtern versteht man in der Akustik Vorrichtungen, die Schwingungen frequenzabhängig mehr oder weniger stark dämpfen. Schall innerhalb eines bestimmten Frequenzbereiches lassen sie passieren. Dieser Bereich wird als Durchlassbereich bezeichnet. Schallanteile außerhalb des Durchlassbereichs werden gedämpft bzw. gesperrt; diese Bereiche werden als Sperrbereich bezeichnet. So genannte Tiefpassfilter lassen Schallanteile bis zu einer oberen Grenzfrequenz passieren und dämpfen Schallanteile mit höheren Frequenzen. Umgekehrt lassen Hochpassfilter Schallanteile bis zu einer unteren Grenzfrequenz passieren und dämpfen darunter liegende Schallanteile. Bandpassfilter (vgl. Abbildung 1) lassen Schallanteile passieren, die innerhalb eines durch eine untere und eine obere Grenzfrequenz definierten Frequenzbands liegen. Sie werden durch ihre Mittenfrequenz und Bandbreite beschrieben.

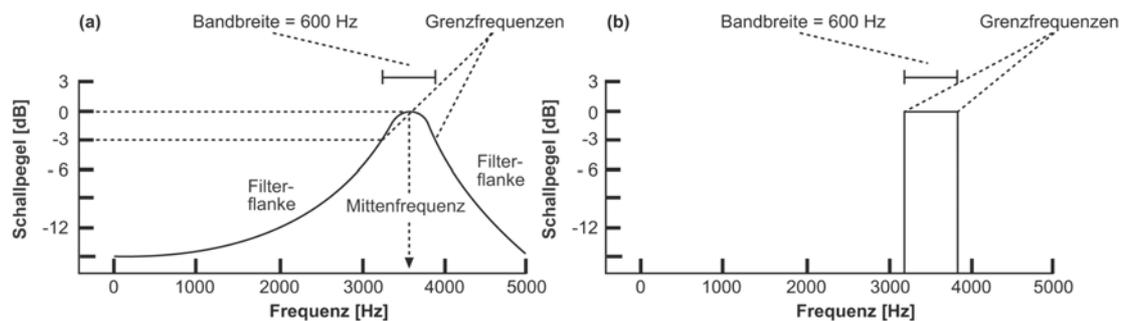


Abbildung 1: Grundbegriffe zur Beschreibung akustischer Filter

(a) ist ein Beispiel für ein reales Filter, (b) ist Beispiel für ein ideales Filter mit rechteckiger Filterfunktion. Die Flankensteilheit idealer Filter ist unendlich.⁷

Die Mittenfrequenz liegt bei der geometrischen Mitte der beiden Grenzfrequenzen (die geometrische Mitte ist die Quadratwurzel aus dem Produkt der beiden Frequenzen). Beidseitig dieser Mittenfrequenz werden die Schallanteile bei allen realen Filtern auch innerhalb des Durchlassbereichs leicht gedämpft. Die Bandbreite des Filters ist die Differenz zwischen den Frequenzen auf beiden Seiten der Mittenfrequenz, bei der die Dämpfung 3 dB beträgt. Ein Filter ist darüber hinaus dadurch gekennzeichnet, wie gut er Schallanteile jenseits des Übertragungsbereichs ausfiltert. Die Filtergüte ist in an der Flankensteilheit des Filters ablesbar. Sie wird in dB pro Oktave angegeben. Je steiler die Flanken jenseits der Grenzfrequenzen abfallen, desto schärfer ist das Filter abgestimmt und desto größer ist seine Filtergüte.

⁷ In Anlehnung an Hellbrück (1993), S. 67

3 Gehöranatomie und -physiologie

3.1 Das auditive System

Aufgabe des *auditiven Systems* (oder *Gehörsystems*) ist die Umsetzung von Schallereignissen in Sinneswahrnehmungen. Druckschwankungen der Luft werden in Signale des Hörnervs umgewandelt, die von den auditiven Zentren des Gehirns als Schallwahrnehmung interpretiert werden. Das Gehörsystem unterteilt sich in *peripheres* und *zentrales* auditives System. Unter dem Begriff peripheres auditives System werden die anatomischen Strukturen des Außenohrs, Mittelohrs und Innenohrs verstanden (siehe Abbildung 2).⁸ Die in der Umgangssprache als „Ohren“ bezeichneten, am Kopf sichtbaren *Ohrmuscheln* spielen dabei eine vergleichsweise geringe Rolle. Das zentrale auditive System umfasst die mit der auditiven Verarbeitung zusammenhängenden Nervenstrukturen: den an der Cochlea beginnenden Hörnerv und die auditiven Verarbeitungszentren im Gehirn. Letztere spielen für diese Arbeit nur am Rande eine Rolle und werden daher nicht behandelt.⁹ Die Begriffe *Gehör* bzw. *Gehörsinn* dagegen werden im folgenden für die Schallwahrnehmungsfähigkeit benutzt. Das beinhaltet die Mechanismen der Schallverarbeitung sowie die Eigenschaften der Schallwahrnehmung. Schwerpunkt der folgenden Abschnitte sind die im Innenohr stattfindenden Prozesse.

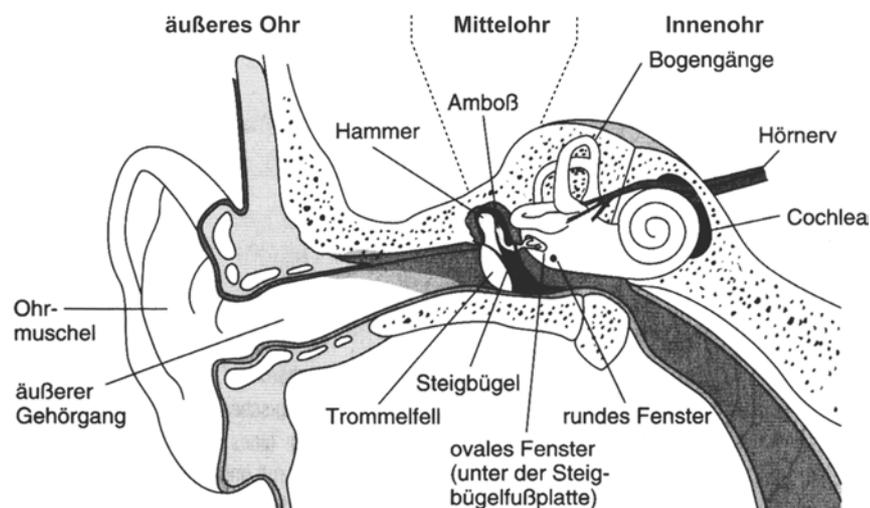


Abbildung 2: Peripheres auditives System

Die Bogengänge beinhalten das Gleichgewichtsorgan, das vom Gehör funktional weitestgehend unabhängig ist.¹⁰

⁸ Auch die Verarbeitung von Schallsignalen im *Innenohr* ist nach dieser Definition Teil der *peripheren* auditiven Verarbeitung.

⁹ Der Hörnerv (VIII. Nerv) wird in der medizinischen Literatur zum peripheren auditiven System gezählt. In dieser Arbeit wird er zum zentralen System gerechnet. Vgl. Zwicker/Fastl (1999).

¹⁰ Aus Goldstein (1997), S. 322

3.2 Das Außenohr

Das Außenohr umfasst die Ohrmuschel (*Pinna*) und den äußeren Gehörgang, der auch als *Ohrkanal* oder *Hörkanal* bezeichnet wird (siehe Abbildung 2). Die Auswirkungen der Ohrmuschel auf das Spektrum spielen eine wichtige Rolle bei der Lokalisation von Schallsignalen. Der etwa 3 cm lange röhrenförmige Gehörgang erfüllt sowohl eine Schutzfunktion für die weiter innen im Kopf liegenden Sinnesorgane, als auch eine akustische Funktion: Durch Resonanzeffekte sorgt der Gehörgang für eine leichte Verstärkung der für die Sprachverständlichkeit wichtigen Frequenzanteile zwischen 2000 und 4000 Hz.

3.3 Das Mittelohr

Am Ende des äußeren Gehörgangs liegt das Trommelfell. Diese Membran überträgt die Schwingungen der Luft an das Mittelohr (siehe Abbildung 2). Dort befinden sich in einem Hohlraum zwischen Trommelfell und Innenohr die Gehörknöchelchen (*Ossikel*). Der direkt am Trommelfell befestigte Hammer (*Malleus*) gibt dessen Schwingungen über den Amboß (*Incus*) an den Steigbügel (*Stapes*) weiter. Die Fußplatte des Steigbügels sitzt auf einer Membran auf, die eine Öffnung des Innenohrs (das *ovale Fenster*) abdeckt. Auf diese Weise werden die Schwingungen an das eigentliche Sinnesorgan, die Hörschnecke (*Cochlea*), weitergeleitet.

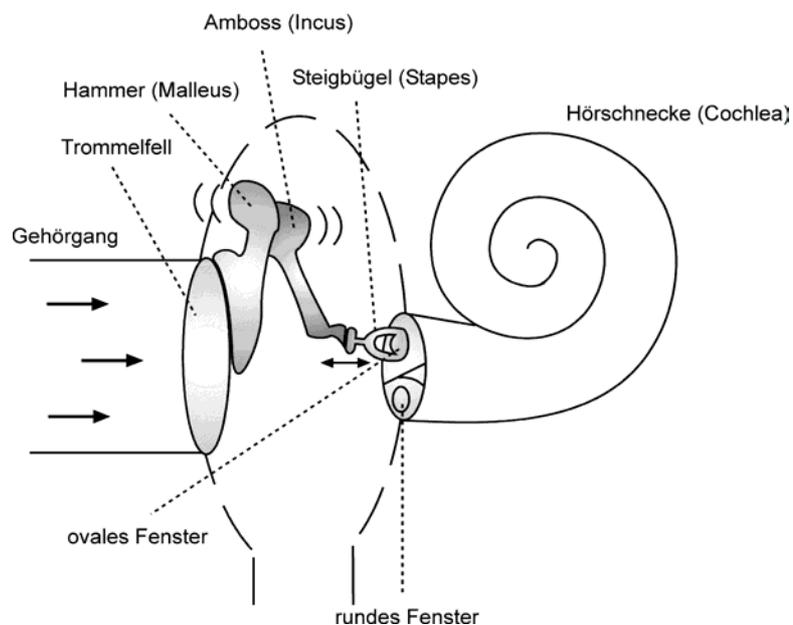


Abbildung 3: Gehörknöchelchen und Schwingungsübertragung im Mittelohr

Die auf das Trommelfell einwirkenden Schallwellen und die Auslenkungen der Gehörknöchelchen sind durch Pfeile und Bewegungslinien gekennzeichnet.¹¹

¹¹ Nach Hellbrück (1993), S. 85

Der Sinn der komplizierten Strukturen des Mittelohrs ergibt sich aus der Tatsache, dass die Cochlea im Gegensatz zum luftgefüllten Außen- und Mittelohr wässrige Flüssigkeiten enthält. Druckschwankungen der Luft können jedoch wegen des Dichteunterschiedes nur schlecht an die Innenohrflüssigkeiten weitergegeben werden. Würden die Schwingungen aus der Luft direkt an diese Flüssigkeiten übertragen, würden nur etwa 3% von ihnen weitergegeben werden. Durch zwei Mechanismen sorgt die Übertragungskette im Mittelohr für die Überwindung dieses Problems: Erstens besteht zwischen der Fläche des Trommelfells (ca. 0,55 cm²) und der Steigbügel Fußplatte (ca. 0,032 cm²) ein erheblicher Größenunterschied im Verhältnis von ca. 17:1. Das erhöht den auf die Fläche der Steigbügel Fußplatte übertragenen Druck. Zweitens wirken durch die Anordnung der Gehörknöchelchen Hebelprinzipien, die die Schwingungen etwa um den Faktor 1,3 verstärken. Der genaue Faktor der Verstärkung variiert in verschiedenen Berechnungen. Unbestritten ist jedoch, dass das Hören ohne das Mittelohr sehr viel schwerer fällt. Bei irreparabler Beschädigung der Gehörknöchelchen wird manchmal eine Operation durchgeführt, nach der der Luftschall direkt ins Innenohr weitergeleitet wird. Damit die Betroffenen genauso hören wie mit Gehörknöchelchen, muss der Schalldruck um einen Faktor zehn bis 50 vergrößert werden.

3.4 Das Innenohr

3.4.1 Die Hörschnecke (Cochlea)

Das Innenohr besteht aus den Bogengängen des Gleichgewichtsorgans und der Hörschnecke (*Cochlea*)(siehe Abbildung 2). Diese beiden Sinnesorgane sind jedoch in ihrer Funktion weitestgehend unabhängig voneinander. Die Cochlea ist eine knöcherne, gewundene Struktur mit etwas mehr als 2 ½ Windungen¹² und ist von besonders hartem Knochengewebe umgeben. Ihre Länge beträgt ca. 35 mm. Ihr Durchmesser beträgt an der Basis ca. 3 mm und verringert sich zur Spitze hin, wo er nur etwa 1,5 mm erreicht (Abbildung 4 und Abbildung 5).¹³ Die Cochlea ist fast durchgängig von drei flüssigkeitsgefüllten Räumen durchzogen: von der Scala Vestibuli, Scala Tympani und Scala Media (siehe Abbildung 7 (a)). Die Scala Media, die sich in der Mitte der Cochlea befindet, ist zur Scala Vestibuli hin durch die Reissner-Membran abgegrenzt, zur Scala Tympani hin durch die Basilarmembran. Die sie ausfüllende Flüssigkeit heißt Endolymphe. Die Scala Media verengt sich zunehmend und endet kurz vor der Spitze der Cochlea, dem Apex.

¹² Es finden sich bei verschiedenen Autoren z.T. unterschiedliche Angaben, z.B. bei Goldstein (1997) 2 ¾ Windungen

¹³ Werte nach Terhardt (1998), S. 57

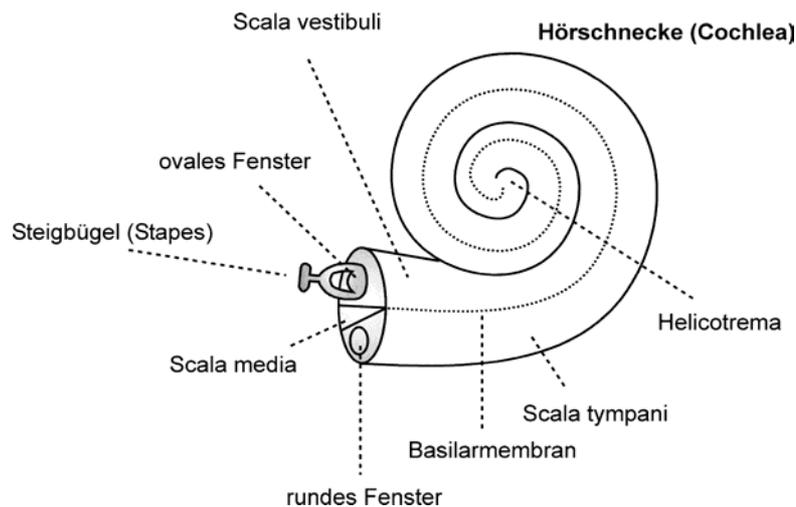


Abbildung 4: Längsschnitt durch die teilweise „entrollte“ Cochlea¹⁴

Dort besteht eine Verbindung zwischen Scala Vestibuli und Scala Tympani, die als Helicotrema bezeichnet wird. Scala Vestibuli und Scala Tympani sind mit Perilymphe gefüllt. Die Scala vestibuli besitzt eine Öffnung zum Mittelohr, das ovale Fenster. Dieses ist durch eine Membran verschlossen, auf die die Fußplatte des Steigbügels aufsetzt. In der Scala Tympani befindet sich ebenfalls eine mit einer Membran verschlossene Öffnung zum Mittelohr hin, das runde Fenster.

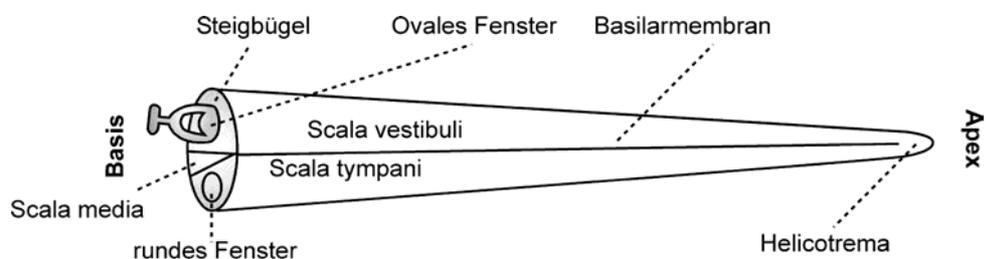


Abbildung 5: Längsschnitt durch die vollständig „entrollte“ Cochlea¹⁵

3.4.2 Basilmembran, Cortisches Organ und Haarzellen

Die Basilmembran ist an der Basis der Cochlea etwa 0,04 mm breit und verbreitert sich bis zu 0,5 mm am Helicotrema (Abbildung 6). Auf ihr befindet sich das cortische Organ (siehe Abbildung 7). Dessen wichtigste Bestandteile sind die Tektorialmembran und die Haarzellen.

¹⁴ Nach Goldstein (1997), S. 324

¹⁵ Nach Goldstein (1997), S. 324

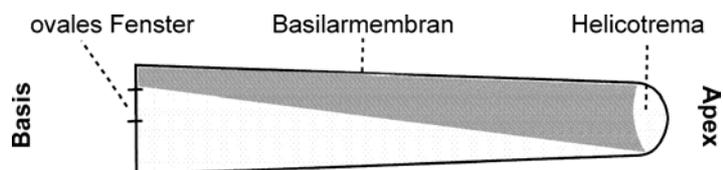


Abbildung 6: Längsschnitt der Cochlea, Breite der Basilarmembran

Die dunkle Fläche kennzeichnet die Basilarmembran. Die Verbreiterung von der Basis bis zum Apex wird deutlich.¹⁶

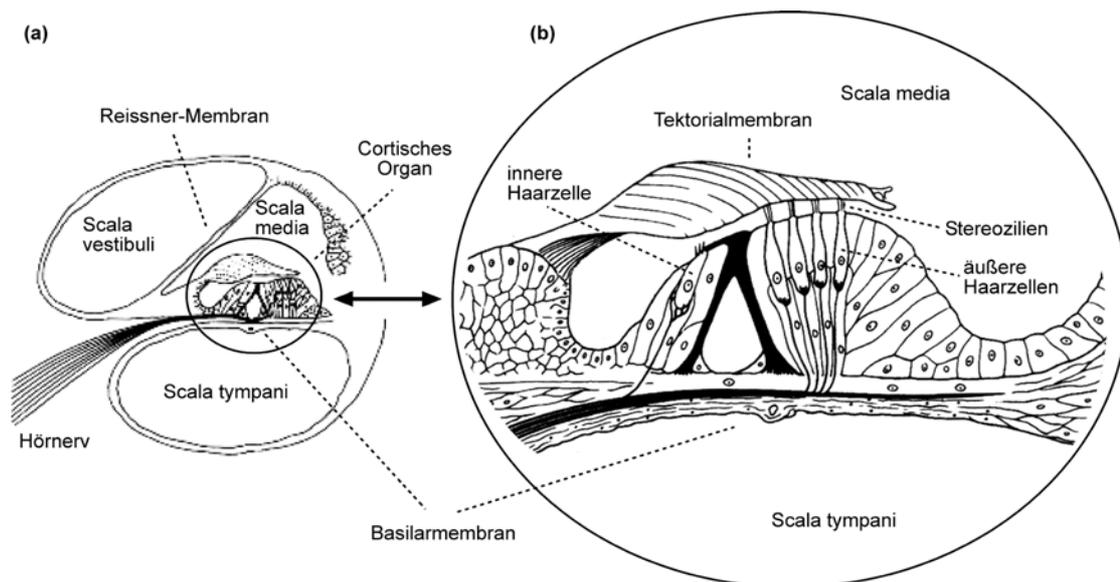


Abbildung 7: Cochlea und Cortisches Organ

(a) Querschnitt durch die Cochlea (b) Ausschnittsvergrößerung Cortisches Organ¹⁷

Die auf der Basilarmembran sitzenden Haarzellen sind die reizaufnehmenden Zellen des Gehörsinns (siehe auch Abschnitt 3.5). Ihre Bezeichnung stammt von den haarartigen Bündeln von Sinneshärchen bzw. Stereozilien, die an ihrem oberen Ende in die Scala Media ragen (siehe Abbildung 7 (b)). Die Haarzellen sind in mehreren Reihen angeordnet. Es gibt zwei Arten von ihnen: Etwa 12.000 äußere Haarzellen, in 3 Reihen angeordnet, sowie etwa 3.500 innere Haarzellen, in einer Reihe (zur Anordnung siehe Abbildung 9).¹⁸ Die je 50-150 Stereozilien der äußeren Haarzellen sind untereinander mit faserigen Strukturen verbunden. Die längsten Stereozilien der äußeren Haarzellen berühren die über der Basilarmembran verlaufende Tektorialmembran; die je etwa 60 Stereozilien der inneren Haarzellen berühren diese jedoch nicht.

¹⁶ Nach Zwicker/Fastl (1999), S. 30

¹⁷ Nach Goldstein (1997), S. 324

¹⁸ Vgl. Warren (1999), S. 9-10. Erstaunlicherweise fanden sich in der verwendeten Literatur neben vielen ähnlichen Werten erhebliche Abweichungen und unklare Angaben.

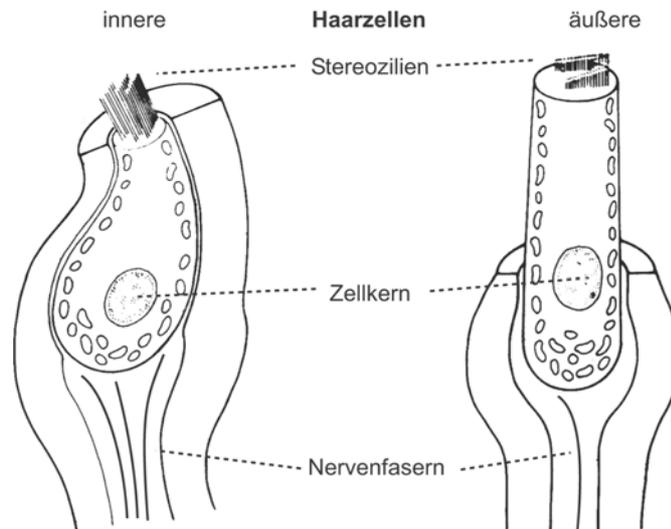


Abbildung 8: Schematische Darstellung innerer und äußerer Haarzellen¹⁹

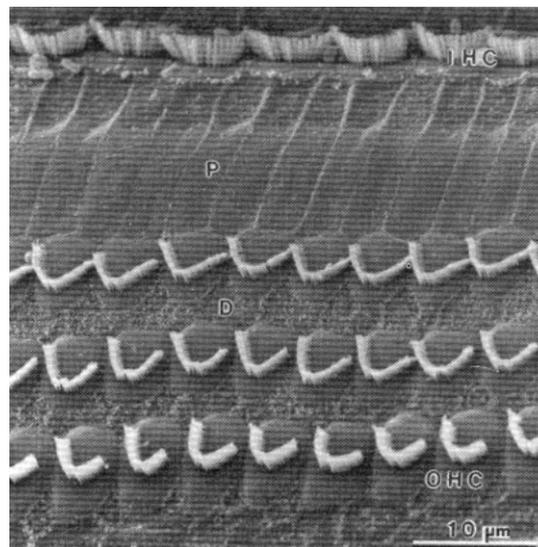


Abbildung 9: Anordnung der Haarzellen und Stereozilien

In dieser Elektronenmikroskopaufnahme sind eine Reihe innerer Haarzellen (oben) und drei Reihen äußerer Haarzellen (unten) sichtbar. Auf den Haarzellen sind die Stereozilien in typischer „W“-Form angeordnet.²⁰

Innere und äußere Haarzellen sind mit den etwa 30.000 Fasern (Axonen) des Hörnervs verbunden, der die Verbindung zu den akustischen Zentren des Gehirns herstellt. Die Mehrzahl ist afferent („aufsteigend“), d.h. die Fasern entspringen in der Cochlea und enden im Cortex. Nur rund 2000 Faserverbindungen sind efferent („absteigend“), d.h. sie entspringen im Cortex und enden in der Cochlea.

¹⁹ Nach Goldstein (1989), S. 420

²⁰ Aus Møller (2000), S. 19

Innere und äußere Haarzellen haben unterschiedliche Funktionen, was sich in ihren Nervenverbindungen widerspiegelt. Jede der inneren Haarzellen ist mit vielen (bis zu ca. 20) afferenten Nervenfasern verbunden (konvergente Innervation). 95 % der Hörnervenfasern erhalten ihre Signale von den inneren Haarzellen, denen somit die Hauptrolle als Sensoren zukommt. Jede afferente Nervenfaser ist *eindeutig* mit *einer* inneren Haarzelle, d.h. einer bestimmten Stelle des Cortischen Organs verbunden. Diese räumliche Zuordnung, die in der Cochlea beginnt (cochleotop bzw. tonotop Organisation, vgl. Abschnitt 4.2) ist im gesamten Hörnerv präsent.

Die efferenten Fasern hingegen verlaufen fast ausschließlich zu den äußeren Haarzellen. Jede Faser ist mit *mehreren* äußeren Haarzellen verbunden (divergente Innervation). Äußere und innere Haarzellen werden durch den gleichen Mechanismus angeregt. Während die inneren Haarzellen daraufhin elektrophysische Signale aussenden (siehe Abschnitt 3.5), reagieren die äußeren Haarzellen vor allem mechanisch.²¹ Sie können im Takt der Anregung ihre Länge und Ausrichtung ändern, wodurch sie das ganze System des Cortischen Organs dynamisch beeinflussen. Die Bewegung der äußeren Haarzellen wirkt sich vor allem in einer Präzisierung der Frequenzanalyse durch die Cochlea aus (vgl. ausführlich in Abschnitt 4.1.2). Zudem sorgt dieser Mechanismus für eine Komprimierung der Signale, d.h. einer Verstärkung bei schwachen und eine Dämpfung bei sehr starken anregenden Reizen. Dies trägt zur Fähigkeit des Gehörs bei, enorme Unterschiede im Schalldruckpegel verarbeiten zu können.²²

3.5 Stimulationsmechanismus des Gehörsinns

Die Haarzellen als Sensoren des Gehörs sprechen auf mechanische Signale an. Sie reagieren auf die Auslenkung der Sinneshärchen an ihrer Spitze mit der Aussendung von Botenstoffen an Nervenfasern, die ihrerseits bioelektrische Impulse an das Nervensystem weiterleiten. Der erforderliche mechanische Reiz wird durch die Schwingungen der Basilarmembran ausgelöst. Diese entstehen auf folgende Weise:

Die Bewegungen des Steigbügels führen zu einer Auslenkung des ovalen Fensters. Dadurch werden Druckwellen in den Flüssigkeiten in der Cochlea ausgelöst. Da diese nahezu unkomprimierbar sind, ist ein Druckausgleich notwendig, der über das runde Fenster erfolgt. Wenn die Membran des ovalen Fensters durch den Steigbügel nach innen in die Scala vestibuli gedrückt wird, wölbt sich das runde Fenster der Scala tympani dementsprechend nach außen ins Mittelohr. Die zwischen Sca-

²¹ Diese Entdeckung wurde erst 1985 gemacht. Vgl. Møller (2000), S. 79

²² Abschnitt 2.2 zu den physiologischen Grundlagen des Gehörs stützt sich insgesamt im wesentlichen auf Schmidt (2000); Goldstein (1997) sowie Møller (2000)

la vestibuli und Scala tympani liegende cochleare Trennwand aus Basilarmembran, Cortischem Organ und Reissner-Membran wird dadurch zunächst nach unten ausgelenkt (Abbildung 10 (a)). Auf den Druck des Steigbügels folgt eine Zugbewegung, in dessen Folge sie wieder nach oben ausgelenkt wird (Abbildung 10 (b)). Auf diese Weise wird die cochleare Trennwand in Schwingungen versetzt, die der Frequenz der ursprünglichen Schallsignale entsprechen (zum Zusammenhang zwischen Frequenz und Schwingung der Basilarmembran siehe Abschnitt 4.1).

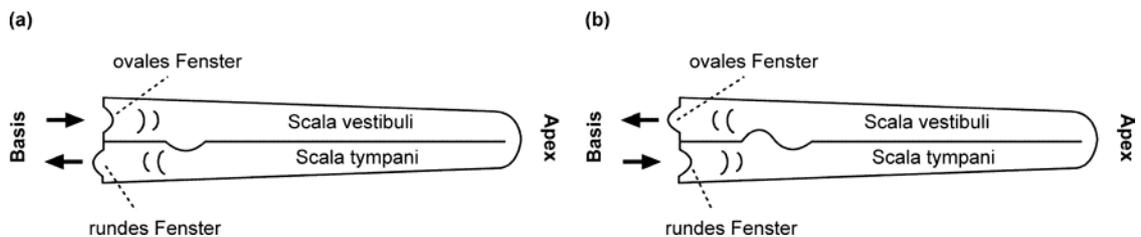


Abbildung 10: Auslenkungsmechanismus der Basilarmembran²³

Das cortische Organ wird dadurch in Relativbewegungen versetzt: Basilarmembran und Tektorialmembran sind so angeordnet, dass sie gegeneinander verschoben werden. Die Stereozilien der auf der Basilarmembran sitzenden Haarzellen sind z.T. an der Tektorialmembran befestigt. Daher werden sie durch die Relativbewegung von Basilarmembran und Tektorialmembran ausgelenkt, z.T. aber auch durch die Druckwellen in der sie umgebenden Endolymphe (Abbildung 11).

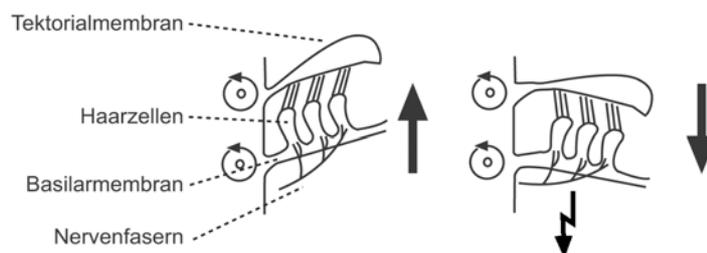


Abbildung 11: Auslenkung der Stereozilien²⁴

Diese Auslenkung ist der auslösende Reiz für die Ausschüttung von chemischen Botenstoffen (Transmittern) an die mit den Haarzellen verbundenen Nervenfasern, die wiederum ein elektrophysiologisches Signal an das zentrale auditive System weiterleiten. Die nötigen Auslenkungsbewegungen sind sehr klein: Eine Sinneswahrnehmung kann bereits durch Bewegungen der Basilarmembran in der Größenord-

²³ Nach Terhardt (1998), S. 233-234

²⁴ Nach Zenner (2000), S. 326

nung von nur 10^{-10} m (das entspricht dem Durchmesser von Wasserstoffatomen) ausgelöst werden.²⁵

4 Die Frequenzanalyse durch das Gehör

Eine der fundamentalsten Eigenschaften des Gehörs ist dessen Frequenzselektivität. Auf ihr beruht die für die Qualität der Hörwahrnehmung bestimmende Größe: die Wahrnehmung der Tonhöhe. Unter Frequenzselektivität versteht man zum einen die Fähigkeit des Gehörs zur *Frequenzunterscheidung*, d.h. zur Wahrnehmung des Frequenzunterschieds zwischen zwei nacheinander präsentierten Tönen. Zum anderen versteht man darunter die Fähigkeit, einzelne Komponenten eines komplexen Signals zu trennen bzw. aufzulösen (*Frequenzauflösungsvermögen*).²⁶ Die von den Sinneszellen in der Cochlea hervorgerufenen elektrophysischen Signale des Hörnervs sind die Informationsquelle, aus der die auditiven Verarbeitungszentren des Gehirns alle Informationen über Schallereignisse beziehen. Die Eigenschaften unserer Hörwahrnehmungen – einschließlich der Tonhöhen – müssen also darin codiert sein. Die Abschnitte 4.1-4.3 beschreiben, welche physiologischen Mechanismen an der psychophysisch messbaren Frequenzselektivität des Gehörs beteiligt sind. Anschließend wird in Abschnitt 4.4 beschrieben, wie diese Mechanismen mit der Repräsentation der Frequenz im neuronalen Code zusammenhängen.

4.1 Die Frequenzselektivität der Basilarmembran

4.1.1 Das Wanderwellen-Prinzip der Basilarmembranschwingung

Für das heutige Verständnis des Hörens waren die Forschungen von Georg von Békésy von elementarer Bedeutung. Er arbeitete in einer Reihe von Studien ab 1928 mit Cochlea-Modellen und -Präparaten und entdeckte so das Wanderwellen-Prinzip der Basilarmembranschwingung.²⁷ Er konnte nachweisen, dass die vom Steigbügel ausgelösten Wellenbewegungen der Basilarmembran von der Basis in Richtung des Apex wandern. Dabei hängt die Strecke, die sie zurücklegen, von der anregenden Schwingungsfrequenz ab. In Abhängigkeit von der Frequenz erreicht die Welle an ganz bestimmten Abschnitten der Basilarmembran ihre maximale Amplitude. Die Aktivität der Haarzellen, die durch diese Bewegung angeregt werden, ist an dieser

²⁵ Vgl. Terhardt (1997), S. 235

²⁶ Neben der Frequenz spielen bei der Tonhöhenempfindung auch andere Faktoren eine Rolle. Vgl. Terhardt (1998), S. 307 ff.

²⁷ Von Békésy bekam für seine Forschungen 1961 den Nobelpreis für Physiologie und Medizin verliehen.

Stelle am stärksten.²⁸ Ein Augenblickszustand einer solchen Wellenbewegung der Basilarmembran ist in Abbildung 12 festgehalten. Es zeigte sich, dass hohe Frequenzen Auslenkungsmaxima an der Basis, tiefe Frequenzen Auslenkungsmaxima näher am Apex erzeugen. Man kann daher der Basilarmembran der Länge nach eine ungefähre Frequenzskala zuordnen (Abbildung 13).

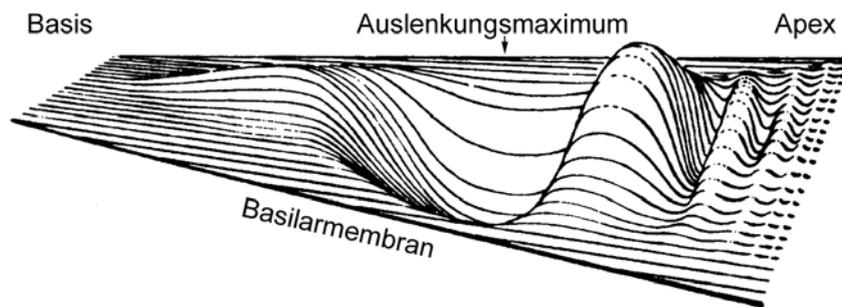


Abbildung 12: Augenblickszustand der Wanderwelle auf der Basilarmembran²⁹

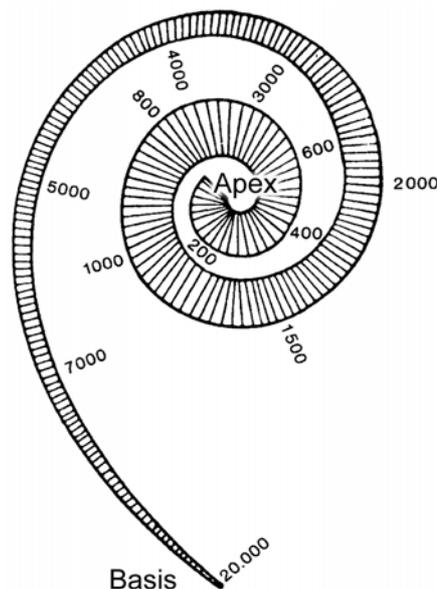


Abbildung 13: Frequenz-Positions-Zuordnung auf der Basilarmembran

Die Zahlen geben die ungefähre Position der maximalen Auslenkung der Basilarmembran für verschiedene Frequenzen an. Das Verhältnis der Breitenänderung der Basilarmembran von der Basis zum Apex ist übertrieben dargestellt.³⁰

²⁸ Die sogenannte *Intensitätskodierung* stellt ein Problem dar, das hier jedoch nicht im einzelnen diskutiert werden kann. Nicht alle Nervenfasern reagieren auf gleich starke Anregung mit gleicher Aktivität. Vgl. Moore (1997), S. 33 ff.

²⁹ Nach Warren (1999), S. 16

³⁰ Nach Warren (1999), S. 11

Hüllkurven geben Aufschluss über die Abhängigkeit der Form der Wellenbewegung von der Frequenz. Sie können ähnlich wie Abbildung 12 als Augenblickszustände der Wellenbewegung angesehen werden. Dabei wird nur der Betrag, aber nicht die Richtung der Auslenkung wiedergegeben. Ermittelt man die Auslenkung der Wellenbewegung entlang der Basilarmembran für bestimmte Frequenzen, erhält man Hüllkurven wie in Abbildung 14. Man erkennt, dass bei einer bestimmten anregenden Frequenz die Basilarmembran zwischen Basis und der Position der maximalen Auslenkung in geringerem Maße mitschwingt. Über die Position der maximalen Auslenkung hinaus fällt die Amplitude der Schwingung schnell auf nahezu Null.

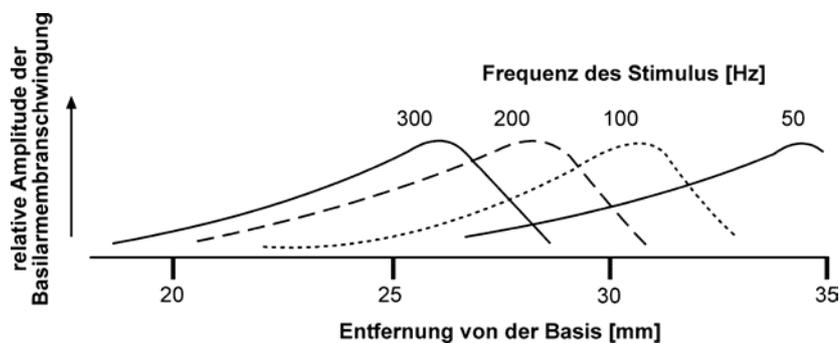


Abbildung 14: Hüllkurven der Basilarmembranschwingung

Die Kurven geben die relative Amplitude der Schwingungen entlang eines Abschnitts der Basilarmembran für verschiedene Sinustöne wieder.³¹

Beobachtet man dagegen eine bestimmte Position der Basilarmembran und ermittelt die Auslenkung in Abhängigkeit von verschiedenen Frequenzen, erhält man Hüllkurven wie in Abbildung 15. An bestimmten Positionen schwingt die Basilarmembran maximal bei bestimmten Frequenzen. Über der Frequenz des Maximums fällt die Amplitude der Schwingung rasch ab. Bei Frequenzen unterhalb des Maximums fällt die Amplitude deutlich weniger steil ab.

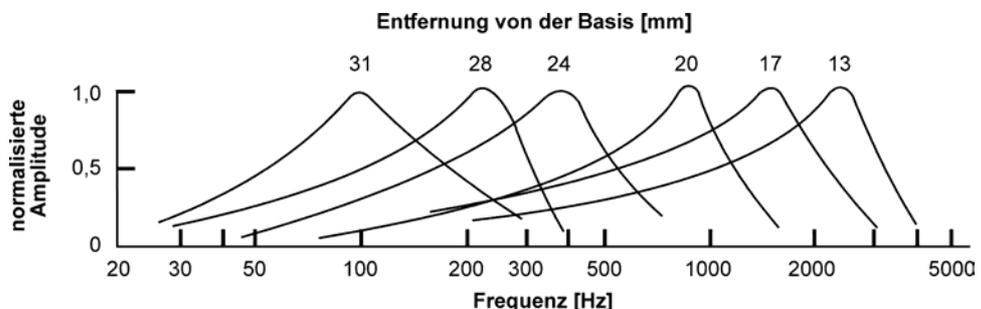


Abbildung 15: Basilarmembranschwingung an verschiedenen Positionen

Die Kurven geben die relative Amplitude (normalisiert) für die Schwingung an verschiedenen Positionen der Basilarmembran in Abhängigkeit von der anregenden Frequenz wieder.³²

³¹ Nach Stevens (1998), S. 208

4.1.2 Die Funktion der äußeren Haarzellen

Von Békésy forschte in den 30er Jahren des 20. Jahrhunderts mit aus heutiger Sicht sehr beschränkten Mitteln. Er führte Beobachtungen an Modellen und mikroskopische Beobachtungen an Leichenpräparaten durch. Um dabei überhaupt eine Auslenkung der Basilarmembran beobachten zu können, waren extreme Schallintensitäten nötig. Von Békésy beobachtete auf diese Weise, dass die verwendeten Schallsignale breite Abschnitte der Basilarmembran zum Schwingen brachten. Er nahm an, dass dieses Schwingungsverhalten linear sei. Daher wurde aus den Ergebnissen die Frequenzselektivität bei geringen Schallintensitäten extrapoliert. Von Békésy selbst war daher der Ansicht, dass der Wanderwellen-Mechanismus allein nicht das psychophysisch messbare, exakte Frequenzauflösungsvermögen des Gehörs erklären könne. In den 70er Jahren nahm man daher an, dass ein zweites Filter wirksam sein müsse, das die Frequenzselektivität der Basilarmembran schärft.³³

Mittlerweile existiert jedoch eine Reihe von weit genaueren Methoden zur direkten Beobachtung der Basilarmembran. Häufig wird die so genannte Mößbauer-Methode angewendet. Dabei wird radioaktives Material auf der Basilarmembran platziert. Deren Schwingungsverhalten lässt sich dann anhand der Auswirkungen des Doppler-Effekts auf die abgegebene Strahlung untersuchen.³⁴ Neuere Untersuchungen ergaben, dass die Frequenzselektivität der Basilarmembran weit präziser ist, als von Békésy annehmen konnte. Sie ist zudem nicht linear, sondern hängt von der Amplitude ab und ist bei niedrigen Intensitäten größer als bei hohen. Des Weiteren ist die Frequenzselektivität im hohen Maße vom physiologischen Zustand der Cochlea abhängig, von der Zufuhr von Blut, Sauerstoff und toxischen Substanzen, und unterscheidet sich daher stark zwischen lebenden Organismen und den Leichenpräparaten, die von Békésy untersuchte.³⁵ Dies ist vor allem auf die lange unerkannte, aktive Rolle der äußeren Haarzellen zurückzuführen.

Erst 1985 wurde die Entdeckung gemacht, dass die äußeren Haarzellen eine aktive Rolle spielen und dadurch die Frequenzselektivität der Basilarmembran entscheidend verstärken. Während die inneren Haarzellen als Signalwandler und –geber für den Hörnerv fungieren, scheinen die äußeren Haarzellen vor allem mechanisch auf die Basilarmembran selbst einzuwirken. Sie reagieren auf die Anregung ihrer Stereozilien mit einer aktiven Beeinflussung der Bewegung des Systems.³⁶ Mittlerweile

³² Nach Stevens (1998), S. 208

³³ Vgl. Evans (1992)

³⁴ Eine Übersicht zu den Methoden findet sich bei Terhardt (1998), S. 232

³⁵ Vgl. Terhardt (1998), S. 232

³⁶ Vgl. Møller (2000), S. 79

wird angenommen, dass die Frequenzselektivität der Basilarmembran dadurch ebenso präzise ist wie die feine Frequenzabstimmung einzelner Hörnervenfasern (siehe Abschnitt 4.2). Zudem kompensiert die Bewegung der äußeren Haarzellen bei schwachen Signalen durch positive Rückkopplung Energieverluste und sorgt für eine niedrigere Wahrnehmungsschwelle. Bei großen Amplituden wirkt der Mechanismus dagegen begrenzend.³⁷

4.2 Die Frequenzselektivität der Hörnervenfasern

Es ist möglich, die elektrische Aktivität einzelner Nervenfasern als Antwort auf bestimmte anregende Schallsignale direkt zu messen. Diese Methode ist technisch sogar einfacher zu realisieren als die direkte Beobachtung der Schwingungen der Basilarmembran.³⁸ Die Messung wird im Tierversuch mittels feiner Elektroden durchgeführt. Hörnervenfasern zeigen auch ohne externe Anregung ständig spontane Entladungen nach zufälligem Muster. Ab einem bestimmten Signalpegel wird jedoch eine Erregungsschwelle überschritten, was zu einer messbaren Änderung der Entladungsrate führt. Die grafische Wiedergabe dieser Änderungsschwelle der Entladungsrate in Abhängigkeit von der Frequenz des anregenden Signals wird als neuronale Frequenz-Tuningkurve (*Frequency Tuning Curve, FTC*) bezeichnet (vgl. Abbildung 16).

Frequenz-Tuningkurven zeigen, dass die Reaktion der Hörnervenfasern auf sehr enge Frequenzbereiche abgestimmt ist. Die Frequenz, bei der die Erregungsschwelle am niedrigsten ist bzw. bei der die Faser bei fester Amplitude eine maximale Impulsfrequenz erreicht, wird als charakteristische Frequenz oder Bestfrequenz der Nervenfaser bezeichnet. Bei zunehmender Schallintensität reagiert die Nervenfaser auf eine zunehmend größere Bandbreite von Frequenzen. Einzelne Nervenfasern verhalten sich somit wie (nicht-lineare) Filter, da sie Eingangssignale frequenzabhängig übertragen. Wie in Abbildung 16 zu erkennen ist, unterscheiden sich Frequenz-Tuningkurven in Abhängigkeit von ihrer charakteristischen Frequenz in der Form. Frequenz-Tuningkurven von Nervenfasern, die auf hohe Frequenzen abgestimmt sind, haben eine asymmetrische Form, mit sehr steiler Flanke in Richtung höherer und einer wesentlich flacheren Flanke zu tiefen Frequenzen hin. Frequenz-Tuningkurven von Nervenfasern mit niedrigerer charakteristischer Frequenz haben eine symmetrischere Form.

³⁷ Für die Existenz aktiver Mechanismen spricht nicht zuletzt die Tatsache, dass das Innenohr selbst messbaren Schall abgeben kann, sogenannte otoakustische Emissionen. Vgl. Møller (2000), S. 80, S. 159; Terhardt (1998), S. 237-239; Zwicker/Fastl (1999), S. 35 ff.

³⁸ Vgl. Møller (2000), S. 157

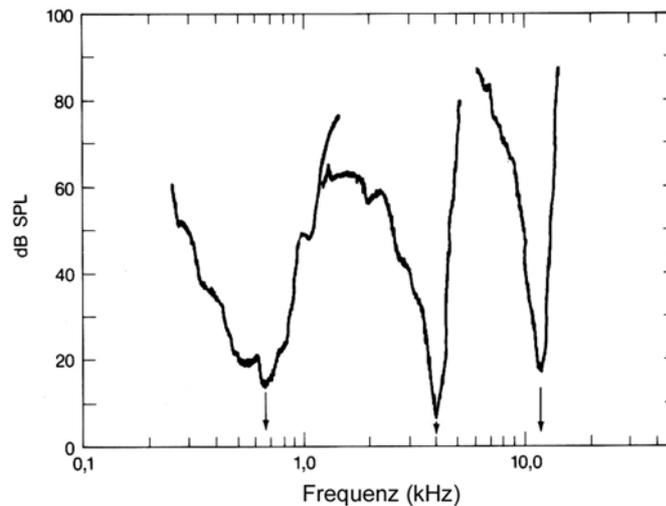


Abbildung 16: Neuronale Frequenz-Tuningkurven (FTC)

Tuningkurven für drei verschiedene Nervenfasern eines Tieres. Die Pfeile markieren die jeweilige charakteristische Frequenz.³⁹

Die Messung der elektrophysischen Aktivität ist an verschiedenen Stellen des auditiven Nervensystems möglich. Es wurde nachgewiesen, dass nicht nur die direkt von den Haarzellen ausgehenden Nervenfasern, sondern auch Fasern an zentraler gelegenen Stellen des Hörnervs ähnliche Tuningkurven zeigen. Wie im Cortischen Organ sind die auf benachbarte Frequenzen gestimmten Nervenfasern im gesamten auditiven Nervensystem benachbart angeordnet. Dies wird als tonotope bzw. cochleo-*tope* Organisation bezeichnet.⁴⁰

4.3 Phasenkopplung der neuronalen Entladungen

Die Frequenzselektivität der Basilarmembran und die durchgängig tonotope Organisation des auditiven Nervensystems bedeuten, dass im peripheren auditiven System eine Umwandlung der Frequenz in ortsbezogene Informationen stattfindet (sogen. *Frequenz-Orts-Transformation*). Eine ebenso wichtige Rolle spielen daneben zeitliche Mechanismen der Frequenzanalyse. Die elektrophysischen Impulse einzelner Haarzellen (vgl. Abschnitt 3.5) erfolgen mit großer Präzision auf die Phase der Basilarmembranschwingung abgestimmt. Damit folgt die Frequenz der Impulse direkt der Frequenz der anregenden Schwingungen. Dies wird als Phasenkopplung (*Phase Locking*) bezeichnet. Dieser Mechanismus unterliegt einer bedeutenden Einschränkung: Die Auslenkung der Stereozilien führt zu einer Depolarisierung in der Haarzelle. Auf diese Entladung folgt zwingend eine Erholungsphase (*Refraktärzeit*), wodurch die Geschwindigkeit der Impulsfolge und somit das zeitliche Auflösungsvermögen einzel-

³⁹ Aus Goldstein (1997), S. 331

⁴⁰ Vgl. Goldstein (1997), S. 151 ff.

ner Nervenfasern auf die Größenordnung von Millisekunden begrenzt ist. Damit liegt die Obergrenze der direkten Übertragung der Frequenz durch Phasenkopplung bei 800-1000 Hz.

4.4 Frequenzcodierung

Wie lassen sich nun das Frequenzunterscheidungsvermögen und das Frequenzauflösungsvermögen des Gehörs anhand der in der Cochlea erzeugten neuronalen Signale erklären? Auf der Grundlage der in den Abschnitten 4.1-4.2 beschriebenen tonotopen Organisation von Basilarmembran und Hörnerv einerseits sowie der in Abschnitt 4.3 beschriebenen Phasenkopplung der Entladungen der Sinneszellen andererseits kommen zwei unterschiedliche Prinzipien in Betracht:

1. *Die Frequenz wird örtlich codiert (Ortstheorie, Ortsprinzip):*
Basis ist die Frequenzselektivität der Basilarmembran. Dort findet eine Frequenz-Ortstransformation statt. Die Information über den Ort der ursprünglichen Stimulation wird durch das gesamte Hörnervensystem bewahrt (tonotope bzw. cochleotopische Organisation); der Ort der ursprünglichen Stimulation signalisiert daher die Frequenz.
2. *Die Frequenz wird zeitlich codiert (Zeittheorie, Periodizitätstheorie)*
Haarzellen werden in Abhängigkeit von der Phase der anregenden Signalkomponenten ausgelenkt. Die Entladungen der Hörnervenfaser sind dadurch synchron mit der Phase (Phasenkopplung). Die zeitlichen Entladungsmuster der Neuronen signalisieren daher die Frequenz.

In der Hörwahrnehmungsforschung war lange umstritten, wie diese Mechanismen mit der Frequenzselektivität des Gehörs zusammenhängen.⁴¹ Wichtigste Grundlage für die Diskussion dieses Problems ist die Analyse der auf verschiedenen Ebenen des Hörnervs nachweisbaren neuronalen Entladungsmuster. Die Cochlea liefert als Ausgangssignal einen Code, der *beide* Repräsentationen der Frequenz enthält, d.h. sowohl Informationen über das Energiespektrum als auch über die Periodizität bzw. Wellenform von Schallsignalen. Um für die Frequenzselektivität des Gehörs in Frage zu kommen, muss ein Codierungsmechanismus folgenden Ansprüchen genügen: Er muss zum einen unabhängig von den weiteren Eigenschaften des Signals sein (z.B. Intensität), zum anderen muss die Information tatsächlich auf dem gesamten Signalweg zu den Zentren der auditiven Verarbeitung in geeigneter Form erhalten bleiben.⁴²

⁴¹ Sowohl im Bereich der Frequenzselektivität der Basilarmembran als auch bei der Untersuchung der Phasenkopplung hat die Forschung rasche Fortschritte gemacht, die dazu führten, dass sich in der Literatur zahlreiche widersprüchliche Angaben zu diesem Thema finden. Die folgenden Ausführungen stützen sich auf die aktuellsten verfügbaren Bücher mit dem Schwerpunkt Sinnesphysiologie: Warren (1999); Møller (2000)

⁴² Für eine zusammenfassende Diskussion vgl. Møller (2000), S. 181-196

4.4.1 Örtliche Codierung

Die präzise Frequenzselektivität der Basilarmembran sowie die cochleotop Organisations in allen Teilen des Hörnervensystems sind das hervorstechende Argument für das Ortsprinzip der Frequenzcodierung. Nachdem lange Zeit die Genauigkeit der Frequenzselektivität der Basilarmembran nicht ausreichend erschien, um das Frequenzanalysevermögen des Gehörs zu erklären, scheint zwar dieser Einwand nach neueren Erkenntnissen widerlegt zu sein (vgl. Abschnitt 4.1). Das entscheidende Argument gegen eine Frequenzcodierung gemäß dem Ortsprinzip der Basilarmembranschwingung ist jedoch deren Abhängigkeit von der Stimulusintensität. Bei großen Intensitäten ist eine beträchtliche Verschiebung des Schwingungsmaximums in Richtung der Basis der Cochlea zu beobachten.⁴³ Daneben bestehen weitere Einwände: So existieren z.B. Befunde bei Patienten, die eine Veränderung des Schwingungsverhaltens der Basilarmembran zeigten. Während das veränderte Schwingungsverhalten eine Frequenzcodierung nach dem Ortsprinzip unmöglich machen würde, schien ihr Sprachwahrnehmungsvermögen dennoch nicht beeinträchtigt.⁴⁴ Dazu kommt ein Phänomen aus der Psychoakustik: Die Tonhöhe komplexer Klänge wird von Versuchspersonen auch dann wahrgenommen, wenn die eigentliche Grundfrequenz fehlt.⁴⁵ Die Ortstheorie der Frequenzcodierung kommt daher als Erklärung für das präzise Frequenzunterscheidungsvermögen des Gehörs nicht in Betracht.

4.4.2 Zeitliche Codierung

Grundlage für die Hypothese der zeitlichen Codierung ist die Kopplung der neuronalen Impulse mit der Wellenform des anregenden Signals. Im Gegensatz zum Ortsprinzip der Basilarmembranschwingung scheint die Phasenkopplung weitgehend intensitätsunabhängig zu sein.⁴⁶ Gegen die Hypothese der zeitlichen Codierung werden vor allem zwei Einwände angeführt: Zum einen das beschränkte zeitliche Auflösungsvermögen von Nervenfasern aufgrund ihrer elektrischen Eigenschaften (vgl. Abschnitt 3.5); zum anderen erscheint es zweifelhaft, dass zeitliche Informationen die neuronale Weiterleitung an höhere Verarbeitungszentren ohne Präzisionsverluste überstehen. In zentraleren Regionen des auditiven Nervensystems scheint die Frequenzobergrenze für Phase Locking immer weiter abzusinken.

Møller (2000) weist jedoch auf Unzulänglichkeiten früherer Untersuchungsansätze hin. In vielen Studien wurde nur das Verhalten einzelner Nervenfasern unter-

⁴³ Vgl. Møller (2000), S. 191

⁴⁴ Ebd., S. 192-193

⁴⁵ Vgl. Warren (1999) S. 66

⁴⁶ Vgl. Møller (2000), S. 187

sucht. Doch während deren Phasenpräzision in höheren Nervenbahnen abnimmt, scheint dort die Fähigkeit zur Integration der Informationen mehrerer Fasern zuzunehmen. Für die prinzipielle Fähigkeit des auditiven Systems zu sehr präziser Verarbeitung zeitlicher Informationen spricht die Fähigkeit der Lokalisation von Schallquellen, die auf z.T. sehr geringen Laufzeitunterschieden beruht.⁴⁷

Ein höheres zeitliches Auflösungsvermögen lässt sich mit der *Salventheorie* erklären.⁴⁸ Bei Frequenzen über der höchsten Übertragungsfrequenz einzelner Nervenfasern feuern diese demnach nicht bei jeder Periode der anregenden Schwingung, sondern in gleichmäßigen Intervallen unter Auslassung einzelner Perioden. Andere Nervenfasern feuern bei anderen Perioden, so dass eine Gruppe von Nervenfasern insgesamt die hohe Frequenz der anregenden Schwingung übertragen kann. Die Fähigkeit zu Phasenkopplung, die auch für die Salventheorie eine unabdingbare Voraussetzung ist, scheint jedoch beim Menschen eine Obergrenze von etwa 5 kHz zu haben. Darüber ist die Wahrscheinlichkeit, dass einzelne Neuronen feuern, für jede Phase der anregenden Schwingung gleich. Hartmann (1997) wertet dies als eine Stärke des Modells der zeitlichen Codierung, da auch die psychophysisch messbare Fähigkeit zur Tonhöhenunterscheidung in diesem Bereich nachzulassen scheint.⁴⁹

Für eine entscheidende Rolle zeitlicher Mechanismen für das Frequenzunterscheidungsvermögen vor allem im Zusammenhang mit komplexen Schallsignalen sprechen auch Befunde im Zusammenhang mit physischen Schädigungen des Gehörs. Schädigungen des Hörnervs, die gleichbedeutend mit einer Schädigung des Nervenleitungsvermögens und damit der Übertragungsfähigkeit des zeitlichen Codes sind, führen zu größeren Verlusten des Sprachunterscheidungsvermögens als Schädigungen der Cochlea, selbst wenn die Verletzungen eine vergleichbare Anhebung der Hörschwelle zur Folge hatten. Die Sprachwahrnehmung ist in solchen Fällen immer stärker eingeschränkt als Audiogramme vermuten lassen, die mit reinen Sinustönen durchgeführt wurden.⁵⁰

4.5 Die auditive Vorverarbeitung durch die Cochlea

In der Cochlea findet eine präzise Frequenzanalyse von Schallsignalen durch die Frequenzselektivität der Basilarmembranschwingung statt. Als Grundlage für die Frequenzcodierung kommt deren Ortsprinzip wegen seiner Intensitätsabhängigkeit

⁴⁷ Vgl. Møller (2000), S. 189-190; Palmer (1995), S. 85

⁴⁸ Vgl. Møller (2000), S. 195

⁴⁹ Vgl. Hartmann (1997), S. 293; Palmer (1995), S. 85; Terhardt (1998), S. 236

⁵⁰ Vgl. Møller (2000), S. 193

jedoch nicht in Frage.⁵¹ Zeitliche Mechanismen scheinen für das Frequenzunterscheidungsvermögen des Gehörs die wichtigere Rolle zu spielen. Angesichts der Eigenschaften der in den Abschnitten 4.1 bis 4.4 diskutierten Mechanismen ist allerdings naheliegend anzunehmen, dass die Frequenzselektivität des Gehörs sowohl auf örtlicher als auch zeitlicher Codierung beruht. Bereits Ende der 40er Jahre wurde vermutet, dass zeitliche Codierung bei niedrigen Frequenzen, örtliche Codierung bei höheren am wichtigsten sei. Im mittleren Frequenzbereich würden beide Prinzipien wirksam sein. Diese Hypothese wurde als *Volley Theory* bekannt.⁵² In der neueren Literatur herrscht die Ansicht vor, dass die spektralen Komponenten komplexer Schallsignale durch die Frequenzselektivität der Basilarmembran in schmalbandige Komponenten aufgeteilt werden, bevor die zeitliche Struktur dieser Komponenten in den Entladungsmustern von verschiedenen Populationen von Nervenfasern codiert wird. Aus physiologischer Sicht scheint dies plausibel, da die Aufteilung in Frequenzbänder eine anschließende zeitliche Codierung erleichtern würde.⁵³

5 Auditive Filter und kritische Bandbreiten

5.1 Historische Entwicklung

5.1.1 Entdeckung des kritische Bandbreiten-Phänomens

In den 30er Jahren des 20. Jahrhunderts beschrieb *Harvey Fletcher* erstmals das Phänomen der kritischen Bandbreiten.⁵⁴ Er widmete sich neben Arbeiten zur erstmaligen systematischen Bestimmung der absoluten Hörschwelle und zur Sprachverständlichkeit intensiv der Lautstärkeempfindung. In Veröffentlichungen in diesem Zusammenhang wurde das Auftreten kritischer Bandbreiten 1933 das erste Mal erwähnt.⁵⁵ Fletcher stellte in zahlreichen psychoakustischen Experimenten fest, dass sich die Hörwahrnehmung unterscheidet, je nachdem ob die Bandbreite der dargebotenen Stimuli über oder unter einem kritischen Wert liegt. Das Phänomen äußert sich bei der Lautstärkeempfindung folgendermaßen: Liegen die Frequenzen der Kompo-

⁵¹ Vgl. Møller (2000), S. 183

⁵² Vgl. Møller (2000), S. 195

⁵³ Vgl. Møller (2000), S. 195 ("duplex hypothesis of frequency discrimination"); Yates (1995), S. 43-44 ("Acoustic Preprocessing"); Terhardt 1998, S. 241-242 ("Serien-Parallel-Umsetzung der Schallsignale"). Für eine ausführliche Diskussion der Frequenzcodierung im auditiven System vgl. Møller (2000), S. 151 ff.

⁵⁴ Fletcher leitete das Acoustic Research Department der Bell Laboratories der Telefongesellschaft AT&T, wo umfangreiche Grundlagenforschungen zu Sprache und Hören durchgeführt wurden. Die Erkenntnisse flossen in die Entwicklung von technischen Übertragungssystemen für Sprache und Musik ein, die u.a. in der Form des Telefonsystems und Funk bzw. Radio in der ersten Hälfte des 20. Jahrhunderts einen Boom erlebten.

⁵⁵ Fletcher/Munson (1933a); Fletcher/Munson (1933b), zit. nach Allen (1995)

nenten eines Schallsignals weiter als eine kritische Bandbreite voneinander entfernt, entspricht die empfundene Gesamtlautstärke des Signals der Summe der *Lautheiten*, d.h. der *empfundenen* Lautstärken, der einzelnen Komponenten. Liegen die Komponenten des Signals innerhalb der kritischen Bandbreite, folgt die empfundene Gesamtlautstärke der Summe von deren physikalischen *Intensitäten* (vgl. Abschnitt 6.1). Des Weiteren spielte die kritische Bandbreite eine entscheidende Rolle bei Maskierungseffekten (vgl. Abschnitt 6.3). Fletcher stellte fest, dass in einem Breitbandrauschen offenbar nur ein schmales Frequenzband um einen Testton dessen Hörbarkeit effektiv verschlechtert. In den 30er Jahren entwickelten Fletcher und sein Kollege Munson auf dieser Grundlage ihr Konzept der kritischen Bandbreiten. Die entscheidenden Veröffentlichungen der Ergebnisse erschienen 1938.⁵⁶

Fletcher schlug ein *Filtermodell des Gehörs* vor, das auf seinen experimentellen Befunden und dem damaligen Kenntnisstand der Gehörphysiologie beruhte: Das periphere auditive System verhält sich wie eine Reihe paralleler, überlappender Bandpassfilter. Die Filter bilden ein Kontinuum entlang der Basilarmembran und beruhen auf deren Frequenzselektivität: Jeder Ort auf der Basilarmembran reagiert auf ein schmales Frequenzband mit jeweils unterschiedlicher Mittenfrequenz.⁵⁷ Die Filter fassen die Schallenergie in Frequenzbändern zusammen. Die gemessenen kritischen Bandbreiten entsprechen der Bandbreite dieser Filter. Fletcher benutzte dabei eine Vereinfachung: Er rechnete mit Bandbreiten idealer Bandpassfilter mit rechteckiger Übertragungsfunktion. Solche Werte werden heute als äquivalente rechteckige Bandbreiten (*Equivalent Rectangular Bandwidths, ERBs*) bezeichnet (vgl. Abschnitt 5.3.2). Fletcher nahm nicht an, dass die *realen* Filter rechteckig seien. Dieser Eindruck kann jedoch entstehen und hat in der Diskussion über kritische Bandbreiten mitunter für Verwirrung gesorgt.⁵⁸

5.1.2 *Energiespektrum-Modell und Critical Ratio*

Fletchers Filtermodell des Gehörs beruhte vor allem auf Maskierungsexperimenten. Bei solchen Experimenten geht darum, unter welchen Bedingungen Komponenten eines Schalls durch andere Komponenten *verdeckt* werden, d.h. unter welchen Bedingungen sie nicht mehr oder weniger deutlich wahrnehmbar sind (vgl. Ab-

⁵⁶ Diese Arbeiten von Fletcher (1938a; 1938b) werden jedoch in der Literatur fast nie zitiert. Dort findet sich gewöhnlich ein Verweis auf eine Veröffentlichung von 1940 (Fletcher 1940), die jedoch lediglich als Überblick gedacht war und ursprünglich 1938 bei einem Dinner der „American Association of Physics Teachers“ vorgetragen wurde. Zur frühen Phase der Forschung zu kritischen Bandbreiten siehe Jont B. Allens Anmerkungen zum 1995 erschienenen ASA-Reprint von H. Fletchers Buch „Speech and Hearing in Communication“: Allen (1995).

⁵⁷ Vgl. Moore (1997), S. 91

⁵⁸ Vgl. Allen (1995)

schnitt 5.2). Diesen Experimenten lag eine Reihe von Annahmen zugrunde, die zusammen als *Energiespektrum-Modell* bezeichnet werden:

- Um ein Signal aus einem Hintergrundgeräusch herauszufiltern, bedient sich das Gehör eines einzelnen Filters, dessen Mittenfrequenz möglichst nahe der Signalfrequenz liegt.
- Der Filter lässt das Signal durch, aber entfernt einen Großteil des Geräuschs.
- Zur Maskierung des Signals tragen nur die Komponenten des Rauschens bei, die denselben Filter passieren wie das Signal.
- Das Gehör nutzt den Filter, der den größten Signal-Rauschabstand aufweist.
- Die Wahrnehmungsschwelle des Signals hängt vom Anteil des Rauschens ab, das vom Filter durchgelassen wird.⁵⁹

Fletcher entwickelte auf dieser Basis eine Methode zur indirekten Schätzung der kritischen Bandbreiten, die sogenannte *Critical Ratio-Methode*. Sie beruht auf Maskierungsexperimenten, bei denen die Wahrnehmungsschwelle eines Sinustons in einem Hintergrund aus Breitband-Rauschen ermittelt wird. Fletcher nahm an, dass unmittelbar bei der Wahrnehmungsschwelle das Verhältnis der Schallenergie des Testtons zur Schallenergie des Rauschens innerhalb der kritischen Bandbreite um den Ton eine Konstante K ergibt. Für Rauschen mit einer Bandbreite unterhalb des kritischen Wertes nahm er an, dass der Wert der Konstante für dieses Signal-Rausch-Verhältnis $K=1$ sein müsse. Ist die Energiedichte des Rauschens bekannt, lässt sich aus diesem Zusammenhang die kritische Bandbreite des Rauschens errechnen. Später stellte sich heraus, dass der Wert $K=1$ und damit auch die errechneten Bandbreiten nicht korrekt sind.⁶⁰

Die Annahmen des Energiespektrum-Modells erweisen sich zur Vorhersage von Maskierungseffekten meist als zuverlässig. Doch heute ist bekannt, dass sie bei weitem nicht immer gültig sind: Erstens lassen sich viele Maskierungseffekte nur durch das Zusammenwirken mehrerer Filter erklären. Zweitens beruhen Wahrnehmungen von Signalen u.U. auf anderen Filtern, die neben der Signalfrequenz liegen (*Off-Frequenz-Listening*). Drittens ignoriert das Energiespektrum-Modell vollständig die zeitliche Dimension von Schall (vgl. Abschnitt 7).⁶¹

5.1.3 Bark-Skala, Frequenzgruppen und Kanalmodell

Nach der Entdeckung des Phänomens wurde ein enormer Aufwand zur experimentellen Bestimmung der kritischen Bandbreiten betrieben. Einflussreich waren vor

⁵⁹ Vgl. Moore (1997), S. 91; Hartmann (1997), S. 256

⁶⁰ Scharf (1970) nennt einen Wert von $K=0,4$; Zwicker/Fastl (1999) stellen fest, dass K zwischen 0,5 (für tiefe Frequenzen) und 0,25 (für hohe Frequenzen) variiert.

⁶¹ Vgl. Hartmann (1997), S. 244-245 und Moore (1995), S. 164-165

allein die Arbeiten Eberhard Zwickers in den 50er und 60er Jahren. Zwicker entwickelte auf der Basis von Fletchers Konzept der kritischen Bandbreiten das *Frequenzgruppenmodell*: Das periphere auditive System wird als eine Reihe von parallelen Bandpassfiltern betrachtet. Diese Filter bewirken eine zunächst eine Spektralanalyse des Schalls: Komponenten komplexer Schallsignale, die in verschiedene Filterbandbreiten fallen, werden als Frequenzgruppe in getrennten parallelen Kanälen verarbeitet. Frequenzgruppen spielen als Analyseeinheiten des Gehörs für die Verarbeitung komplexer Schallsignale eine zentrale Rolle.⁶² Ausgangspunkt für Zwickers Konzept sind die Ergebnisse umfangreicher psychoakustischer Experimente zur Bestimmung der Frequenzgruppenbreite. Er veröffentlichte seine Ergebnisse 1961 als Bark-Skala, in der 24 Frequenzgruppen nebeneinander über den hörbaren Frequenzbereich aufgereiht werden. Die Bark-Skala beruht auf einer ganzen Reihe von Studien, u.a. zur Hörschwellenmessung bei Stille, zur Maskierungshörschwelle, zur Wahrnehmungsschwelle von Modulationen und zur Lautheit.⁶³

Die grafische Darstellung der Bark-Skala (Abbildung 17) verdeutlicht eine grundlegende Eigenschaft der kritischen Bandbreiten: ihre Abhängigkeit von der Frequenz. Zwickers Skala zeigt keine kontinuierlich ansteigende Kurve. Vielmehr nennt die Bark-Skala für die kritischen Bandbreiten bis etwa 500 Hz einen konstanten Wert von etwa 100 Hz; erst darüber hinaus nimmt die Bandbreite in Abhängigkeit von der Mittenfrequenz zu.

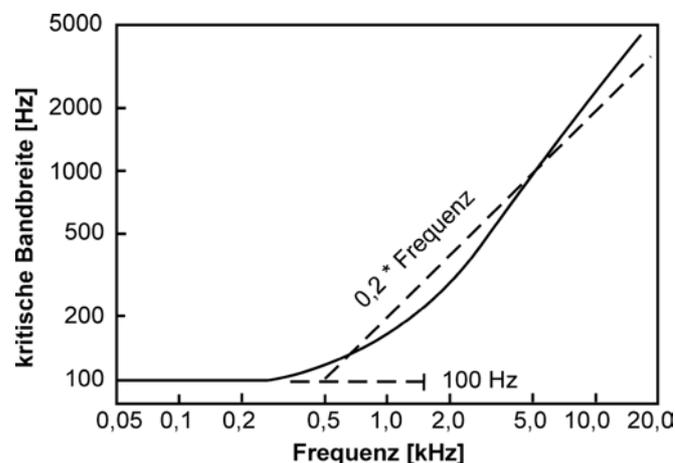


Abbildung 17: Abhängigkeit der kritischen Bandbreite von der Frequenz

Unterhalb von 500 Hz beträgt die kritische Bandbreite nach Zwicker annähernd konstant 100 Hz. Darüber beträgt sie ca. 20% der Mittenfrequenz (gestrichelte Linien). Die durchgezogene Linie zeigt die Werte der Bark-Skala.⁶⁴

⁶² Zwicker verwendet in seinen englischsprachigen Veröffentlichungen für „Frequenzgruppe“ den Ausdruck „Critical Band“.

⁶³ Vgl. Beschreibung der einzelnen Methoden in Zwicker und Fastl (1999); S. 150 ff.

⁶⁴ Nach Zwicker/Fastl (1999), S. 158

Zwicker bringt die Frequenzgruppen der Bark-Skala mit gleich großen Abschnitten von etwa 1,3 mm Länge auf der Basilarmembran in Verbindung. Darüber hinaus stellt er eine Verbindung mit der Mel-Skala der Tonhöhenempfindung her.⁶⁵ Zwicker konstruierte anhand der Frequenzgruppen technische Filtermodelle mit parallelen Kanälen, um mögliche Mechanismen der peripheren Schallverarbeitung zu erforschen.⁶⁶ Die Werte der Bark-Skala galten jahrzehntelang als akzeptiert. Sie bilden die Grundlage einer ISO-Norm zur Lautheitsberechnung (vgl. Abschnitt 8).

5.1.4 *Unsicherheit über physiologische Zusammenhänge*

Die physiologischen Mechanismen, die dem Phänomen der kritischen Bandbreite zugrunde liegen, sind bis heute nicht in allen Einzelheiten bekannt.⁶⁷ Seit den Anfängen der Entwicklung des Filtermodells auf der Basis von Beobachtungen aus psychoakustischen Versuchen haben sich die Erkenntnisse über die Physiologie des Gehörs jedoch bemerkenswert entwickelt. Aufgrund vielfältiger Parallelen zwischen psychoakustischen und physiologischen Erkenntnissen gab es schon seit Fletchers Arbeiten verschiedene ähnlich lautende Annahmen zu möglichen Zusammenhängen. Jedoch ließen sich die Filterbandbreiten aus den psychoakustischen Experimenten nur schlecht mit den damaligen Erkenntnissen über die Frequenzselektivität der Basilarmembran in Einklang bringen.

Übereinstimmung herrschte früh in der Annahme, dass die kritischen Bandbreiten einer exponentiellen Funktion folgen und jeweils einem Abschnitt konstanter Länge auf der Basilarmembran entsprechen.⁶⁸ In der Literatur sind vielfach Längenangaben für diese Abschnitte der Basilarmembran gemacht worden; sie liegen üblicherweise um ca. 1 mm (beim Menschen).⁶⁹ Greenwood stellte 1990 fest, dass die von ihm 1961 entwickelte Frequenz-Positions-Funktion auf der Basis eigener psychoakustischer Experimente auch mit neueren Daten übereinstimmt, die seitdem in verschiedenen Tierversuchen mit verschiedenen Spezies gewonnen wurden.⁷⁰ Mit der Frequenz-Positions-Funktion eng zusammen hängt der Vorschlag, kritische Bandbrei-

⁶⁵ Diese Verbindung würde bedeuten, dass die Ortstheorie der Frequenzcodierung zutrifft und wird daher angezweifelt. Vgl. Hartmann (1997), S. 295

⁶⁶ Vgl. Zwicker/Fastl (1999)

⁶⁷ Hellbrück (1993), Zwicker/Fastl (1999)

⁶⁸ Fletcher (1940); Zwicker et. al. (1957); Scharf (1970), S. 188

⁶⁹ Die Werte variieren zwischen 0,5 mm (Fletcher) und 1,3 mm (z.B. Scharf). Meist fehlt der Hinweis, dass sie von den jeweils zugrundeliegenden Werten für die kritischen Bandbreiten abhängen sowie selbstverständlich auch von der absoluten Länge der Basilarmembran. Greenwood (1990), S. 2604 weist in seiner Zusammenfassung einer großen Zahl von physiologischen Studien zu Frequenz-Positions-Funktionen auf beträchtliche Variationen in der individuellen Länge des Cortischen Organs hin.

⁷⁰ Voraussetzung ist die Normalisierung der Länge der Abschnitte der Basilarmembran und eine Anpassung an den jeweiligen hörbaren Frequenzbereich. Vgl. Greenwood (1990)

ten mit einer absoluten Zahl von beteiligten Neuronen (ca. 1300) in Verbindung zu bringen, woraus früh die Vorstellung neuronaler Erregungsmuster entlang der Basilarmembran entstand.⁷¹ Sowohl Fletcher als auch Zwicker entwickelten Modelle der neuronalen Erregung entlang der Basilarmembran (vgl. Abschnitt 5.2), die geeignet sind, das physiologische Prinzip kritischer Bandbreiten in Maskierung und Lautheitssummation zu erklären.

Es blieb das Problem der mangelnden Übereinstimmung der psychoakustischen messbaren und der beobachtbaren physiologischen Frequenzselektivität. Da von Békésy wesentlich größere Bandbreiten der Basilarmembranschwingung beobachtet hatte, sah sich u.a. Fletcher dazu veranlasst, nach einer anderen Erklärung für seine Filterbandbreiten zu suchen.⁷² Von Békésys Beobachtungen sollten noch jahrzehntelang für Unsicherheit über den Mechanismus der kritischen Bandbreiten sorgen. Aus diesem Grunde findet sich in der Literatur oftmals eine strikte Trennung zwischen der Behandlung des psychoakustischen Phänomens der kritischen Bandbreiten und den Filtereigenschaften der Gehörphysiologie wie z.B. in der vielzitierten Arbeit von Bertram Scharf, in der er 1970 den damaligen Forschungsstand zusammenfasste.⁷³

5.1.5 Neuausrichtung der Filtertheorie und neue Bandbreitenwerte

Seit den späten 70er Jahren, vor allem aber in den 80er Jahren führten neue Erkenntnisse der Physiologie des Gehörs und neuere psychoakustische Studien, wie die Notched Noise-Maskierungsexperimente von Patterson (vgl. Abschnitt 5.3.1.2), zu einer Weiterentwicklung der Filtertheorie. Vor allem Brian C.J. Moore und Kollegen in Cambridge tragen viel dazu bei. Moore stellt detaillierte Vergleiche von Zwickers Versuchen aus den 50er und 60er Jahren mit neueren Studien an und bewertet dessen Ergebnisse neu. Er bestätigt, dass das Phänomen der kritischen Bandbreiten in einer großen Zahl verschiedener Experimente auftritt. Die meisten Experimente ergeben jedoch keinen klaren Bruch in der Hörwahrnehmung, der mit der kritischen Bandbreite korrespondiert. Die Hörwahrnehmung verändert sich vielmehr allmählich und kontinuierlich in Abhängigkeit von der Bandbreite. Die Annahme, dass die Experimente einen Bruch ergeben sollten, die sich z.T. bei Zwicker findet, führt Moore auf die seit Fletcher als Vereinfachung benutzte Annahme rechteckiger Filter zurück.⁷⁴

⁷¹ Vgl. Scharf 1970, S. 188

⁷² Vgl. Allen (1995)

⁷³ Vgl. Scharf (1970)

⁷⁴ Zu den Unterschieden zwischen „Münchener“ Werten nach Zwicker und den „Cambridge“-Werten nach Moore vgl. Hartmann (1997), S. 249

Fletcher und Zwicker konnten die kritischen Bandbreiten anhand verschiedener Phänomene experimentell nachweisen und abschätzen, deren Zusammenhang mit der Frequenzselektivität des Gehörs blieb jedoch unklar (vgl. Abschnitt 5.1.4). Bei Moore steht dagegen der Begriff der auditiven Filter im Mittelpunkt. Er führt die kritischen Bandbreiten direkt auf die Eigenschaften der psychophysisch messbaren Frequenzselektivität zurück und betrachtet die kritische Bandbreite als die *effektive Bandbreite* der auditiven Filter. Moore kann sich auf neue Erkenntnisse zur Mikromechanik der Basilarmembran stützen (vgl. Abschnitt 4.1) und geht davon aus, dass die *psychophysisch* messbaren Filtereigenschaften des Gehörs ausreichend durch die *physiologisch* nachweisbare Frequenzselektivität des peripheren auditiven Systems erklärt werden können. Obwohl auch Moore angibt, dass die physiologische Grundlage der kritischen Bandbreiten noch unsicher ist, geht er davon aus, dass sie auf der Frequenzselektivität der Basilarmembran beruht.

Moore u.a. stellen umfangreiche Studien zur Form der auditiven Filter an und ermitteln eine neue Funktion der Bandbreitenwerte, die von der „traditionellen“ Bark-Skala abweicht. Ihre Bandbreiten sind kleiner, und ihre Frequenzabhängigkeit setzt sich unterhalb von 500 Hz fort, wo Zwicker eine konstante Bandbreite angibt (vgl. Abschnitt 5.3.2).⁷⁵ Patterson et. al. entwickeln Anfang der 90er Jahre die Gammaton-Filterfunktion, die für Modelle des peripheren auditiven Systems eine wichtige Rolle erlangt (vgl. auch Abschnitt 8 - Exkurs: Anwendungen des Filter; sowie Abschnitt 9.2 - Der heutige Stellenwert des Filtermodells).⁷⁶

5.1.6 Zur Definition des Begriffs „kritische Bandbreite“

Der Gebrauch des Begriffs *kritische Bandbreite* in der Literatur kann leicht für Verwirrung sorgen. Dazu trägt die historische Entwicklung bei. Ursprünglich war der Begriff reserviert für ein psychoakustisches Phänomen: Die Änderung der Hörempfindung, je nachdem ob das Frequenzspektrum eines Schalls eine bestimmte Bandbreite überschreitet oder nicht. Die Bandbreite, bei der die Änderung eintritt, wird kritische Bandbreite genannt. Scharf (1970) unterscheidet noch strikt zwischen dieser empirischen Definition und dem Filtermodell des Gehörs. Dagegen wird in der neueren Literatur die kritische Bandbreite direkt auf die Filtereigenschaften des peripheren auditiven Systems bezogen. Sie wird als die effektive Bandbreite der auditiven Filter betrachtet.⁷⁷ Der Vergleich mit deutschsprachiger Literatur wird dadurch verkompliziert, dass Zwicker in seinen deutschsprachigen Veröffentlichungen den Begriff *Frequenz-*

⁷⁵ Vgl. Hartmann (1997), S. 238 ff.

⁷⁶ ebd., S. 254

⁷⁷ Vgl. Hartmann (1997)

gruppe prägte, in seinen englischsprachigen Veröffentlichungen aber *Critical Band* und *Critical Bandwidth* benutzte und beibehielt (vgl. Zwicker/Fastl 1999), während sich die Bedeutung des Begriffs wandelte.

In dieser Arbeit werden häufig zwei Bandbreite-Skalen gegenübergestellt. Zum einen die Bark-Skala nach Zwicker, zum anderen die Bandbreitenwerte nach Moore. Moores Werte werden in der Literatur oft als ERB-Werte bezeichnet, was insofern unglücklich ist, als sowohl Zwicker als auch bereits Fletcher in ihren Werten nichts anderes als ERB-Werte sahen (vgl. Abschnitt 5.1.1 und Abschnitt 5.3.2). In der neueren Literatur finden sich daher für die Bark-Skala auch die Bezeichnungen „traditionelle“ (Moore 1995, 1997) oder „Münchener“ (Hartmann 1997) Bandbreiten.⁷⁸ In dieser Arbeit werden diese Werte im Folgenden als *Zwicker-Werte* oder *Werte der Bark-Skala* bezeichnet. Der Begriff *ERB-Werte* oder *Cambridge-ERBs* bleibt für Verweise auf die neueren Werte nach Moore reserviert.⁷⁹

5.2 Maskierung und neuronale Erregungsmuster

5.2.1 Maskierung und das Frequenzauflösungsvermögen des Gehörs

Das Phänomen der Maskierung (bzw. Verdeckung) gehört zu unseren alltäglichen Hörerfahrungen. Ein Beispiel ist die Hörbarkeit von Sprache bei einer Unterhaltung neben einer befahrenen Straße. Beim Vorbeifahren lauter Fahrzeuge muss der Sprecher im Vergleich zu ruhigen Momenten wesentlich lauter reden, um verstanden oder um überhaupt gehört zu werden, da der Lärm den Sprachschall „verdeckt“. Maskierung bezeichnet zum einen den *Vorgang*, durch den die Schwelle der Hörbarkeit eines Signals durch die gleichzeitige Anwesenheit eines weiteren Schallsignals angehoben wird, zum anderen das *Maß*, um das die Hörbarkeitsschwelle eines Schallsignals durch die gleichzeitige Anwesenheit eines weiteren Schallsignals angehoben wird.⁸⁰ Maskierungsexperimente spielen eine wichtige Rolle bei der experimentellen Erforschung der Funktionsweise des Gehörs. Sie spielten eine zentrale Rolle bei Fletchers Arbeiten zum Phänomen der kritischen Bandbreiten (vgl. Abschnitt 5.1.1). Ein Signal wird am ehesten durch ein zweites verdeckt, wenn dessen Komponenten in der Nähe der Frequenz des ersten Signals liegen. Darauf beruht die Annahme, dass Maskierungseffekte Aufschlüsse über das Frequenzauflösungsvermögen des Gehörs liefern können. Wenn ein Ton den anderen gerade eben verdeckt, bedeutet dies, dass das Gehör daran scheitert, die beiden Töne aufzulösen.

⁷⁸ Nach der TU München, an der Zwicker arbeitete

⁷⁹ Nach Moores Wirkungsstätte, der University of Cambridge (GB)

⁸⁰ Nach der Definition der American Standards Association; vgl. Moore (1997), S. 89

Neuronale Erregungsmuster sind ein Modell, das dieses Frequenzauflösungsvermögen auf der Grundlage der Frequenzselektivität der Basilarmembran erklärt und eine Erklärung für Maskierungseffekte liefert.⁸¹ Neuronale Erregungsmuster werden als eine interne Repräsentation des Spektrums des Stimulus betrachtet.⁸² Sie geben das Maß der Reaktion der Neuronen entlang der Basilarmembran auf einen Stimulus wieder. Die Tuningkurven (FTCs) einzelner Nervenfasern zeigen, dass diese auf ein schmales Band von Frequenzen um eine charakteristische Frequenz reagieren (vgl. Abschnitt 4.2). Das heißt, dass ein einzelner Ton nicht nur die Reaktion eines einzelnen Neurons auslöst, sondern auch – in geringerem Maße - Aktivität in Neuronen mit angrenzenden charakteristischen Frequenzen erzeugt. Deren Aktivität nimmt mit der Entfernung von der Frequenz des Tons ab. Wenn man die Neuronen entsprechend ihrer charakteristischen Frequenz auf einer Achse anordnet und entlang dieser Achse deren Entladungsaktivität aufträgt, erhält man Muster der Verteilung neuronaler Aktivität auf der Basilarmembran (Abbildung 18).

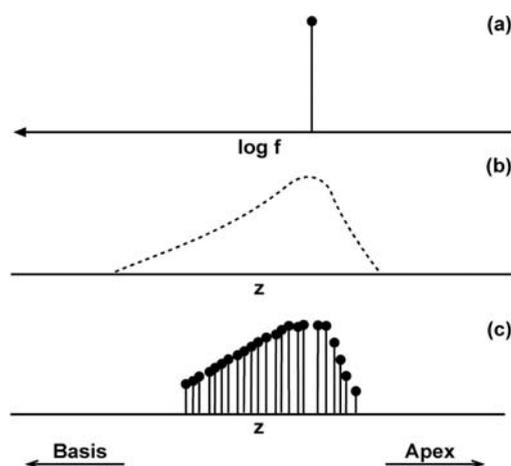


Abbildung 18: Neuronale Erregungsmuster

(a) Spektrum eines Sinus-Tons; (b) Hüllkurve der dadurch ausgelösten Basilarmembranschwingung; (c) neuronale Erregungsmuster entlang der Basilarmembran.⁸³

5.2.2 Psychophysische Maskierungsmuster

Die Form neuronaler Erregungsmuster lässt sich beim Menschen, wo keine direkte Messung neuronaler Aktivität möglich ist, in Form so genannter psychophysischer Maskierungsmuster (*Masking Pattern, Masked Audiogram*) annähern. Sie werden in einem Experiment ermittelt, bei dem ein schmales Rauschband mit fester Mittenfrequenz und Bandbreite als Maskierungssignal dient. Für Sinustöne mit verschiedenen Frequenzen werden dann die Mithörschwellen ermittelt.

⁸¹ Vgl. Pickles (1988), S. 258

⁸² Vgl. Moore (1997), S. 36 ff.; Hartmann (1997), S. 242 ff.

⁸³ Nach Hartmann (1997), S. 243

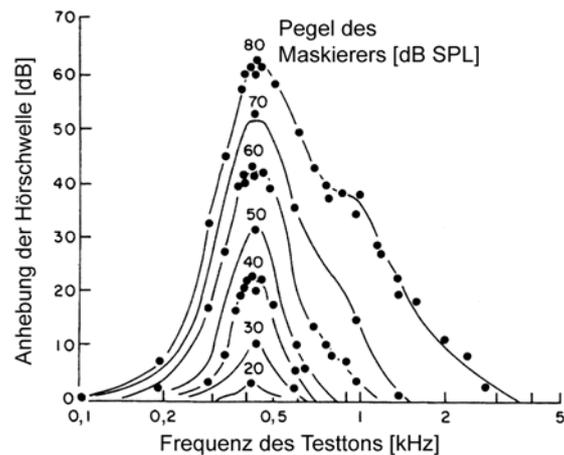


Abbildung 19: Maskierungsmuster (Masked Audiograms)

Mithörschwellen für ein Bandpassrauschen (Bandbreite 90 Hz, Mittenfrequenz 410 Hz) mit verschiedenen Pegeln.⁸⁴

Das Maskierungsmuster erhält man, indem man in einem Diagramm für jede getestete Frequenz die ermittelte Mithörschwelle einträgt (vgl. Abbildung 19). Eine Interpretation dieses Musters ist, den Verlauf der Mithörschwelle als ein Maß für die Aktivität der Neuronen zu betrachten, die das Maskierungssignal auslöst. Die Aktivität ist innerhalb der Bandbreite des Rauschens am stärksten und fällt jenseits davon ab, so dass Testtöne schlechter gehört werden, je näher ihre Frequenz an den Grenzfrequenzen des Rauschens liegt. Auffällig ist, dass der Maskierungseffekt in Richtung der hohen Frequenzen mit höherem Pegel des Maskierers überproportional wächst. Diese Asymmetrie verhält sich umgekehrt zu der der FTCs einzelner Neuronen. Dies lässt sich dadurch erklären, dass FTCs die Antwort *eines* Neurons auf mehrere Stimuli verschiedener Frequenz wiedergeben. Maskierungsmuster sind das Ergebnis des Antwortverhaltens von *mehreren* Neuronen verschiedener charakteristischer Frequenzen (CF) auf ein Maskierungssignal fester Frequenz. Das Maskierungssignal aktiviert auch Neuronen mit höherer CF, da diese in geringem Maße auch auf niedrigere Frequenzen ansprechen. Umgekehrt aktiviert das Maskierungssignal weniger Neuronen mit niedrigerer CF, da diese wesentlich weniger auf höhere Frequenzen reagieren (deutlich zu sehen in der steileren Flanke der FTC in dieser Richtung; vgl. Abbildung 16).⁸⁵

5.2.3 Berechnung neuronaler Erregungsmuster

Die auf diese Weise ermittelten Muster sind nur eine ungenaue Repräsentation neuronaler Erregungsmuster. Erregungsmuster lassen sich – ebenfalls auf der

⁸⁴ Aus Pickles (1988), S. 259

⁸⁵ Vgl. Pickles (1988), S. 258-259

Grundlage von Maskierungsexperimenten - auch genauer annähern.⁸⁶ Man geht dabei von den einzelnen auditiven Filtern aus, deren Form mit der Notched-Noise Methode ermittelt werden kann (vgl. Abschnitt 5.3.2). Das Erregungsmuster für ein Signal wird aus den Ausgangssignalen von denjenigen Filtern berechnet, die in Abhängigkeit von ihrer jeweiligen Mittenfrequenz und Durchlassfunktion verschieden stark auf das Signal ansprechen (vgl. Abbildung 20: Berechnung des Erregungsmusters für ein 1 kHz - Sinussignal).

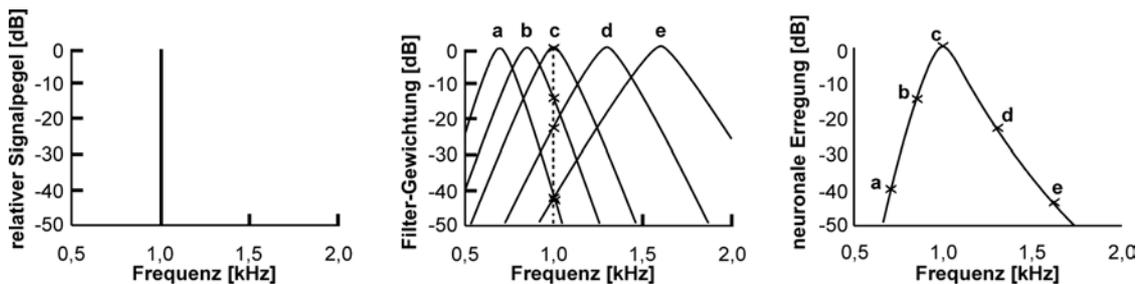


Abbildung 20: Berechnung des Erregungsmusters für ein 1 kHz - Sinussignal

Ein einzelner Sinuston mit einer Frequenz von 1 kHz (linke Abbildung: Spektrum) fällt in die Übertragungsbereiche mehrerer auditiver Filter, da sich diese überlappen (mittlere Abbildung). Das Erregungsmuster entlang der Basilarmembran ergibt sich aus den Ausgangssignalen der einzelnen Filter.⁸⁷

Obwohl in der Regel angenommen wird, dass die Filterform symmetrisch ist, ergibt sich eine asymmetrische Form des Erregungsmusters. Die Asymmetrie entsteht, weil die Filter höherer Mittenfrequenz größere Bandbreiten haben.

5.2.4 Maskierung als Effekt der Überlagerung neuronaler Aktivität

Maskierungseffekte treten auf, wenn sich die neuronalen Erregungsmuster zweier Signalkomponenten überlagern. Bis zu einem kritischen Frequenzabstand werden Signalkomponenten klar getrennt wahrgenommen, da sich die von ihnen ausgelösten Erregungsmuster der Basilarmembran nicht überschneiden (Abbildung 21, (a) und (b)). Erst wenn der Frequenzabstand der Signalkomponenten ein kritisches Maß unterschreitet, beeinflusst die dann auftretende Überschneidung der Erregungsmuster die Hörwahrnehmung (Abbildung 21, (c) und (d)). Die Asymmetrie der Erregungsmuster äußert sich darin, dass höhere Frequenzen leichter verdeckt werden als tiefere (*Upward Spread of Masking*).

⁸⁶ Moore/Glasberg (1983); nach Moore (1997), S. 111 ff.

⁸⁷ Nach Moore (1995), S. 181

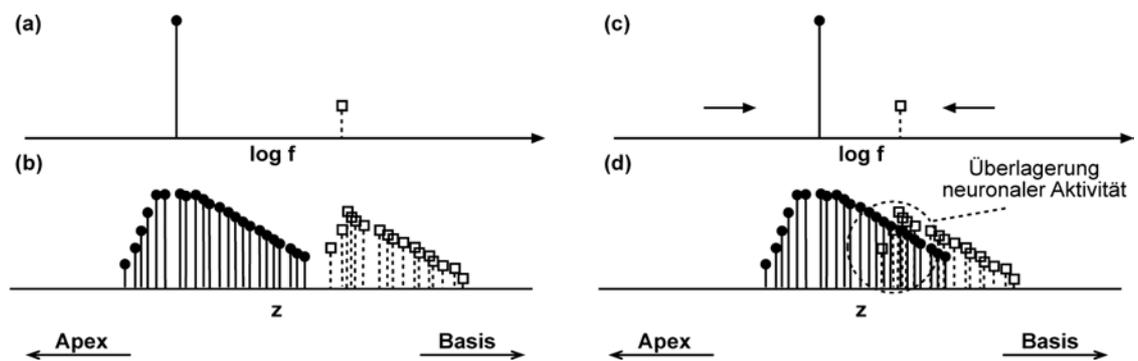


Abbildung 21: Überlagerung neuronaler Erregungsmuster

Zwei Sinustöne mit ausreichendem Frequenzabstand (a) lösen auf der Basilarmembran unabhängige Erregungsmuster aus (b). Wenn der Abstand einen kritischen Wert unterschreitet, überlagern sich die Erregungsmuster, Maskierungseffekte in der Hahrnehmung treten auf.⁸⁸

Die Einzelheiten der physiologischen Mechanismen der Maskierung sind ungeklärt. Eine Hypothese ist die der „Überschwemmung“ (*Swamping*) neuronaler Aktivität: Alle Frequenzkomponenten eines Signals lösen neuronale Aktivität aus. Jedoch kann die von einzelnen Komponenten ausgelöste Aktivität weit überwiegen. Durch das ungleiche Verhältnis wird die durch die schwächere Komponente ausgelöste neuronale Aktivität „überschwemmt“. Wird dadurch ihr Maximum verdeckt, kann sie nur noch über neuronale Aktivität in Regionen wahrgenommen werden, die weiter vom Erregungsmuster des dominanten Signals entfernt liegen; dies wird als *Off-Frequency Listening* bezeichnet.⁸⁹ Ein weiterer möglicher Mechanismus der Maskierung ist der der Hemmung (*Suppression*): Neuronale Aktivität eines Neurons als Antwort auf eine Signalkomponente mit dessen charakteristischer Frequenz kann durch eine weitere Komponente unterdrückt bzw. gehemmt werden, die selbst keine Aktivität des besagten Neurons auslöst. Ein Mechanismus der Hemmung würde eine bedeutende Nicht-Linearität darstellen. Dass lineare Filtermodelle jedoch in der Regel ausreichen, um die Mithörschwellen von Signalen bei simultaner Maskierung vorherzusagen, spricht für eine entscheidende Rolle des Verschüttungs-Mechanismus.⁹⁰

5.2.5 Modelle der Lautheit

Die erste Erwähnung von kritischen Bandbreiten stand im Zusammenhang mit der empfundenen Lautstärke von der Bandbreite des Stimulus (vgl. Abschnitt 5.1.1). Ähnlich wie bei Maskierungseffekten bieten neuronale Erregungsmuster auch ein Erklärungsmodell für die Lautheit komplexer Schallstimuli. Bereits Fletcher schlug vor,

⁸⁸ Nach Hartmann (1997), S. 243

⁸⁹ Vgl. Hartmann (1997), S. 244; Moore (1997), S. 117 ff.

⁹⁰ Vgl. Moore (1997), S. 117 ff.

die empfundene Gesamtlautstärke von breitbandigen Stimuli mittels der Summation neuronaler Aktivität zu erklären.⁹¹

Die Gesamtlautheit komplexer Stimuli hänge demnach von der neuronalen Aktivität über mehrere Frequenzkanäle hinweg ab. Ein Maß der tatsächlichen Aktivität der Neuronen als Reaktion auf bestimmte Stimuli ist schwer definierbar, da sich sowohl die spontane als auch die maximalen Reaktionen einzelner Neuronen erheblich unterscheiden. Daher werden neuronale Erregungsmuster oft durch das Maß der effektiven *Anregung* repräsentiert. Modelle der Lautheit haben die gleiche grundlegende Struktur aus mehreren Stufen (siehe Abbildung 22).

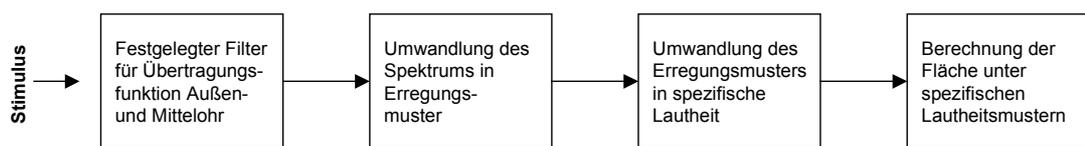


Abbildung 22: Grundstruktur von Modellen der Lautheit⁹²

Die erste Stufe ist ein Filter, das der Übertragungsfunktion des äußeren Ohres entspricht. Die zweite Stufe ist eine Berechnung des neuronalen Erregungsmusters für den jeweiligen Stimulus auf der Basis eines Maßes der Anregung. Auf der nächsten Stufe erfolgt eine Umwandlung von der Anregung in eine spezifische Lautheit. Dies ist eine Repräsentation der Lautheit, die über einen bestimmten festgelegten Abschnitt der Basilarmembran summiert wird. Die Gesamtlautheit eines breitbandigen Signals wird auf der vierten Stufe aus der Fläche der spezifischen Lautheitsmuster errechnet.⁹³

5.3 Eigenschaften der auditiven Filter

5.3.1 Die Form der auditiven Filter

5.3.1.1 Psychophysische Frequenz-Tuningkurven

Die Ermittlung psychophysischer (oder *psychoakustischer*) Frequenz-Tuningkurven (PTC) ist die direkteste Methode, eine Annäherung der Form auditiver Filter zu erhalten. Das Vorgehen ähnelt dem zur Ermittlung von neuronalen Frequenz-tuningkurven (FTC, vgl. Abschnitt 4.2). Grundlage für die Annahme, dass sie die Form der auditiven Filter beschreiben, ist das Energiespektrum-Modell (vgl. Abschnitt 5.1).

⁹¹ Ausführlich: Vgl. Fletcher (1940) über *auditory patterns*

⁹² Nach Moore (1997), S. 60

⁹³ Vgl. Zwicker/Fastl (1999), Moore (1997)

Es wird angenommen, dass für die Entdeckung eines Testsignals zumindest bei geringem Pegel lediglich *ein* auditiver Filter benutzt wird, dessen Mittenfrequenz der Signalfrequenz entspricht. Des Weiteren wird angenommen, dass der auditive Filter bei der Mithörschwelle des Testsignals ein konstantes Ausgangssignal liefert.

Bei der experimentellen Ermittlung von PTCs wird Versuchspersonen ein Ton mit geringem Schalldruckpegel und konstanter Frequenz dargeboten. Dann bietet man Maskierungssignale, meist schmalbandiges Rauschen, mit variierender (Mitten-) Frequenz an und ermittelt, bei welchem Schalldruckpegel der Testton gerade noch hörbar ist.⁹⁴ Eine Filtercharakteristik wird normalerweise bestimmt, indem das Ausgangssignal in Abhängigkeit von einem *Eingangssignal* mit *konstantem* Pegel und variierender Frequenz gemessen wird. Daher liefert die beschriebene Art der Messung psychophysischer Tuningkurven - die Linearität des Filters vorausgesetzt - theoretisch die invertierte Form des Filters.

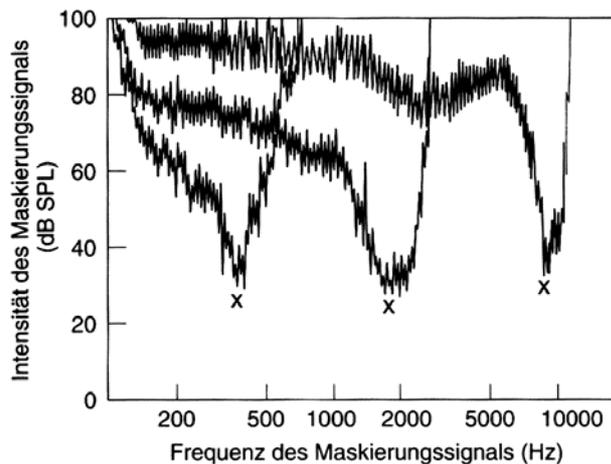


Abbildung 23: Psychophysische Frequenz-Tuningkurven⁹⁵

Psychoakustische Frequenz-Tuningkurven ähneln in der Form den in Abschnitt 4.2 beschriebenen neuronalen Frequenz-Tuningkurven, so dass es naheliegend erscheint, dass sie die Form der auditiven Filter annähernd zutreffend beschreiben. Jedoch gibt es Einwände: Neuronale Frequenz-Tuningkurven geben die Antwort eines einzelnen Neurons auf einen einzigen anregenden Ton wieder, während bei der Messung von PTCs mehrere Töne oder komplexe Signale gleichzeitig vorhanden sind. Es ist jedoch wahrscheinlich, dass das Gehör Filter nutzt, deren Mittenfrequenz nicht der Signalfrequenz entspricht, wenn dies für die Wahrnehmung eines Testsignals von Vorteil ist (sogenanntes *Off-Frequency Listening*). Dies kann z.B. der Fall sein, wenn

⁹⁴ Die Verwendung von Rauschen soll die bei Sinustönen auftretenden Schwebungen verhindern, da diese die Wahrnehmungsschwelle für den Testton beeinflussen können.

⁹⁵ Aus Goldstein (1997), S. 369

das Maskierungssignal oberhalb der Frequenz des Testtones liegt und der Filter einen relativ flachen Frequenzgang mit abfallenden Seiten besitzt. Die Nutzung eines Filters etwas unterhalb der Mittenfrequenz würde den Anteil des Maskierungssignals am Output beträchtlich dämpfen, das Testsignal selbst aber nur unwesentlich.⁹⁶ Des Weiteren hat sich gezeigt, dass sich auditive Filter nicht linear verhalten; ihre Form ändert sich bei höheren Schallpegeln. Auf diese Nicht-Linearität wird im Folgenden noch näher eingegangen.

5.3.1.2 Die Notched-Noise Methode

Die sogenannte *Notched-Noise* Methode zur Untersuchung der Form der auditiven Filter soll Off-Frequency Listening ausschließen. Dabei wird als Maskierer ein Rauschen verwendet, das eine Aussparung im Frequenzgang hat, eine Lücke, deren Mittenfrequenz der Frequenz des Testtons entspricht. Die Breite der Aussparung wird variiert und dabei die Wahrnehmungsschwelle für das Testsignal gemessen. Das optimale Verhältnis von Signal zu Maskierungsgeräusch wird bei dieser Methode ausschließlich mit dem auditiven Filter erreicht, deren Mittenfrequenz der Signalfrequenz entspricht. Diese Methode ist allerdings nicht dazu geeignet, Asymmetrien in der Filterform aufzuzeigen. Jedoch legen die PTCs nahe, dass die auditiven Filter zumindest im Bereich der Spitze weitgehend symmetrisch sind, so dass die Notched-Noise Methode eine brauchbare Annäherung liefert.⁹⁷

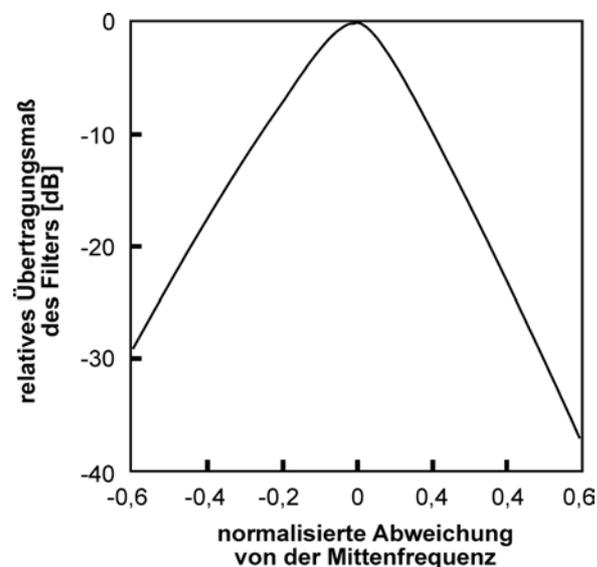


Abbildung 24: Typische Form eines auditiven Filters
Ermittelt mit der Notched Noise-Methode.⁹⁸

⁹⁶ Vgl. Moore (1997), S. 103-105; Hartmann (1997), S. 243-244

⁹⁷ Vgl. nach Moore (1997), S. 105 ff.

⁹⁸ Nach Moore (1995), S. 174

5.3.1.3 Asymmetrie und Pegelabhängigkeit der Filterform

Eine Abwandlung der Notched-Noise Methode erlaubt es, Asymmetrien in der Filterform festzustellen. Hierbei wird die Lücke im Spektrum des Maskierers asymmetrisch um die Signalfrequenz platziert. Die Ergebnisse bestätigen, dass die Filter bei moderaten Schallpegeln eine annähernd symmetrische Form haben. Bei höheren Pegeln wird allerdings deutlich, dass sich die auditiven Filter nicht linear verhalten: Ihre Form verändert sich in Abhängigkeit vom Pegel des Eingangssignals.

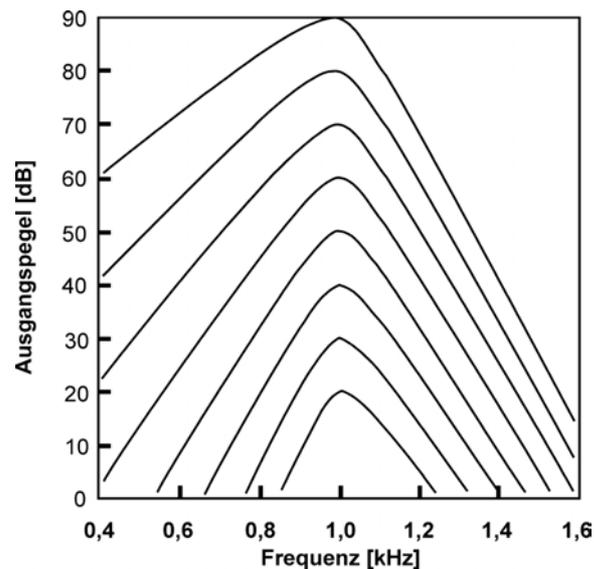


Abbildung 25: Abhängigkeit der Filterform vom Schallpegel⁹⁹

Bei höheren Pegeln wird die Form zunehmend asymmetrisch: Die Flanke zu niedrigen Frequenzen hin wird flacher als die in Richtung der hohen Frequenzen. Die Veränderungen für die Seite hin zu hohen Frequenzen sind nicht eindeutig, vor allem im unteren Frequenzbereich. Für Mittenfrequenzen im Bereich von 1-4 kHz wird sie jedoch mit zunehmendem Pegel steiler.¹⁰⁰

5.3.2 Die Bandbreite der auditiven Filter

Da die Form der auditiven Filter nicht rechteckig ist, reicht die Angabe der kritischen Bandbreite alleine nicht aus, um sie präzise zu beschreiben. Aus Gründen der einfacheren Handhabung ist es jedoch üblich, bei Filtern allgemein lediglich eine Bandbreite anzugeben (vgl. Abschnitt 2.4). In fast allen Studien zu auditiven Filtern wird mit *äquivalenten* Bandbreiten hypothetischer rechteckiger Filter (*Equivalent Rec-*

⁹⁹ Aus Moore (1995), S. 178

¹⁰⁰ Vgl. Moore (1995), S. 177-178

tangular Bandwidths, ERB) gearbeitet.¹⁰¹ Äquivalenz bedeutet in diesem Fall, dass der angenommene rechteckige Filter bei weißem Rauschen als Eingangssignal exakt genauso viel akustische Energie passieren lässt wie der reale Filter. Die ERB lässt sich auf Grundlage der Übertragungsfunktion des realen Filters $H(f)$ folgendermaßen berechnen:

$$B_{ER} = \int_0^{\infty} df |H(f)|^2 \quad [3]$$

B_{ER} = äquivalente rechteckige Filterbandbreite

Voraussetzung ist, dass $H(f)$ normiert ist und höchstens 1 erreicht

Formel 3: Äquivalente rechteckige Filterbandbreite (ERB)¹⁰²

Neuere Messungen der Bandbreite weichen insbesondere im unteren Frequenzbereich deutlich von den Ergebnissen Zwickers ab. Moore (1997) vergleicht eigene Messungen mittels Pattersons Notched-Noise Methode und andere neuere Ergebnissen mit den Werten der Bark-Skala.¹⁰³

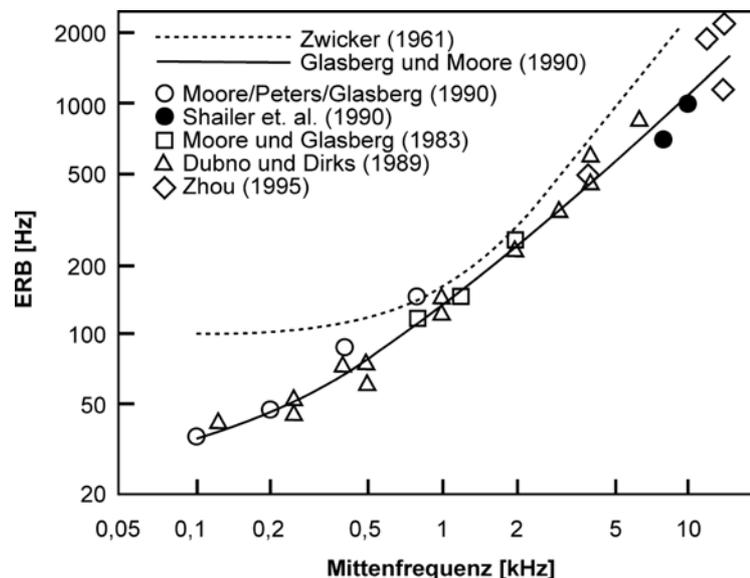


Abbildung 26: Übersicht über kritische Bandbreiten nach verschiedenen Studien

Die Symbole kennzeichnen Ergebnisse aus den verschiedenen angegebenen Studien. Diese Werte liegen alle deutlich näher an Moores Werten als an der Bark-Skala.¹⁰⁴

¹⁰¹ Diese Vereinfachung wird vielfach implizit vorausgesetzt. Dadurch kann unter Umständen der Eindruck entstehen, dass die rechteckige Filterform den realen Filtern entspricht. Dieser wird in einigen (früheren) Studien scheinbar bestätigt, wo z.T. (fälschlicherweise) abrupte Änderungen der Wahrnehmung in Abhängigkeit von der kritischen Bandbreiten festgestellt wurden (vgl. Abschnitt 9.1).

¹⁰² Vgl. Hartmann (1997), S. 245-246

¹⁰³ Vgl. Moore (1997), S. 107-109

¹⁰⁴ Nach Moore (1997) S. 108

Aus der Grafik ist deutlich die Abhängigkeit der Bandbreite der auditiven Filter von der Mittenfrequenz erkennbar. Im Gegensatz zu Zwickers Ergebnissen (gestrichelte Linie) scheint diese Abhängigkeit jedoch auch für den Bereich unterhalb von 500 Hz gültig zu sein; laut den hier aufgetragenen Werten aus neueren Studien nehmen auch dort die Werte kontinuierlich ab.¹⁰⁵ Moore (1997) vermutet, dass Zwickers Werte von „indirekten“ Methoden wie der Critical Ratio – Berechnung und Messungen der kritischen Modulationsfrequenz beeinflusst waren. Er merkt an, dass Zwicker offenbar besonders für den unteren Frequenzbereich nur wenige Daten aus direkten Messungen vorlagen. Die ERBs der auditiven Filter nach Moore liegen typischerweise zwischen 11% und 17% der Mittenfrequenz.¹⁰⁶ Bandbreiten für beliebige Mittenfrequenzen f_c bis 10 kHz können mit folgender Funktion ermittelt werden (oberhalb von 10 kHz liegen wenige Erkenntnisse zur kritischen Bandbreite vor):

$$B_C = 24,7(1 + 4,37f_c) \quad [4]$$

B_C = Bandbreite des Filters
 f_c = Mittenfrequenz [kHz]

Formel 4: Bandbreite-Funktion für Cambridge-ERB

Sowohl die Werte der Bark-Skala als auch die Ergebnisse Moores entsprechen in etwa Filtern mit einer Bandbreite von 1/3 Oktave. Diese werden als leicht handhabbare Annäherung der menschlichen auditiven Filter z.B. in der Audio-Industrie genutzt (vgl. Abbildung 27; zu Anwendungen siehe auch Abschnitt 8).¹⁰⁷

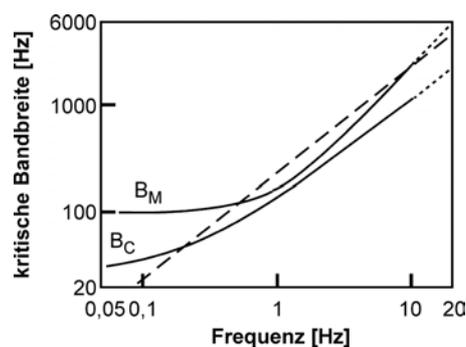


Abbildung 27: Bark-Skala, „Cambridge“-ERBs und 1/3-Oktav-Filter

Die mit BM gekennzeichnete Linie stellt die Werte der Bark-Skala dar, die mit BC bezeichnete Linie die ERB-Werte nach Moore („Cambridge“-ERB). Die gestrichelte Linie stellt 1/3-Oktav-Filter dar, die in der Audio-Industrie als brauchbare Annäherung an die auditiven Filter benutzt werden.¹⁰⁸

¹⁰⁵ Vgl. Moore (1995), S. 175

¹⁰⁶ Vgl. Moore (1997), S. 93

¹⁰⁷ Vgl. Hartmann (1997), S. 250

¹⁰⁸ Nach Hartmann (1997), S. 250

5.3.3 Nicht-simultane Maskierung und auditive Filter

Die in den vorhergehenden Abschnitten beschriebenen Erkenntnisse zu Bandbreite und Form der auditiven Filter wurden hauptsächlich aus Maskierungsexperimenten abgeleitet, bei denen Signal und Maskierer gleichzeitig präsentiert wurden (simultane Maskierung). Eine Reihe von Veränderungen treten jedoch auf, sobald Maskierer und Testsignal zeitlich getrennt sind (nicht-simultane Maskierung). Mögliche Variationen sind, das Signal vor dem Maskierer (*Premasking*)¹⁰⁹ oder nach dem Maskierer zu präsentieren (*Postmasking*)¹¹⁰.

Während die Ergebnisse von Studien zu Premasking widersprüchlich und schwer zu interpretieren sind, liegen zu Postmasking einige übereinstimmende Erkenntnisse.¹¹¹ Postmasking tritt auf, wenn ein Testsignal innerhalb eines kurzen Zeitintervalls von ca. 100-200 ms nach dem Maskierer präsentiert wird. Je näher das Testsignal zeitlich an das Maskierungssignal herangerückt wird, desto stärker ist der Maskierungseffekt. Auch bei nicht-simultaner Maskierung lässt sich die Form der auditiven Filter mittels der Notched-Noise Methode untersuchen.

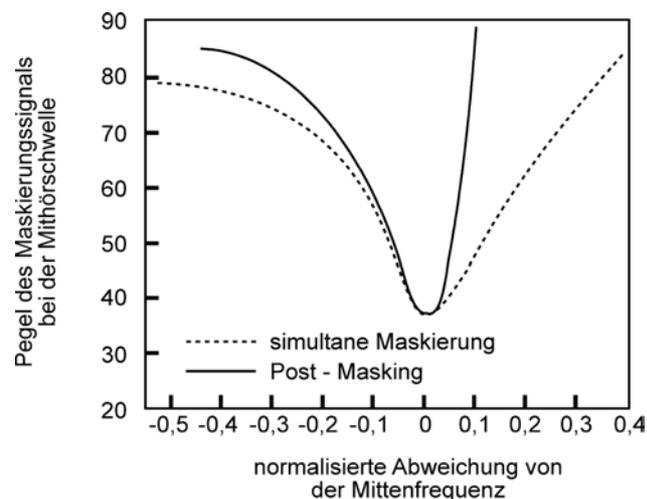


Abbildung 28: PTCs für simultane und nicht-simultane Maskierung
Mittenfrequenz = 2 kHz.¹¹²

Die Form der auditiven Filter ist bei Postmasking schmaler (Abbildung 28). Daraus ergeben sich folglich geringere Werte für die kritischen Bandbreiten bzw. ERBs.¹¹³ Mögliche Erklärungsansätze für das Auftreten nicht-simultaner Maskierung sind die Verringerung der Empfindlichkeit der vom vorausgehenden Maskierungssig-

¹⁰⁹ Auch: *Backward Masking*; vgl. Moore (1997)

¹¹⁰ Auch: *Forward Masking* oder *Poststimulatory Masking*; vgl. Moore (1997)

¹¹¹ Vgl. Moore (1995), S. 194-195

¹¹² Nach Moore (1995), S. 198

¹¹³ Vgl. Pickles (1988), S. 261; Moore (1995), S. 198

nal stimulierten Neuronen oder das Fortdauern von dadurch hervorgerufenen Mustern neuronaler Aktivität. Dazu kommt, dass die Anregung der Basilarmembran eine gewisse Zeit benötigt, um abzuklingen.¹¹⁴ Die Präzisierung der Frequenzselektivität bei nicht-simultaner Maskierung ist insofern von großer Bedeutung, als dass natürliche akustische Reize eine ausgeprägte zeitliche Struktur und rasche Fluktuationen aufweisen. Das bedeutet, dass man in vielen natürlichen Hörsituationen möglicherweise von schmaleren Bandbreiten der auditiven Filter ausgehen muss.¹¹⁵

6 Kritische Bandbreiten als empirisches Phänomen

Nachdem Fletcher das Auftreten kritischer Bandbreiten bei grundlegenden Aspekten der Hörwahrnehmung - Maskierung und Lautstärkewahrnehmung – nachgewiesen hatte (vgl. Abschnitt 5.1.1), wurde eine Vielzahl von Studien durchgeführt, um das Phänomen in den verschiedensten Zusammenhängen zu untersuchen. Da Unsicherheit darüber herrschte, wie die physiologischen Mechanismen der Frequenzanalyse mit diesem Phänomen in Verbindung gebracht werden könnten, wurde es lange rein empirisch definiert: Die kritische Bandbreite ist die Bandbreite eines akustischen Reizes, bei deren Über- bzw. Unterschreiten in psychoakustischen Experimenten eine Änderung der Hörempfindung eintritt. Scharf (1970) und Zwicker/Fastl (1999) fassen mehrere Studien zusammen, in denen übereinstimmend kritische Bandbreiten nachgewiesen wurden. Der kritische Wert der Bandbreite, bei der eine Änderung der Hörwahrnehmung eintritt, erwies sich bei vielen dieser Zusammenhänge als gleich oder ähnlich. Zwicker entwickelte eine Reihe von beispielhaften Experimenten, die als Methoden zur Messung von Bandbreite-Werten dienen können.¹¹⁶

6.1 Lautheit komplexer Schallsignale

Die Abhängigkeit der empfundenen Gesamtlautstärke von Schallsignalen von ihrer Bandbreite gehört seit Fletchers Studien in den 30er und 40er Jahren zu den am gründlichsten untersuchten Aspekten kritischer Bandbreiten. Ein komplexes Signal wird als lauter wahrgenommen, wenn seine Gesamtenergie über mehrere kritische Bandbreiten verteilt ist als wenn die Schallernergie innerhalb eines kritischen Bands konzentriert ist. Bleibt die Bandbreite eines Schallsignals insgesamt unterhalb des kritischen Wertes, ist die Lautstärkeempfindung unabhängig von der Bandbreite. Innerhalb einer kritischen Bandbreite hängt der Gesamteindruck der Lautstärke von der Summe der Reizintensitäten ab. Wenn die Komponenten des Schalls weiter als eine

¹¹⁴ Vgl. Moore (1995), S. 196

¹¹⁵ Vgl. Pickles (1988), S. 261

¹¹⁶ Vgl. Zwicker/Fastl (1999)

kritische Bandbreite auseinander liegen, ergibt sich der Gesamteindruck der Lautstärke aus der Summe der empfundenen Lautstärken der einzelnen Komponenten.

Für ein Beispiel von zwei Sinustönen mit gleichem Pegel wirkt sich dies folgendermaßen aus: Liegen die Frequenzen der beiden Töne weiter als eine kritische Bandbreite auseinander, wird der Zusammenklang als doppelt so laut wahrgenommen wie jeder der beiden Töne einzeln. Liegen die Frequenzen innerhalb einer kritischen Bandbreite, wird der Zusammenklang als so laut wahrgenommen wie ein einzelner Ton mit einer Schallintensität entsprechend der Gesamtintensität der beiden Töne. Der Eindruck der Gesamtlautstärke ist deutlich geringer als im ersten Fall (vgl. Abschnitt 2.3).¹¹⁷

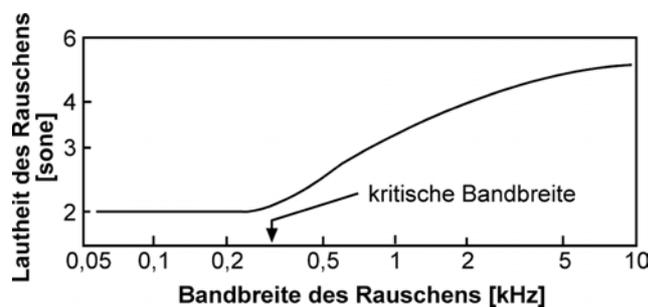


Abbildung 29: Lautheit eines Bandpass-Rauschens in Abhängigkeit von der Bandbreite
Für ein Rauschsignal mit konstanter Gesamtenergie. Mittenfrequenz = 2 kHz, SPL = 47 dB.¹¹⁸

Eindrucksvolle Ergebnisse zur Abhängigkeit der Gesamtlautheit von der Bandbreite liefert beispielhaft ein Experiment von Zwicker. Die Bandbreite eines Rauschens mit konstanter Energie wird dabei Schritt für Schritt verändert. Die Gesamtenergie bleibt jedoch konstant, so dass bei größerer Bandbreite die Energiedichte des Stimulus geringer ist. Das Ergebnis ist in Abbildung 29 verdeutlicht: Die Lautheit des Rauschbands mit einer Mittenfrequenz von 2 kHz bleibt bis zu einer Bandbreite von 300 Hz konstant, was der kritischen Bandbreite bei dieser Frequenz entspricht. Bei größeren Bandbreiten steigt die Lautheitsempfindung kontinuierlich. Sie kann bis zur dreifachen Stärke anwachsen, wenn das Rauschband auf sehr große Bandbreiten ausgedehnt wird.¹¹⁹

6.2 Hörschwelle komplexer Signale

Die absolute Hörschwelle eines komplexen Stimulus wird durch dessen Gesamtintensität bestimmt, solange alle Komponenten innerhalb der kritischen Bandbrei-

¹¹⁷ Vgl. Moore (1997), S. 93

¹¹⁸ Nach Zwicker/Fastl (1999), S. 157

¹¹⁹ Vgl. Zwicker/Fastl (1999), S. 157

te liegen. Dies lässt sich unmittelbar mit dem Prinzip der Summation der empfundenen Gesamtlautstärke in Verbindung bringen (vgl. Abschnitt 6.1). Das kann sich dahingehend auswirken, dass zwei (oder mehr) Schallkomponenten mit nahe beieinander liegenden Frequenzen selbst dann gehört werden können, wenn jede einzelne unterhalb der Hörschwelle liegt, da sich ihre Schallenergien aufsummieren.¹²⁰ Auch anhand von Hörschwellenmessungen können kritische Bandbreiten direkt gemessen werden.

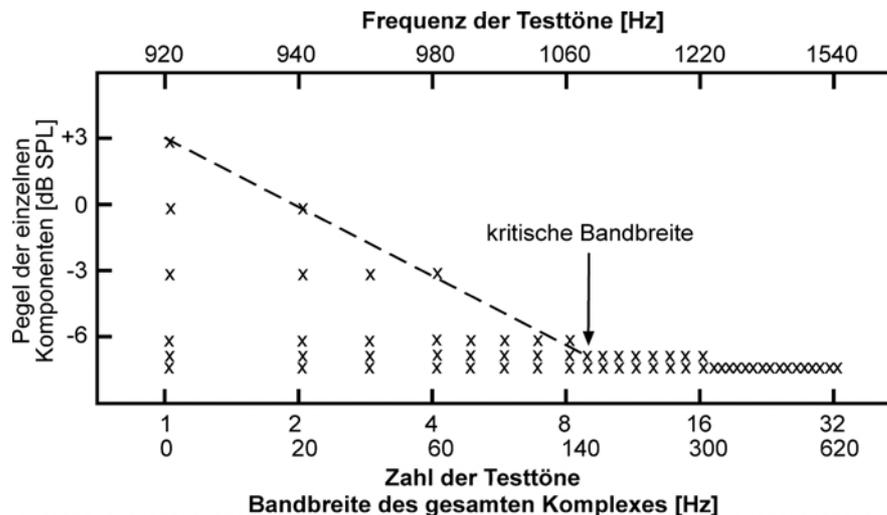


Abbildung 30: Absolute Hörschwelle für einen Komplex von Sinustönen

Die farbigen Punkte markieren einzelne Töne der dargebotenen Komplexe von Sinustönen. Die gestrichelte Linie zeigt die Hörschwelle für den gesamten Komplex. Ab dem 9. Ton verringert sich die Hörschwelle nicht mehr. Dies wird als Maß für die kritische Bandbreite betrachtet, sie beträgt hier etwa 160 Hz (in der Grafik markiert durch den Pfeil).¹²¹

Zwicker beschreibt dies für einen Komplex von Sinustönen gleicher Amplitude.¹²² Zunächst wird die Hörschwelle für einen einzelnen Testton bei 920 Hz gemessen. Nach und nach kommt jeweils ein weiterer Testton mit gleicher Amplitude und jeweils 20 Hz höherer Frequenz hinzu (940, 960, 980,... Hz). Jedes mal wird die Hörschwelle für den gesamten Komplex neu gemessen. Zwicker stellte fest, dass die Hörschwelle mit zunehmender Zahl der Testtöne absinkt, und zwar um 3 dB pro Verdopplung der Anzahl der Testtöne. Jedoch setzt sich dies ab einer bestimmten Zahl von Tönen - gleichbedeutend mit einer bestimmten Bandbreite - nicht fort. Im beschriebenen Experiment ist dies ab dem 9. Ton der Fall. Die Bandbreite des Stimulus beträgt bei 9 Tönen 160 Hz. Dies entspricht der kritischen Bandbreite bei 1 kHz nach Zwickers Bark-Skala. Zwickers Ergebnisse konnte in neueren Studien allerdings nicht

¹²⁰ Vgl. Zwicker/Fastl (1999), S. 151

¹²¹ Nach Zwicker/Fastl (1999), S. 151

¹²² Vgl. Zwicker/Fastl (1999), S. 151

vollständig bestätigt werden. Neuere Daten zeigen keinen mit den kritischen Bandbreiten korrespondierenden *Bruch*. Die Schwellen werden lediglich monoton höher.¹²³

6.3 Maskierungshörschwelle und bandbegrenztes Rauschen

Das so genannte *Band Widening-Experiment* wurde von Fletcher neben der *Critical Ratio-Berechnung* als zweite, direkte Methode zur Messung der kritischen Bandbreite beschrieben.¹²⁴ Das Experiment verdeutlicht das Auftreten von kritischen Bandbreiten bei einem weiteren zentralen Aspekt der Hörwahrnehmung, der Maskierung. Zur Verdeckung eines Testtones trägt nur der Teil der Schallenergie eines Maskierungssignals bei, der innerhalb einer kritischen Bandbreite in der Nähe von dessen Frequenz liegt.

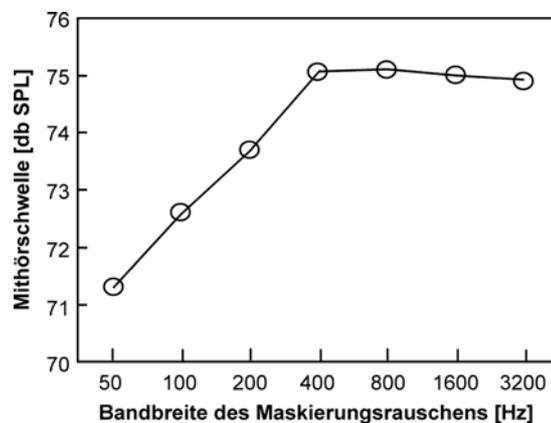


Abbildung 31: Maskierungshörschwelle und bandbegrenztes Rauschen
Frequenz des Sinussignals und Mittenfrequenz des Rauschens bei 2 kHz.¹²⁵

Fletcher maß die Mithörschwelle für einen Sinuston bei gleichzeitig dargebotenem bandbegrenztem Rauschen. Die Mittenfrequenz des Rauschbandes entsprach der Frequenz des Sinustons. Die Energiedichte des Rauschens wurde konstant gehalten, seine Bandbreite jedoch variiert. Auf diese Weise nahm die Gesamtenergie des Rauschens mit größerer Bandbreite zu. Fletcher stellte fest, dass sich die Mithörschwelle für den Testton mit der Bandbreite des Rauschens zunächst bis zu einer bestimmten Grenze erhöht. Weitere Vergrößerung der Bandbreite darüber hinaus führte jedoch zu keiner Erhöhung der Wahrnehmungsschwelle mehr (Abbildung 31).¹²⁶

¹²³ Vgl. Moore (1997), S. 96

¹²⁴ „Direkt“ im Gegensatz zur indirekten Bandbreitenberechnung mittels der Critical Ratio-Methode (vgl. Abschnitt 5.1.2). Vgl. Fletcher (1940)

¹²⁵ Nach Moore (1995), S. 163

¹²⁶ Eine ausführlichere Zusammenfassung der Experimente Fletchers und seiner Folgerungen findet sich bei Allen (1995).

6.4 Two-Tone Masking

Ein weiteres Maskierungsexperiment, das zur Messung der kritischen Bandbreite dienen kann, ist als *Two-Tone Masking* bekannt.¹²⁷ Zwicker maß die Wahrnehmungsschwelle für ein bandbegrenztetes Rauschen bei gleichzeitiger Darbietung von zwei Sinustönen gleichen Pegels mit gleichem Abstand unter- und oberhalb von der Mittenfrequenz des Rauschens (Abbildung 32 (b)). Die Bandbreite des Rauschens entsprach der kritischen Bandbreite bei dessen Mittenfrequenz. Zwicker variierte den Frequenzabstand zwischen den beiden Sinustönen und stellte fest, dass die Wahrnehmungsschwelle für das Rauschen konstant blieb, bis der Abstand einen kritischen Wert überschritt, an der die Schwelle abrupt sank (Abbildung 32 (a)). Er interpretierte diesen Wert als ein Maß der kritischen Bandbreite.

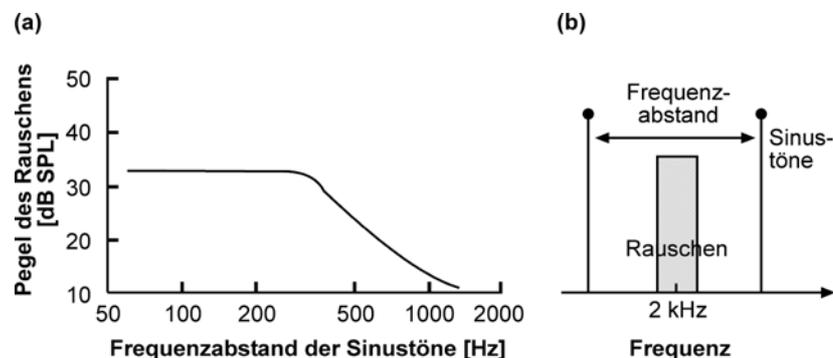


Abbildung 32: Two-Tone-Masking¹²⁸

Das Experiment lässt jedoch auch andere Interpretationen zu. Es ist wahrscheinlich, dass die dargebotenen Signale miteinander interagieren; die gemessene Wahrnehmungsschwelle könnte somit von der Wahrnehmung von Kombinationsprodukten der verwendeten Signale abhängig sein. Wenn experimentelle Vorkehrungen zur Verhinderung solcher Kombinationsprodukte getroffen werden, tritt kein abruptes Absinken der Wahrnehmungsschwelle auf. Diese nimmt vielmehr kontinuierlich mit größer werdendem Frequenzabstand ab.¹²⁹

6.5 Phasen- und Modulationseffekte

Als weitere Methode zur Messung der kritischen Bandbreiten schlug Zwicker ein Experiment zur Wahrnehmungsschwelle von Amplitudenmodulationen (AM) sowie Frequenzmodulationen (FM) eines Trägersignals in Abhängigkeit von der Modulation

¹²⁷ Vgl. Zwicker/Fastl (1999), S. 153 ff.

¹²⁸ Aus Zwicker/Fastl (1999), S. 153

¹²⁹ Zu Literaturangaben über neuere Studien zu Two-Tone-Masking vgl. Moore (1997), S. 96-97

onsfrequenz vor.¹³⁰ Bei gleicher Trägerfrequenz und gleicher Modulationsfrequenz besteht im Spektrum kein Unterschied zwischen AM und FM, außer in der *Phase* der durch die Modulation entstehenden Seitenbänder. Daher müssen wahrnehmbare Unterschiede zwischen AM und FM Signalen unter den genannten Bedingungen auf die Wahrnehmbarkeit von Phasenunterschieden zurückzuführen sein.

Für geringe Modulationsfrequenzen sind die Schwellenwerte für AM und FM sehr unterschiedlich, nähern sich mit wachsender Frequenz jedoch an, bis sie bei einem bestimmten Wert gleich sind. Bei der von Zwicker benutzten Trägerfrequenz von 1 kHz war dieser Wert 80 Hz; er wird als kritische Modulationsfrequenz (CMF) bezeichnet. Zwicker nimmt an, dass das Gehör nur für Phasenunterschiede empfänglich ist, wenn alle Komponenten des modulierten Signals innerhalb einer kritischen Bandbreite liegen. Er schlägt vor, dass die CMF mit der Hälfte der kritischen Bandbreite korrespondiert; d.h. die CMF wurde als erreicht angesehen, wenn die Bandbreite des gesamten Signals die kritische Bandbreite erreicht.

Moore (1997) weist auf Studien hin, die Zwickers Interpretation in Frage stellen. So korrespondiert die CMF statt mit der kritischen Bandbreite offenbar mit der Schwelle, an der die tiefste Komponente des Spektrums wahrnehmbar wird. Die Wahrnehmungsschwelle für diese tiefste Komponente hängt mehr von dem Filter ab, dessen Mittenfrequenz der Frequenz dieser Komponente entspricht, als vom Filter, der auf der Trägerfrequenz liegt. Des Weiteren scheint die Annahme, dass Phasenunterschiede nur zwischen Signalkomponenten innerhalb einer kritischen Bandbreite wahrnehmbar sind, nicht haltbar: Wenn die Pegel aller Komponenten reichlich über der Hörschwelle liegen, sind Unterschiede zwischen AM und FM-Signalen wahrnehmbar, deren Komponenten beträchtlich weiter auseinander liegen.¹³¹ Hartmann (1997) führt mehrere Arbeiten an, deren Ergebnisse dagegen sprechen, dass die Wahrnehmung von AM und FM auf einen einheitlichen, auf dem Signalspektrum beruhenden Prozess zurückzuführen ist, wie von Zwicker angenommen. Zumindest bei niedrigen Trägerfrequenzen, wo neuronale Phasenkopplung möglich ist (vgl. Abschnitt 4.3), und bei niedrigen Modulationsfrequenzen (unter 10 Hz) scheinen Mechanismen der neuronalen Synchronizität eine Rolle zu spielen.¹³²

6.6 Rauigkeit und Schwankungsstärke

Rauigkeit und Schwankung sind Hörempfindungen, die durch Fluktuationen von Schallsignalen ausgelöst werden. Die Rauigkeit hängt im unteren Frequenzbe-

¹³⁰ Vgl. Zwicker/Fastl (1999), S. 154 ff.

¹³¹ Zusammenfassung der Kritik an Zwickers Interpretation bei Moore (1997), S. 99-100

¹³² Zusammenfassend zur Wahrnehmung von AM und FM vgl. Hartmann (1997), S. 449 ff.

reich von der kritischen Bandbreite ab. Der Zusammenklang zweier Frequenzkomponenten hängt von der Frequenzdifferenz der Teiltöne ab. Innerhalb einer bestimmten Bandbreite verschmelzen die Töne zu einem einzigen Klangeindruck. Dessen Qualität verändert sich mit größer werdendem Abstand und durchläuft mehrere, ineinander übergehende Bereiche, bevor zwei deutlich getrennte Töne hörbar werden:

- Frequenzdifferenz 0-15 Hz: Einzelner Ton mit schwankender Amplitude (Schwebung). Bei 4 Hz ist der Eindruck der Schwankung am stärksten.
- Frequenzdifferenz 15 Hz - 70 Hz: Ein mit größerer Frequenzdifferenz zunehmend rauher Ton. Die Rauigkeitsempfindung ist bei etwa 70 Hz Modulationsfrequenz am stärksten.
- Frequenzdifferenz über 70 Hz: Zwei zunehmend getrennte Teiltöne; die Rauigkeit des Klangs nimmt ab.
- Frequenzdifferenz über 200-300 Hz: Der Eindruck der Rauigkeit verschwindet. Die beiden Teiltöne sind vollständig getrennt wahrnehmbar.

Rauhigkeit tritt nur auf, solange der Frequenzabstand innerhalb der kritischen Bandbreite liegt. Der Zusammenhang besteht so jedoch nur für Trägerfrequenzen im Frequenzbereich unterhalb von ca. 1 kHz. Oberhalb von 2 kHz geht die Obergrenze in einen konstanten Wert von 250-300 Hz über (Abbildung 33).

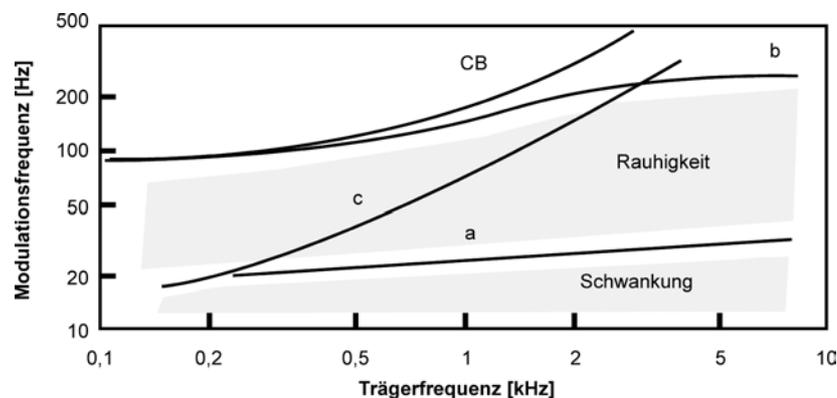


Abbildung 33: Existenzbereich von Rauigkeit und Schwankungsempfindung

CB kritische Bandbreite; a Grenze für Schwankungsempfindung; b Grenze für Rauigkeit; c Grenze für Aufspaltung in zwei einzelne Töne.¹³³

Das Rauigkeitsempfinden entsteht aufgrund des beschränkten Frequenzauflösungsvermögens der Basilarmembran. Bei geringem Frequenzabstand überschneiden sich dort die Erregungsmuster zweier Töne. Wenn der Frequenzabstand größer wird und man zwei getrennte Töne wahrnimmt, bedeutet dies, dass die Töne auf der Basilarmembran zwei unabhängige Frequenzmaxima hervorrufen. Die Obergrenze von 300 Hz für die Rauigkeitsempfindung sind offenbar eine Konstante des menschlichen Gehörs. Dort scheint die allgemeine Grenze für die Wahrnehmbarkeit von

¹³³ Nach Terhardt (1998), S. 294

Schallfluktuationen zu liegen. Der Grund dafür ist die Trägheit des neuronalen Systems.¹³⁴

6.7 Unterscheidbarkeit von Teilschwingungen komplexer Signale

Als ein weiteres Maß für die kritische Bandbreite wird die Fähigkeit des Gehörs zur Analyse von Teilschwingungen komplexer Signale betrachtet. Ein entsprechendes Experiment wird von Scharf (1970) beschrieben.¹³⁵ Darin werden komplexe Schallsignale aus zwölf Einzeltönen mit gleicher Lautheit angeboten. Die Einzeltöne waren in einer Variation des Experiments harmonisch verwandt; in einer anderen nicht. Probanden sollen ermitteln, welcher von zwei weiteren dargebotenen Tönen im Komplex enthalten war. Einer dieser Töne entspricht einer Teilschwingung; der andere liegt in der Mitte zwischen dieser Teilschwingung und der nächsthöheren. Als Maß für die kritische Bandbreite wird angenommen, dass die höchste identifizierbare Teilschwingung des Komplexes von der benachbarten Teilschwingung mindestens um eine kritische Bandbreite entfernt liegt. Ein Beispiel: Bei einer Grundschwingung von 250 Hz können 6,5 Teilschwingungen (in 75% aller Fälle) identifiziert werden; d.h. bis zu einer Frequenz von 1625 Hz ($6,5 \cdot 250$ Hz) genügt eine Frequenzdifferenz von 250 Hz zur Identifizierung.

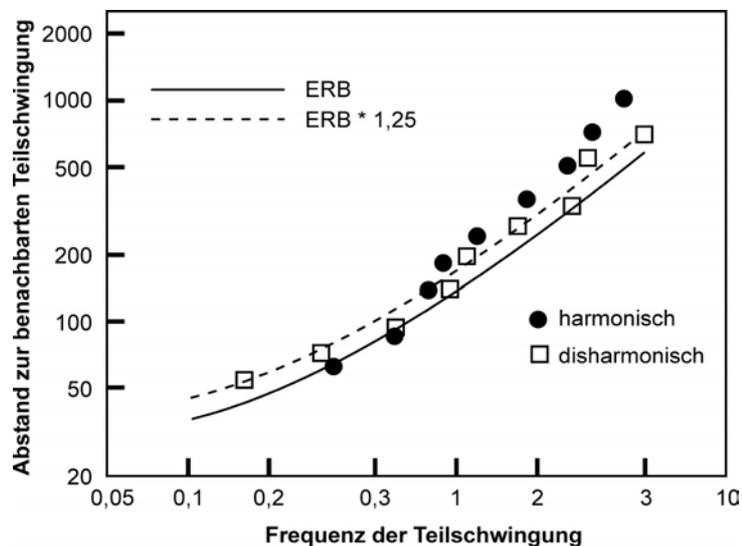


Abbildung 34: Unterscheidbarkeit von Teilschwingungen als Maß für kritische Bandbreiten¹³⁶

Daten aus diesem Experiment stimmen unterhalb von 1000 Hz mit den Werten der Bark-Skala nach Zwicker nicht überein. Multipliziert man jedoch die Werte nach

¹³⁴ Vgl. Terhardt (1998), S. 292 ff.; Hellbrück (1993); S. 117

¹³⁵ Vgl. Scharf (1970), S. 173-174

¹³⁶ Nach Moore (1997), S. 101

Moore (vgl. Abschnitt 5.3.2) mit dem Faktor 1,25, erhält man eine gute Deckung der ERB-Funktion mit Daten aus diesem Experiment. Weitere Experimente bestätigten, dass die Identifizierbarkeit von Teilschwingungen offenbar mit der ERB-Funktion korrespondiert. Das Frequenzauflösungsvermögen des Gehörs, das sich in den kritischen Bandbreiten widerspiegelt, reicht jedoch als alleiniges Erklärungsmodell nicht aus. In einem ähnlichen Versuch wie dem oben beschriebenen wurde eine deutlich bessere Leistung von Musikern im Vergleich zu Nicht-Musikern festgestellt. Die mögliche Deutung, dass Musiker engere kritische Bandbreiten haben, wurde durch vergleichende Maskierungsexperimente widerlegt. Daher scheint es wahrscheinlich, dass in diesem Zusammenhang ein weiterer Mechanismus eine Rolle spielt, den Musiker aufgrund ihrer Erfahrung besser einsetzen können.¹³⁷

7 Filterübergreifende und zeitliche Verarbeitungsprozesse

Dem Energiespektrum-Modell liegt die Annahme zugrunde, dass Hörer zur Analyse einzelner Signalkomponenten das Ausgangssignal eines auditiven Filters heranziehen. Damit ist die „stillschweigende Folgerung“ verbunden, dass die Ausgangssignale aller anderen Filter ignoriert werden. Dies ist jedoch nicht der Fall, wie die im Folgenden beschriebenen Versuche zeigen. Die dort auftretenden Phänomene belegen das Vorhandensein auditiver Verarbeitungsprozesse über mehrere Filter hinweg. Darüber hinaus ignoriert das Energiespektrum-Modell zeitliche Eigenschaften und Veränderungen des Signals wie Phasen, Amplitudenschwankungen und Kurzzeitfluktuationen in Maskierungsrauschen, die bei den genannten Effekten eine fundamental wichtige Rolle spielen.

7.1 Profilanalyse

Profilanalyse bezeichnet die Fähigkeit des Gehörs, Unterschiede in der Form des Spektrums zu entdecken. Das Gehör ist offenbar in der Lage, unabhängig von der Gesamtlautstärke eines Schalls dessen spektrale Zusammensetzung zu vergleichen. D.h. es existiert ein Vergleichsmechanismus, der erkennt, ob sich Pegel einzelner Komponenten des Schalls relativ zu den Pegeln der anderen Komponenten verändern. Experimentelle Hinweise dafür wurden Anfang der 80er Jahre beschrieben. Green (1988) führte eine Reihe von Experimenten durch, in denen die Wahrnehmbarkeit von Pegelanhebungen einer einzelnen Komponente im Spektrum eines komplexen Signals untersucht wurde. Probanden waren in der Lage, sehr geringe Pegelveränderungen einzelner Komponenten des Stimulus relativ zu den Pegeln der anderen

¹³⁷ Vgl. Moore (1997), S. 100-103

Komponenten wahrzunehmen. Um auszuschließen, dass die Wahrnehmung auf dem zeitlichen Vergleich des Ausgangssignals eines einzigen Filters beruht, wurde der Pegel des Gesamtsignals bei jeder Darbietung verändert. Die Unterschiedswahrnehmung muss daher auf der Analyse der Form des Spektrums basieren.

Die Existenz eines Mechanismus der Profilanalyse ist angesichts unserer Alltagserfahrung nicht überraschend: Die Unterscheidbarkeit von Klangfarben bzw. Klangqualität von Musikinstrumenten oder die Unterscheidung von Vokalen unabhängig von deren Gesamtlautstärke erscheint uns selbstverständlich. Diese Fähigkeit muss auf einem Vergleich der spektralen Form beruhen. Dies bedeutet, dass ein Vergleich der relativen Ausgangspegel mehrerer auditiver Filter unabhängig vom Gesamtpegel stattfindet, so dass die Annahmen des Energiespektrum-Modells nicht zutreffen können.¹³⁸

7.2 Comodulation Masking Release

Das Phänomen Comodulation Masking Release deutet darauf hin, dass die Mithörschwelle nicht nur von der Zusammensetzung des Energiespektrums von Testsignalen und "Störschall" abhängt. Nicht nur die Pegelverhältnisse und Frequenzabstände, sondern auch zeitliche Veränderungen des Störschalls spielen eine Rolle. Comodulation Masking Release bezeichnet einen Effekt, der zeigt, dass sich die Wahrnehmbarkeit eines Testtons in Breitbandrauschen verbessern kann, wenn die Rauschkomponenten gemeinsam moduliert werden. Der Effekt tritt auch ein, wenn sich die Rauschkomponenten in großem Frequenzabstand zum Testton befinden. Bedingung ist, dass die Modulationen in verschiedenen Frequenzbereichen miteinander zusammenhängen (*Co-Modulated Noise*). Ähnlich wie die Profilanalyse kann dieser Effekt mit dem Vergleich der Ausgangssignale mehrerer auditiver Filter erklärt werden. Er widerspricht somit ebenfalls den Annahmen des Energiespektrum-Modells.

Der Effekt wurde beispielhaft in einem Maskierungsexperiment nachgewiesen, bei dem die Mithörschwelle für ein Sinussignal in Abhängigkeit von der Bandbreite eines maskierenden Rauschhintergrunds gemessen wurde (Abbildung 35).

¹³⁸ Vgl. Moore (1997), S. 128

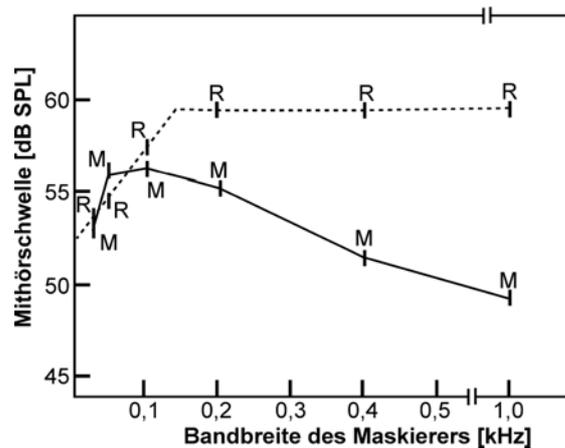


Abbildung 35: Mithörschwellen bei unmoduliertem und comoduliertem Rauschen

Die mit *R* gekennzeichneten Punkte sind die Mithörschwellen für ein 1 kHz Signal in einem bandbegrenzten, unmoduliertem Rauschen mit einer Mittenfrequenz von ebenfalls 1 kHz. Die mit *M* gekennzeichneten Punkte markieren die Mithörschwelle, wenn die Amplitude des Rauschens mit niedriger Frequenz moduliert wird.¹³⁹

Der Vergleich zwischen den Mithörschwellen für moduliertes und unmoduliertes Rauschen zeigt, dass die Mithörschwelle bei moduliertem Rauschen mit zunehmender Bandbreite sinkt. Der Effekt ist am stärksten, je mehr comodulierte Signalanteile vorhanden sind. Entscheidend scheint zu sein, wie viele auditive Filter Informationen über den Rauschhintergrund liefern.¹⁴⁰

7.3 Modulation Discrimination Interference

Modulation Discrimination Interference ist die Bezeichnung für eine Verschlechterung der Wahrnehmung von Veränderungen der Modulation einer (Test-) Trägerfrequenz, wenn gleichzeitig weitere modulierte (Maskierungs-) Signale in weitem Abstand des Testsignals dargeboten werden. Modulation Discrimination Interference ist ein Hinweis darauf, dass das Gehör unter bestimmten Umständen nicht in der Lage ist, das Ausgangssignal eines bestimmten auditiven Filters auszuwerten, obwohl dies für eine beabsichtigte Unterscheidungsleistung nützlich wäre. Die Existenz von Modulation Discrimination Interference widerspricht somit den Annahmen des Energiespektrum-Modells (vgl. Abschnitt 5.1.2). Der Effekt tritt auch dann auf, wenn der Abstand groß genug ist, dass sich die Maskierungssignale nicht mehr auf den auf der Trägerfrequenz liegenden auditiven Filter auswirken können. Er ist am stärksten, wenn die Maskierungssignale mit ähnlichen Raten moduliert werden, tritt jedoch bei breiten Variationen der Modulationsrate auf. Modulation Discrimination Interference wird mit einem Prinzip der perzeptiven Gruppierung in Verbindung ge-

¹³⁹ Nach Moore (1997), S. 122

¹⁴⁰ Vgl. Moore (1997), S. 122

bracht. Dieses besteht darin, dass diejenigen Komponenten eines komplexen Signals zusammengefasst werden, die gemeinsam moduliert werden oder gemeinsam beginnen bzw. enden. Dies entspricht dem subjektiven Höreindruck von Probanden in entsprechenden Versuchen.¹⁴¹

7.4 Auditive Objektwahrnehmung

Bei den bisher behandelten Phänomenen der Hörwahrnehmung und psychoakustischen Experimenten ging es um elementare Merkmale einzelner Schallsignale wie Tonhöhe und Lautstärkeempfindung. Diese Merkmale sind leicht auf physikalische Eigenschaften von akustischen Reizen zurückzuführen. In alltäglichen Hörsituationen werden diese Schalleigenschaften jedoch nicht derart isoliert wahrgenommen. Vielmehr nehmen wir eine komplexe akustische Umwelt wahr, in der wir oft mühelos und eindeutig Schall bestimmten Quellen und Ereignissen zuordnen und oft als bekannt identifizieren. Die Summe der Schallereignisse in unserer Umwelt stellt für uns eine akustische Szene dar. Die Erforschung der Hörwahrnehmung unter *gestaltpsychologischen* Gesichtspunkten ist ein relativ junger Ansatz, der einen Weg zur Erforschung solcher komplexerer psychoakustischer Phänomene weisen. Dort stößt der Ansatz der klassischen Psychophysik, die Hörwahrnehmung aus der Summe der Wahrnehmungen von einzelnen Elementen der Empfindung zu erklären, an seine Grenzen.

Der Vorgang der Analyse, durch den Frequenzanteile Quellen zugeordnet und in Schallereignisse zeitlich zu sogenannten Streams zusammengefasst werden, wird als auditive Gruppierung (Auditory Grouping) bezeichnet. Schallereignisse, die einer Quelle zugeordnet werden und im zeitlichen Verlauf einen Stream bilden, werden als akustische Objekte bezeichnet. Bei der Gruppierung akustischer Reize zu akustischen Objekten wirken offenbar eine Reihe von Prinzipien der perzeptuellen Organisation, nach denen das Gehör die physikalischen Eigenschaften der Reize analysiert. Die im Folgenden aufgeführten Prinzipien wurden zuerst im Zusammenhang mit der visuellen Wahrnehmung formuliert. Sie wurden an auditive Wahrnehmung angepasst. Nicht immer gelten alle Regeln im Einzelnen, doch führt eine auf ihnen beruhende Analyse insgesamt meist zur zutreffenden Interpretation der Reize.

¹⁴¹ Vgl. Moore und Shailer (1992); vgl. Hall et. al. (1995)

- Ähnlichkeitsprinzip: Töne, die einander in der Klangfarbe oder Tonhöhe ähnlich sind, werden als zusammengehörig wahrgenommen (z.B. Melodielinien, die von einem Musikinstrument gespielt werden).
- Gestaltgerechte Linienfortsetzung: Zeitliche Veränderungen in Tonhöhe, Spektrum, wahrgenommener Richtung der Schallquelle (Lokalisation) etc. sind meist kontinuierlich. Kontinuierliche Übergänge signalisieren Zusammengehörigkeit, abrupte Veränderungen deuten auf andere Quellen hin.
- Gemeinsames Schicksal: Veränderungen von Frequenzkomponenten, die von einer Quelle stammen, betreffen meist alle Frequenzkomponenten gemeinsam. Sie beginnen gemeinsam, verändern sich und enden meist gemeinsam.
- Gute Verlaufsgestalt: Wird eine Komponente des Schalls zeitweise maskiert, wird sie dennoch als kontinuierlich wahrgenommen, solange keine direkten sensorischen Hinweise vorliegen, dass sie unterbrochen wurde.

Die Prinzipien der perzeptuellen Organisation können Hinweise darauf liefern, welche physikalischen Eigenschaften akustischer Reize auch scheinbar einfache Bereiche der Hörwahrnehmung wie Maskierung und Lautstärkewahrnehmung beeinflussen.¹⁴²

8 Exkurs: Anwendungen des Filtermodells

8.1 Allgemeine Anwendungsbereiche

Das Filterverhalten ist eine grundlegende Eigenschaft des menschlichen Gehörs. Seine praktische Bedeutung schlägt sich daher in allen Bereichen nieder, die mit der darauf beruhenden frequenzabhängigen Lautheitssummation oder mit Maskierungseffekten im Zusammenhang stehen. Die Bark-Skala beispielsweise bildet die Grundlage für einen ISO-Standard der Lautheitsberechnung.¹⁴³ Neben den im den folgenden Abschnitten beschriebenen Anwendungen bei der Entwicklung von Hörhilfen und bei der digitalen Audiocodierung sind weitere allgemeine Anwendungsbereiche die Bewertung von technischen Schallübertragungssystemen, die Abschätzung von Sprachverständlichkeit, Bewertung von Lärmwirkung und Gebäudeakustik (z.B. bei Konzert- und Kinosälen).¹⁴⁴ In der Audio-Industrie werden 1/3-Oktavfilter, die den Bandbreiten der auditiven Filter nahe kommen, als Annäherung für das menschliche Gehör benutzt (vgl. Abbildung 27).

8.2 Mathematische Modelle

Mathematische Modelle der peripheren auditiven Filterung bilden die Grundlage der technischen Nutzung dieser Eigenschaft des menschlichen Gehörs. Es existie-

¹⁴² Vgl. Goldstein (1997), S. 163 ff; Goldstein (1997), S. 371 ff; Moore (1997), S. 245 ff.; vgl. Bregmann (1990)

¹⁴³ Vgl. Hartmann (1997), S. 73

¹⁴⁴ Vgl. Hartmann (1997); Moore (1997), S. 306 ff.

ren eine Reihe von Filtermodellen, die sich in der Grundstruktur gleichen. Sie bestehen aus einem vorgeschalteten Filter zur Nachbildung der Gehörgangsresonanzen und einer anschließenden Reihe paralleler Filter, die die Eigenschaften der Filterung in der Cochlea nachbilden. Sie bilden die Eingänge von darauf folgenden Verarbeitungskanälen, die den einzelnen Fasern des Hörnervs entsprechen. Die Umsetzung dieser Struktur unterscheidet sich je nach konkretem Modell. Die einfachste Beschreibung der Form der einzelnen Filter liefert die roex(p)-Funktion nach Patterson und Moore, die hier erwähnt wird, um das Prinzip zu verdeutlichen. Die typische Form der auditiven Filter mit runder Spitze und relativ steil abfallende Seiten lässt sich annähernd mit einer einfachen mathematischen Funktion beschreiben, der sogenannten gerundet-exponentiellen Funktion (*rounded-exponential, roex*):

$$|H(f)|^2 = (1 + pg)e^{-pg} \quad [5]$$

mit

$$g = \left| \frac{f - f_c}{f_c} \right|$$

f_c = Mittenfrequenz

p = Filterbandbreite

g = normalisierte Abweichung von der Mittenfrequenz f_c

Formel 5: roex(p) - Filter

Der Parameter p entspricht der kritischen Bandbreite. Nach ihm wird die Funktion als roex(p)-Filter bezeichnet. Sie ist die einfachste mögliche Funktion zur mathematischen Beschreibung der Filterform und reicht oft als Annäherung aus. Es existieren eine Reihe von mathematischen Anpassungen der roex(p)-Funktion, um Asymmetrien sowie Schwierigkeiten bei hohen Schallpegeln und bei Maskierungsschwellen nahe der absoluten Hörschwelle zu berücksichtigen.¹⁴⁵ Während die roex(p)-Funktion lediglich zur mathematischen Beschreibung einzelner Filter dient, gibt es eine Reihe von Modellen zur Beschreibung des Übertragungsverhaltens des gesamten peripheren auditiven Systems. Eine weit verbreitete Funktion ist die Gammatone-Filterfunktion nach Patterson (auch bekannt als *Patterson-Holdsworth Auditory Filter Bank*).¹⁴⁶

¹⁴⁵ Zusammenfassend zu roex(p) und verwandten mathematischen Filtermodellen vgl. Moore (1995), S. 168 ff.

¹⁴⁶ Vgl. Hartmann (1997), S. 254 ff.

8.3 Hörgeräte und Cochlea-Implantate

Erkenntnisse und Messmethoden der Psychoakustik stellen nicht nur eine Grundlage für die Untersuchung von Hörstörungen, sondern auch für die Entwicklung technischer Hilfen für deren Überwindung dar. So können in bestimmten Fällen von Schwerhörigkeit so genannte Multiband-Hörgeräte von Vorteil sein, die Schallsignale frequenzabhängig verstärken. Eine spektakulärere Rolle spielt die Filtertheorie bei der Entwicklung von so genannten Cochlea-Implantaten. In vielen Fällen von Taubheit oder extremer Schwerhörigkeit liegt die Ursache dafür in der Cochlea, während der Hörnerv mehr oder weniger intakt ist. In einigen Fällen ist es möglich, durch direkte Stimulation des Hörnervs der Patienten eine brauchbare Hörempfindung zu erzeugen. Dies wird durch Elektroden erreicht, die in die Hörschnecke implantiert werden. Es gibt verschiedene Ansätze zur Entwicklung der Implantate. Einer besteht darin, dem Hörnerv elektrische Signale zuzuführen, die den Signalen entsprechen, die eine intakte Cochlea erzeugen würde. Um dies zu erreichen, macht man Gebrauch von Bandpassfiltern, die die Filtereigenschaften des gesunden Ohrs so gut wie möglich simulieren. Bei der Entwicklung von Cochlea-Implantaten zeigten sich auch andere Ansätze, vor allem aber Kombinationen verschiedener Ansätze unter Einbeziehung des Filtermechanismus, erfolgreich.¹⁴⁷

8.4 Perzeptive Digitale Audiocodierung

Mit der zunehmenden Verbreitung der digitalen Audiodatenverarbeitung gelangten Verfahren zur Komprimierung der dabei anfallenden Datenmengen zu großer Bedeutung, da Speicherkapazitäten und Übertragungsraten vielfachen Beschränkungen unterliegen. Ausreichende Komprimierungsraten lassen sich dabei nur durch verlustbehaftete Verfahren erreichen, bei denen Teile der ursprünglichen Informationen verloren gehen. So genannte perzeptive Codierungsverfahren machen sich dabei die Filter- und Maskierungseigenschaften des menschlichen Gehörs zunutze. Die mit diesen Systemen erzeugten Signale erhalten objektiv beträchtliche Störungen und Verzerrungen, die jedoch vom gewünschten Signal verdeckt werden und dadurch die subjektive Qualität nicht beeinträchtigen. Die bekanntesten Beispiele sind die MPEG Audiokomprimierungsverfahren (z.B. MP3) und Minidisc, weitere sind z.B. DAB oder Dolby AC3.

Für die perzeptive Codierung wird das Signal zunächst in mehrere aneinandergrenzende Frequenzbänder aufgeteilt und anschließend digitalisiert. Der entscheidende Vorteil der Filterung liegt darin, dass es nicht notwendig ist, jedes Frequenz-

¹⁴⁷ Vgl. Moore (1997), S. 306 ff.

band mit der bei CD-Audio üblichen Auflösung von 16 Bit zu codieren. Liegt die Filterbandbreite in der Größenordnung der auditiven Filterbandbreite, ist nur eine geringe Bitauflösung (5-6 Bit) notwendig. Voraussetzung ist, dass der Pegel der einzelnen Frequenzbänder für kurze Zeitabschnitte (Blöcke von 1.5 – 12 ms) separat codiert wird, wofür nur wenige Daten benötigt werden. Meist sind die Frequenzbänder im unteren Frequenzbereich etwas größer als die auditiven Filterbandbreiten, da der Gebrauch schmalerer Filter zu Zeitverzögerungen führt. Diese Bänder müssen daher im Gegensatz zu den höheren Frequenzbereichen mit höheren Auflösungen als 5-6 Bit codiert werden. Eine weitere Reduzierung der Datenmenge kann erreicht werden, indem für jeden Zeitabschnitt (Block) eines Bands berechnet wird, wie weit das darin enthaltene Signal über der Maskierungsschwelle liegt, die von den angrenzenden Bändern verursacht wird. Die dem Band zugewiesene Auflösung ist umso höher, je weiter das Signal die Maskierungsschwellen überschreitet. Liegt das Band unter der berechneten Maskierungsschwelle, werden ihm keine Bits zugewiesen. Die Zuweisung der Bits ändert sich dynamisch von Block zu Block. Mit perzeptiven Codierern ist eine Reduzierung der Datenmenge um den Faktor 5-10 im Vergleich zu CD-Audio möglich, ohne dass für die große Mehrheit der Hörer Qualitätsverluste zu erkennen sind.¹⁴⁸

9 Abschlussbetrachtung Teil A

9.1 Anmerkungen zur Literatur

Die Recherche zu dieser Arbeit ergab in überraschendem Ausmaß Unterschiede, zum Teil auch Unklarheiten und Widersprüchen in der verwendeten Literatur. Widersprüche finden sich zur Frequenzselektivität der Basilarmembran und damit bei der entscheidenden Frage der Bewertung der Mechanismen der Frequenzanalyse durch das Gehör (vgl. Abschnitt 4). Verblüffendes ergibt sich beim Vergleich des Lehrbuchs von Zwicker/Fastl (1999) mit den Überblickswerken von Moore (1997), Hartmann (1997) und Terhardt (1998). Zwicker stützt sein Modell der Frequenzgruppen nach wie vor auf eigene experimentelle Studien der 50er und 60er Jahre. Neuere Ergebnisse werden ignoriert, obwohl die Werte der Bark-Skala seit längerem in Frage gestellt werden. Die Arbeiten von Zwicker haben jedoch offenbar nach wie vor großen Einfluss: Bark-Skala und Frequenzgruppenmodell finden sich ohne Erwähnung der Kritik z.B. bei Hellbrück (1993). Moore (1995; 1997) hingegen stützt sich auf Zwickers Arbeiten und diskutiert sie ausführlich; Hartmann (1997) und Terhardt (1998) stellen die Arbeiten beider dar. Große Unterschiede finden sich in der Schwerpunktsetzung

¹⁴⁸ Vgl. Moore (1997), S. 306 ff.

und Perspektive der Überblicksliteratur. Zwicker/Fastl (1999) betreiben „klassische Psychophysik“, d.h. sie behandeln Schritt für Schritt verschiedene Größen der Schallempfindung, für die sie Skalen und gesetzmäßige Zusammenhänge entwickeln. Der Schwerpunkt bei Terhardt (1998) und Hartmann (1997) liegt in der mathematisch-formelhaften Beschreibung von Schalleigenschaften und Entwicklung von Modellen der peripheren Schallverarbeitung. Hellbrück (1993), Moore (1997) verfolgen einen interdisziplinären Ansatz. Arbeiten mit dem Schwerpunkt auf der Gehörphysiologie wie Møller (2000) widmen sich ebenfalls der auditiven Filterung, beschränken sich jedoch auf deren Mechanismen und die neuronale Codierung von Schallsignalen, ohne auf die Hörwahrnehmung einzugehen.

Diese Ausgangssituation machte es schwer, die Rolle der kritischen Bandbreiten und der auditiven Filterung in den Mittelpunkt einer Übersichtsarbeit zu stellen und dabei die Arbeiten einzelner Autoren einzuordnen und zu bewerten. Die Ursache ist sicherlich zum einen in der Interdisziplinarität, zum anderen in der raschen Entwicklung der Forschung zu suchen. Dies ist der Grund für den großen Umfang der verwendeten Literatur. Dabei stellte sich heraus, dass als einzige einführende Arbeit Moore (1997) einen interdisziplinären Überblick bietet, der umfassend andere Perspektiven behandelt und die bis dahin verfügbaren neueren Erkenntnisse zur Gehörphysiologie konsequent einbezieht.

9.2 Der heutige Stellenwert des Filtermodells

Das Verständnis des Begriffs *kritische Bandbreiten* hat sich seit seiner ersten Erwähnung deutlich gewandelt. Ursprünglich war er rein empirisch definiert, als das Phänomen, dass sich die Hörwahrnehmung unterscheidet, je nachdem ob die Bandbreite eines akustischen Reizes einen kritischen Wert überschreitet oder nicht. Das Phänomen wurde bereits von Beginn an mit der Frequenzselektivität der Basilarmembran in Verbindung gebracht wurde. Jedoch herrschte aufgrund des damaligen Kenntnisstandes zur Gehörphysiologie Unsicherheit über diesen Zusammenhang, so dass eine strikte Trennung zwischen dem psychoakustischen Phänomen und dem Filtermodell des peripheren auditiven Systems aufrechterhalten wurde (vgl. Abschnitt 5.1.4). Dies wandelte sich mit den Erkenntnissen zur Frequenzselektivität des peripheren auditiven Systems. Das Phänomen der kritischen Bandbreite wird heute als Auswirkung der Filtereigenschaften des peripheren auditiven Systems betrachtet. Kritische Bandbreiten sind die effektive Bandbreite der auditiven Filter, deren Eigenschaften durch die Frequenzselektivität der Basilarmembran und der Hörnervenfasern bestimmt werden (vgl. Abschnitt 5.1.5).

Dieses Filtermodell ist ein grundlegender Bestandteil in allen Modellen der Hörwahrnehmung. Vor allem die Bedeutung der auditiven Filterung für Maskierung und Lautstärkewahrnehmung scheint am sichersten geklärt zu sein. Einige der psychoakustischen Experimente, die in der Vergangenheit als Methoden zur Messung der kritischen Bandbreite benutzt wurden, werden aus heutiger Sicht jedoch anders interpretiert. Vor allem der Einfluss von zeitlichen Veränderungen akustischer Reize spielt dabei eine entscheidende Rolle (vgl. Abschnitt 6;7). Für viele technische Anwendungen, bei denen vor allem Maskierungseffekte und Lautstärkewahrnehmung berücksichtigt werden müssen, reicht die Modellierung der heute bekannten Filtereigenschaften des peripheren Systems aus (vgl. 8).

In Bezug auf die Rolle der Filterung für die Wahrnehmung insgesamt hat sich der Fokus der Forschung verschoben. Heute wird angenommen, dass die Filtermechanismen eine Vorverarbeitung des Schalls bewirken und die zentraleren Prozesse der Schallanalyse und –interpretation auf der peripheren Frequenzanalyse aufbauen. Das Filtermodell basiert auf einer unteren Stufe des Systems der auditiven Verarbeitung. Bei der Verarbeitung natürlicher Schallsignale spielen jedoch komplexe Operationen der zentralen Verarbeitung, des Vergleichs und der Interpretation von Schallsignalen offenbar bereits in scheinbar einfachen Zusammenhängen eine wichtigere Rolle als früher angenommen (vgl. Abschnitt 7.4).

Teil B: Anwendungsplanung und -Entwicklung

10 Grundlagen Multimedia-Lehranwendungen

10.1 Warum Lehren und Lernen mit Software?

Seit den frühesten Versuchen, Computer zur Wissensvermittlung einzusetzen, wurden – vor allem in der Erziehungswissenschaft - exzessive Diskussionen über den Sinn und die (möglicherweise gefährlichen) Auswirkungen des Einsatzes geführt. Seither haben sich jedoch mit den Möglichkeiten und der allgemeinen Akzeptanz der Computertechnologie auch die Motivation und die Art des Einsatzes von Lehr- und Lernsoftware grundlegend gewandelt. So ermöglichen die heutigen Multimedia-Fähigkeiten auch vergleichsweise preisgünstiger Computer Präsentationsformen, aufwändige Visualisierungen und Simulationen, die mit anderen Mitteln gar nicht realisierbar wären. Die verbesserten Hardwarekapazitäten erlauben darüber hinaus große Komplexität und Flexibilität der Software sowie Interaktion mit dem Computer. Der Begriff Lehr- und Lernsoftware ruft nicht mehr das Bild der „Lernmaschine Computer“ hervor, die den Lehrer ersetzt, sondern ist ein weit verbreitetes Mittel der individuellen Wissensaneignung, das zusätzlich zu anderen Angeboten zur Verfügung steht. Die Argumente für den Einsatz von Multimedia-Software für Lehre und Lernen können folgendermaßen zusammengefasst werden:

- Möglichkeit großer Anschaulichkeit durch Kombination verschiedener Medien (Text, Bild, Ton, Animation, Video)
- Möglichkeit der Simulation von Prozessen, die in der Realität nicht direkt beobachtbar oder beeinflussbar sind
- Berücksichtigung verschiedener Nutzereigenschaften und Nutzerinteressen durch Flexibilität der Software
- Interaktion erlaubt individuelle Steuerung und Bedienung durch Nutzer
- Interaktion erlaubt starke Einbindung der Nutzer
- Unabhängigkeit von der Person des Lehrenden, zeit- und ortsunabhängige Verfügbarkeit
- Vorhandene Lehrinhalte lassen sich beliebig oft gleichzeitig oder nacheinander nutzen

Diese Argumente werden mit konkretem Bezug zum vorliegenden Projekt in den folgenden Abschnitten wieder aufgegriffen und diskutiert.

10.2 Lernparadigmen und Vermittlungskonzepte von Lernsoftware

Lehr- und Lernsoftware orientiert sich an den Erkenntnissen der Lernpsychologie und Didaktik. Im 20. Jahrhundert erlangten vor allem drei Lerntheorien großen

Einfluss (in Klammern jeweils die „Blütezeit“ der entsprechenden Theorieschule): Behaviorismus (1920-1960), Kognitivismus (1960-1990) und Konstruktivismus (seit 1990). Die folgenden Anmerkungen zu diesen Theorien sollen dazu dienen, verschiedene Vermittlungskonzepte von Lehr- und Lernsoftware zu klassifizieren und eine Orientierung für die Entwicklung des eigenen Projekts zu bieten. Sie sind jedoch keinesfalls als eine erschöpfende Beschreibung des jeweiligen Konzepts gedacht; dies würde den Rahmen dieser Arbeit sprengen. Auch kann angesichts der Vielfalt der verschiedenen Arten des erlernbaren Wissens, der Lernstile und der Lernsituationen keine dieser Theorien allein vollständig richtig und angemessen sein; sei es auch nur für das Konzept einer Software-Lehranwendung.

Der *Behaviorismus* widmet sich allein der Analyse beobachtbaren Verhaltens. „Lernen ist das Bilden und Festigen von Reiz-Reaktions-Ketten“.¹⁴⁹ Lernprogramme nach der behavioristischen Theorie teilen Wissen in kleinste Einheiten, die präsentiert und abgefragt werden. Durch positive bzw. negative Rückmeldungen sollen dabei Fehler „wegtrainiert“ werden. Für die Vermittlung komplexerer Zusammenhänge ist dieses Vorgehen nicht geeignet. Der *Kognitivismus* untersucht psychische Prozesse der Wahrnehmung und Informationsverarbeitung. Lernen ist das Entwickeln und Verändern von Formen der Repräsentation kognitiver Strukturen. Lernziel ist demnach „der Aufbau einer dem Gegenstand adäquaten mentalen Repräsentation beim Lernenden“.¹⁵⁰ Darauf basieren Lernprogramme, die durch Simulationen und/oder flexible Präsentation von Informationen dem Lernenden helfen sollen, durch Einsicht zu einer Problemlösung zu gelangen. Nach der Theorie des *Konstruktivismus* wird Wissen „vom Lernenden aktiv in komplexen, realen Lebenssituationen konstruiert. Lernen bezeichnet nach konstruktivistischer Sichtweise diesen Prozess des Aufbaus von Erfahrungen.“¹⁵¹ Um der Komplexität realer Wissensbereiche und der notwendigen aktiven Rolle des Lernenden gerecht zu werden, spielen demnach Interaktivität und nicht-lineare Verknüpfungen eine entscheidende Rolle in entsprechenden Softwarekonzepten. Tabelle 1 gibt beispielhaft eine Klassifizierung verschiedener Arten von Lehr- und Lernsoftware wieder.

¹⁴⁹ Steinmetz (1998), S. 819

¹⁵⁰ Steinmetz (1998), S. 820

¹⁵¹ Ebd.

Lernprogramme (steuern den Fortgang des Lernens)

- **Drill- und Übungsprogramme**
Festigung von zuvor erworbenem Wissen durch Wiederholung und Anwendung.
Beispiele: Vokabeltrainer, Übungsprogramme für Mathematik, Aufgabensammlungen.
- **Test-Software**
Dient der automatisierten Lernkontrolle und Leistungsüberprüfung, zur Entlastung von Prüfern oder zur Kontrolle des Selbststudiums. *Beispiele:* Verschiedene Multiple-Choice-Tests.
- **Tutorielle Systeme / Computer-Based Training (CBT)**
Übernehmen sämtliche Lehrfunktionen: Vermittlung des Wissens, Lernkontrolle, Rückmeldung über Lernfortschritte, Auswahl des Lernwegs in Abhängigkeit von gegebenen Antworten. *Beispiele:* Sprachkurse für das Selbststudium.

Lernumgebungen (haben Angebotscharakter; der Lerner steuert das Lernen selbst)

- **Hypertext-Umgebungen**
Sind eine Form der Realisierung elektronischer Nachschlagewerke. Enthalten auf vielfache Weise miteinander verknüpfte inhaltliche Elemente. Dienen zum Recherchieren von Informationen, setzen Vorwissen und selbständiges Lernen voraus.
- **Animationen und Simulationen**
Veranschaulichen dynamisch Sachverhalte, die in der Natur, Gesellschaft oder Wirtschaft (je nach Themenbereich) so nicht verfügbar oder direkt beobachtbar sind. Der Lernende kann ggf. den Ablauf beeinflussen und die zugrunde liegenden Zusammenhänge und Wirkungsmechanismen erkennen. Simulationen sind oft eingebettet in umfangreiche inhaltliche Vorbereitung.
- **Problemlösungsumgebungen**
Lernende werden von der Software mit einer komplexen Aufgabe konfrontiert; Problemlösungsumgebungen dienen der Ausbildung von Handlungsstrategien und Problemlösefähigkeiten.
- **Lernspiele und Edutainment**
Der Wissenserwerb wird durch unterhaltsame Präsentation motiviert; z.B. durch Wettbewerbselemente oder besondere Gestaltung.

Tabelle 1: Lernsoftware¹⁵²

¹⁵² Steinmetz (1998), S. 815; Schwier und Misanchuk benutzen die Termini „Prescriptive“ und „User-Controlled Environment“; vgl. Schwier/Misanchuk (1993). Die Zahl der Kategorien und die Terminologie variieren in der Literatur; durch die Anmerkungen sollte die Einteilung jedoch mit anderen vergleichbar sein.

10.3 Zum Vorgehen bei Anwendungsplanung und -entwicklung

Die Vielfalt und die Komplexität von Multimedia-Lehranwendungen machen es sinnvoll, ihre Entwicklung zu systematisieren. Dabei kann darauf zurückgegriffen werden, dass Multimedia-Systeme Software-Systeme sind – „es gelten daher auch die elementaren Grundregeln der Software-Entwicklung“.¹⁵³ Bei Multimedia-Lehranwendungen sind jedoch darüber hinaus auch Didaktik und Psychologie sowie besonders hohe Anforderungen an Gestaltung und Benutzerfreundlichkeit zu beachten.

Die Entwicklung des vorliegenden Projekts orientiert sich an einem *Phasenmodell* der Softwareentwicklung.¹⁵⁴ Es gliedert sich in *Planungs- und Definitionsphase*, *Entwurfsphase* sowie *Implementierungsphase*. Diese Phasen sind in der Praxis nicht immer trennbar. Insbesondere Tests von Zwischenergebnissen führen oftmals zu einer Schleife aus Entwurf, Implementierung und Test von Prototypen. Das Phasenmodell hat den Vorteil, dass einzelne Phasen mit Zwischenergebnissen abschließen, die jeweils aufeinander aufbauende Schritte hin zur Fertigstellung der Anwendung darstellen.

Die *Planungs- und Definitionsphase* beginnt mit einer Analyse der Ausgangssituation. Ist der Inhalt für die Umsetzung in eine Multimedia-Anwendung geeignet, und eignet sich eine Multimedia-Anwendung für den geplanten Einsatz bzw. die angestrebte Zielgruppe? Daran schließt eine grobe, aber eindeutige Formulierung von Zielen und Anforderungen an, auf deren Grundlage die technische Durchführbarkeit des Projekts beurteilt wird. In der *Entwurfsphase* wird das Design der Anwendung entwickelt. Der Begriff Design bezieht sich nicht nur die grafische Gestaltung, sondern betrifft auch die Struktur. Bei Multimedia-Lehranwendungen betrifft der Entwurf der Struktur zwei Ebenen: die Inhalte der Lehranwendung sowie die Programmarchitektur der Software. Ergebnis der Entwurfsphase sind die Design-Spezifikationen. Diese enthalten detaillierte Anweisungen zur Gestaltung der Benutzeroberfläche. In der *Implementierungsphase* werden schließlich die Komponenten des Entwurfs in eine nutzbare Anwendung umgesetzt. Der entstehende Quellcode ist Bestandteil der Projektdokumentation.

Ein wichtiger Bestandteil des Entwicklungsprozesses ist die *formative Evaluation*, d.h. das Testen und die Bewertung von Teilabschnitten der Umsetzung, die zu Veränderungen in den Anforderungen oder einzelnen Design-Spezifikationen führen kann. Je nach Umfang des gesamten Projekts, nach den verfügbaren technischen,

¹⁵³ Vgl. Holzinger (2001), S. 23

¹⁵⁴ Vgl. Holzinger (2001)

personellen bzw. finanziellen Mittel und nach den Anforderungen an die Qualität des Endprodukts können sich Art und Umfang formativer Evaluation unterscheiden.¹⁵⁵ Über die grobe Phaseneinteilung hinaus stützt sich die weitere Anwendungsentwicklung auf einen Ansatz von Schwier und Misanchuk (1993): Das grobe Gerüst eines Phasenmodells wird dort speziell für die Entwicklung interaktiver Multimedia-Lehranwendungen im Detail ausgearbeitet. Im Mittelpunkt stehen dabei stets die Inhalte der Anwendung, das Ziel der Vermittlung sowie die konsequente Berücksichtigung der Nutzerperspektive bei der Gestaltung.¹⁵⁶

11 Anwendungsplanung

11.1 Prüfung der Anwendbarkeit einer Multimedia-Lösung

Zu Beginn der Anwendungsplanung steht die grundsätzliche Prüfung, ob angesichts des vorgegebenen Inhalts zum einen ein Problem der Wissensvermittlung vorliegt und zum anderen, ob dieses Problem mit einem Multimedia-Lehrprogramm lösbar ist. Die Bearbeitung von Teil A dieser Arbeit ergab, dass keine Literatur existiert, die die Theorie der auditiven Filterung und die Rolle der kritischen Bandbreiten zugleich einführend und umfassend darstellt und dabei aktuelle wissenschaftliche Befunde berücksichtigt (vgl. Teil A, Abschnitt 9). Daher wird die erstgenannte Voraussetzung als erfüllt betrachtet. Teil A zeigt darüber hinaus, dass das gegebene naturwissenschaftliche Thema nicht zu komplex oder abstrakt ist, sondern in Teilaspekte gegliedert und in strukturierter, anschaulicher Form behandelt werden kann. Daher scheint es als Gegenstand einer Multimedia-Lehranwendung grundsätzlich geeignet. Über die bloße Eignung hinaus ist abzuwägen, ob das Thema nicht mit weniger aufwendigen Mitteln als einer Multimedia-Lehranwendung behandelt werden kann. Offenbar ist es möglich, große Teile der wichtigen Sachverhalte in der Form von illustriertem Text anschaulich darzustellen. Allerdings spricht die Beschaffenheit einiger zentraler Aspekte des Themas für den Multimedia-Einsatz: Zum einen werden Phänomene der Psychoakustik behandelt, so dass das Angebot von Hörbeispielen nahe liegend ist. Einigen Lehrbüchern zum Thema sind aus diesem Grunde CDs beigelegt. Zum anderen werden eine Reihe von physiologischen Prozessen beschrieben, die anhand von Animationen veranschaulicht werden könnten.

Neben der Eignung des Inhalts muss geprüft werden, ob eine Multimedia-Lehranwendung auch angesichts von Zielgruppe und geplantem Einsatzgebiet prinzipiell anwendbar ist. Als Zielgruppe kommen Studierende verschiedener Fachgebiete

¹⁵⁵ Vgl. Holzinger (2001), S. 224 ff.

¹⁵⁶ Schwier/Misanchuk (1993)

in Betracht, die das behandelte Thema berühren. Nahe liegend ist auch der Einsatz einzelner Multimedia-Elemente zu Demonstrationszwecken durch Lehrende. Diese Zielgruppe lässt ein hohes Maß von Vorbildung erwarten, die jedoch durch die unterschiedlichen fachlichen Hintergründe und den Spezialisierungsgrad der Nutzer unterschiedlich ausfallen kann (vgl. Abschnitt 1.2). Eine Lehranwendung, die den Anwendern die Kontrolle über Zugang zu den Inhalten und die Bearbeitungsweise überlässt, kann ein geeignetes Mittel sein, dieser heterogenen Zielgruppe gerecht zu werden.¹⁵⁷ Angesichts mehrerer in Frage kommender Fachbereiche und des speziellen Themas können darüber hinaus eine große Zahl und eine weite geografische Verteilung von möglichen Nutzern der Software erwartet werden, so dass auch diese Aspekte für den Einsatz einer Multimedia-Lösung sprechen.

11.2 Analyse existierender Anwendungen

Vor einer Entscheidung über die weitere Entwicklung soll überprüft werden, ob bereits Anwendungen mit ähnlicher Ausrichtung existieren und das Vorhaben ggf. überflüssig machen. Eine Internet-Recherche förderte eine Reihe von Anwendungen aus dem engeren Themenbereich zu Tage (vgl. Tabelle 2).

- **ars auditus** (http://www.dasp.uni-wuppertal.de/ars_auditus/)
Online-Tutorial mit Hörbeispielen zum Thema Gehör und Hören. Von Martina Kremer, Universität Wuppertal
- **Acoustics and Vibration Animations**
(<http://www.kettering.edu/%7Edrussell/Demos.html>)
Visualisierungen zu Grundlagen der Akustik von Dan Russell, Kettering University, Flint, MI
- **Basilarmembran-Wellen-Demonstrator**
(<http://www.physik3.gwdg.de/%7Estrube/bmjava/bmdemo.html>)
von Hans Werner Strube, Universität Göttingen
- **Cochlear Mechanics**
(<http://hearing.psychol.cam.ac.uk/Demos/demos.html>)
Einige Animationen, die dem Einfluss der äußeren Haarzellen auf die Basilarmembranschwingung zeigen. Von Stephen Neely, Universität Cambridge
- **An Auditory Tour**
(<http://www.neurophys.wisc.edu/aud/auditorytour.html>)
Sammlung von Grafiken und Animationen, die verschiedene Aspekte des peripheren auditiven Systems illustrieren. Von John Brugge, Dan Geisler, Joe Hind, David Rose und Steven Greenberg, University of Wisconsin-Madison

Tabelle 2: Übersicht über existierende Lehranwendungen

¹⁵⁷ Schwier und Misanchuk (1993)

Die Recherche erbrachte nur wenige Ergebnisse, die zudem zum Teil schon recht alt sind („An Auditory Tour“ z.B. ist von 1996). Es fällt auf, dass alle Anwendungen aus nicht-kommerziellen Projekten stammen. Es wurden keine kommerziellen Angebote (z.B. CD-ROMs) gefunden, die auch nur ungefähr das Thema abdecken. Für derartige Anwendungen existiert offenbar kein lohnender Markt. Eine Reihe von Anwendungen widmet sich Teilaspekten des Themas, wobei der Schwerpunkt mehr auf Prozessen der Gehörphysiologie liegt. Es wurde jedoch keine gefunden, die das Thema mit dem angestrebten Schwerpunkt und in der gewünschten umfassenden Form behandelt. Die genannten Anwendungen können hier nicht im Einzelnen analysiert werden. Sie wiesen allerdings z.T. Schwachpunkte auf, die in der weiteren Anwendungsplanung vermieden werden sollen (vgl. Tabelle 3). Darüber hinaus existieren noch einige Sammlungen von Matlab-Funktionen,¹⁵⁸ die jedoch nur für weit in der Materie (und in der Bedienung von Matlab) fortgeschrittene Anwender verständlich sind. Einige sollen dem Verständnis der Filtereigenschaften des Gehörs dienen; wegen der hohen Nutzungsvoraussetzungen sind sie jedoch nicht vergleichbar und werden daher nicht als Lehranwendungen aufgelistet.

- Web-Präsentationen genügten z.T. nicht den gängigen Design-Anforderungen (Screen-Design, Übersichtlichkeit d. Navigation)
- Integration von Multimedia-Elementen ist nicht gelungen bzw. technisch umständlich und konfigurationsabhängig (Abspielen von WAV-Samples in neuen Browser-Fenstern, Samples müssen erst heruntergeladen werden, Software muss erst installiert werden)
- Anwendung nicht selbsterklärend (einzelne Software-Simulation zum Herunterladen ohne begleitende Dokumentation, thematische Einbettung etc.)

Tabelle 3: Schwachpunkte existierender Anwendungen

11.3 Zielformulierung und technische Realisierbarkeit

Zusammenfassend wird bei der weiteren Entwicklung der Anwendung von folgenden Voraussetzungen ausgegangen:

- Es besteht Bedarf an einer verständlichen, umfassende Einführung in das Phänomen der kritischen Bandbreiten (vgl. Abschnitt 9.1)
- Die Zielgruppe besitzt Vorwissen, ist jedoch heterogen (unterschiedliche fachliche bzw. Nutzungsinteressen)
- Die Zielgruppe besitzt eine hohe Nutzungsmotivation (spezielles fachliches Interesse)
- Als Art der Nutzung ist eine gezielte eigenständige Nutzung – als Selbstlernumgebung, Nachschlagewerk oder für Demonstrationszwecke durch Lehrende zu erwarten
- Es existieren einige frei verfügbare Animationen zur Gehörphysiologie sowie mathematische Modelle zum Filtermodell des Gehörs im WWW (vgl. Abschnitt 11.2)

¹⁵⁸ z.B. Slaney (1993)

Auf diesen Voraussetzungen beruht die Zielformulierung für das vorliegende Projekt: Ziel ist die Entwicklung einer Multimedia-Lehrumgebung, die die Rolle der kritischen Bandbreiten für die auditive Wahrnehmung in den Mittelpunkt stellt. Sie soll

- einen umfassenden Überblick über dieses Thema bieten, einführenden Charakter haben und leicht verständlich sein
- Grundlagen berücksichtigen, die für das Verständnis unverzichtbar sind und angesichts der Zielgruppe nicht vorausgesetzt werden können
- die systematische, selbständige Erarbeitung des gesamten Inhalts anhand eines „Lehrpfads“ erlauben
- den gezielten Zugriff auf einzelne Elemente des Inhalts erlauben
- so gestaltet werden, dass sie intuitiv bedienbar ist und keine besondere Software oder Programmierkenntnisse erfordert
- auf möglichst vielen Systemen - unterschiedlichen privaten Rechnern und auf Hochschulrechnern – lauffähig sein.¹⁵⁹

Daher wird für das vorliegende Thema die Erstellung einer offenen Lernumgebung angestrebt. Auf eine Steuerung des Lernwegs, Abfrage- und Kontrollelemente wird verzichtet. Multimedia-Elemente sollen sinnvoll zur Veranschaulichung der Sachverhalte eingesetzt werden. Da für sie ein hoher Entwicklungsaufwand angenommen werden kann und bereits frei verfügbare Visualisierungen einiger physiologischer Prozesse der Schallverarbeitung existieren, wird der Schwerpunkt zunächst auf die Demonstration psychoakustischer Effekte durch Hörbeispiele gelegt. Weniger zentrale, aber für das Verständnis unverzichtbare Themenbereiche werden mittels illustrierten Texts behandelt.

Die genannten Anforderungen sind mit gängigen Internet-Technologien erfüllbar. HTML bietet ausreichende Möglichkeiten, um die Grundstruktur der Anwendung zu realisieren und große Teile des Inhalts darzustellen. HTML bietet darüber hinaus die Möglichkeit der Einbindung von Multimedia-Elementen wie interaktiven Animationen und benutzerkontrollierten Audio-Elementen in der Form von Macromedia Flash – Filmen. Mit Flash produzierte Multimedia-Elemente sind nahtlos in HTML-Dokumente integrierbar. Der entscheidende Vorteil einer Anwendung auf HTML – Basis ist, dass sie problemlos auf allen Rechnersystemen nutzbar ist, auf denen moderne WWW-Browser mit einem Flash-Plugin installiert sind. Dies ist für die bei weitem überwiegende Zahl der Systeme anzunehmen, für die der Einsatz der geplanten Anwendung zu erwarten ist. Ein weiterer Vorteil ist ein relativ geringer Entwicklungsaufwand, da geeignete Entwicklungswerkzeuge die Erstellung ohne Programmierung (unter Zuhil-

¹⁵⁹ D.h. die Anwendung muss auf möglichst vielen Systemen lauffähig sein (neben Windows auch Linux / Unix, das an Universitäten verbreitet ist sowie Mac OS, da viele Studierende Apple-Rechner nutzen) und geringe Anforderungen an die Hardwareleistung stellen (ggf. veraltete Universitätsrechner, wenig leistungsfähige private Rechner bei Studierenden).

fenahme einer Scriptsprache für interaktive Flash-Elemente) auf der Basis von ausgereiften grafischen Benutzeroberflächen erlauben. Ein weiterer Vorteil der HTML-Plattform ist, dass die Anwendung sowohl lokal auf einem Rechner gespeichert als auch von Datenträgern abgerufen als auch als Website veröffentlicht werden kann. Letzteres würde die Inhalte problemlos einem großen Nutzerkreis zugänglich machen und die Möglichkeit bieten, bereits vorhandene Angebote im WWW durch Links einzubinden.

Für die Gestaltung der einzelnen HTML-Dokumente und die Verwaltung des gesamten Projekts wird Macromedia Dreamweaver 4 eingesetzt. Das Programm bietet sowohl eine übersichtliche grafische Oberfläche, die schnelles und bequemes Arbeiten erlaubt, als auch die volle Kontrolle über den Quelltext. Es erlaubt ebenfalls den effizienten Einsatz von CSS¹⁶⁰ als Werkzeug zur Systematisierung und Organisation des Layouts. Die Hörbeispiele werden mit Macromedia Flash MX erzeugt.¹⁶¹ Flash ist wegen seiner grafischen Oberfläche ein leicht bedienbares Werkzeug und bietet darüber hinaus mittels einer eigenen, leicht verständlichen Scriptsprache die Möglichkeit der Realisierung interaktiver Elemente.

12 Anwendungsentwurf und Umsetzung

Die weitere Anwendungsentwicklung gliedert sich in zwei Bereiche: die Aufbereitung und Strukturierung der Lehrinhalte sowie die Gestaltung der Benutzeroberfläche, d.h. Layout und Navigation. Im Folgenden wird das Vorgehen skizziert und mit Beispielen aus der entstandenen Anwendung illustriert. Die Designspezifikationen und die Übersicht der einzelnen Komponenten des Inhalts, die die Grundlage für die praktische Umsetzung bildeten, sind im Anhang enthalten.

12.1 Inhalt und inhaltliche Struktur

In Anlehnung an die Gliederung von Teil A wurde eine Aufteilung des Inhalts auf mehrere, thematisch eng zusammenhängende Abschnitte (*Lehrmodule*) vorgenommen. Insgesamt decken die Abschnitte das Themengebiet vollständig ab. Innerhalb der Abschnitte erfolgt eine Aufteilung in inhaltliche Einheiten (*Kernkomponenten*), die die Basis für die Präsentation der Lehrinhalte bilden. Die inhaltlichen Einheiten behandeln jeweils abgeschlossen einen einzelnen Aspekt des Themas. Ihr Umfang und ihre Gestaltung ist auf Anschaulichkeit und Verständlichkeit ausgerichtet. Daher sind sie meist weniger umfangreich und oft anders formuliert als die entsprechenden

¹⁶⁰ Cascading Style Sheets

¹⁶¹ Sie werden als Flash-Filme für den Flash-Player 5 exportiert

Abschnitte aus Teil A. Sie enthalten die wichtigsten Aussagen in wenigen Sätzen Anhand der Kernkomponenten kann das gesamte Lehrmaterial in übersichtlichen Schritten erarbeitet werden. Eine Übersicht aller Komponenten des Inhalts bietet die *Site-map*, die in der Anwendung selbst als Orientierungshilfe zugänglich und im Anhang der Arbeit abgedruckt ist. Die inhaltliche Struktur spiegelt sich in Form der Navigation wieder und bietet so dem Nutzer ein hohes Maß an Transparenz. Einen Ausschnitt zeigt Abbildung 36.

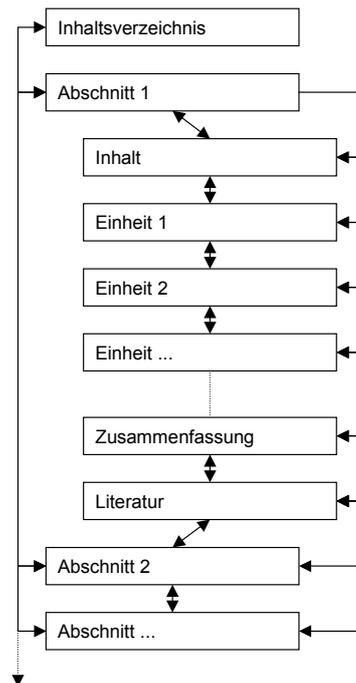


Abbildung 36: Ausschnitt aus der inhaltlichen Struktur der Anwendung

Die Pfeilverbindungen deuten die Verlinkung an. Aus jeder Seite des Inhalts heraus sind alle Abschnitte anwählbar. Innerhalb eines Abschnitts sind alle Seiten dieses Abschnitts anwählbar. Die aufeinander folgenden Abschnitte sind jeweils miteinander verbunden.

Die Gliederung der einzelnen Abschnitte enthält jeweils eine Einführung in den Inhalt. Am Ende des Abschnitts folgen jeweils eine Zusammenfassung des Inhalts sowie die Literatur- und Quellenangaben nebst kommentierten Verweisen auf weiterführende Literatur und ggf. WWW-Links.

12.2 Benutzeroberfläche: Layout und Navigation

Die Benutzeroberfläche definiert die Interaktionsmöglichkeiten des Anwenders mit der Software. Sie umfasst Elemente zur Navigation in den Inhalten sowie die Präsentation der Inhalte. Eine *benutzerfreundliche* Oberfläche ermöglicht den effizienten

Einsatz der Software zum vorgesehenen Zweck durch die angestrebte Zielgruppe von Nutzern. Der Aufwand für das Verständnis der Steuerung des Programms soll im Vergleich zum eigentlichen Zweck der Anwendung minimal sein.

Benutzerfreundliche Gestaltungskonzepte lassen durch die Einhaltung weniger grundlegender Prinzipien erreichen: Bewahrung von Einfachheit bzw. Reduzierung auf das Wesentliche; Klarheit und Eindeutigkeit; Konsistenz der Gestaltungsprinzipien und Funktionsmechanismen sowie konsequente Berücksichtigung der ästhetischen Konzepte Balance, Harmonie und Einheit.¹⁶² Diese Prinzipien werden sowohl auf die Anordnung der Bildschirminhalte als auch auf die grafische Gestaltung aller Elemente der Oberfläche angewendet. Nach mehreren Entwürfen wurden die Einzelheiten der endgültigen Form der Gestaltung in detaillierten Designspezifikationen. Diese bildeten die Vorschriften für die praktische Umsetzung aller Inhalte und gewährleisteten eine einheitliche Gestaltung der gesamten umfangreichen Anwendung. Designspezifikationen wurden für die *Bildschirmmaske*, für *Text/Typographyie*, *Navigationselemente*, *Abbildungen* sowie *Text/Inhalt und Stil* erstellt (siehe Anhang der Arbeit).

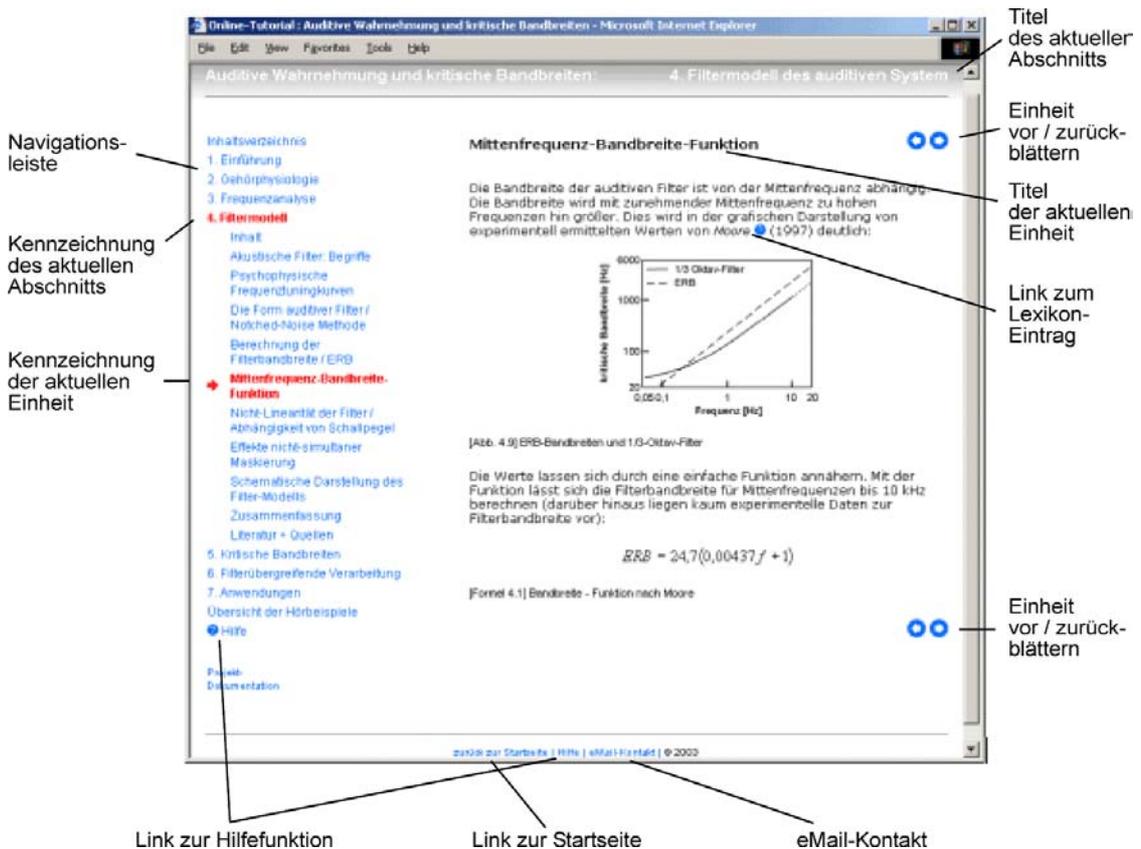


Abbildung 37: Die Benutzeroberfläche
Darstellung eines inhaltlichen Kernelements aus *Abschnitt 4: Das Filtermodell des Gehörs*.

¹⁶² vgl. Schwier/Misanchuk (1993), S. 212 ff.

12.3 Die Hilfe-Funktion

Die Möglichkeit von Querverweisen durch HTML-Links wurde genutzt, um eine Hilfe-Funktion zu realisieren, die in der gesamten Anwendung durch Platzierung von Links an festen Orten ständig präsent ist (vgl. Abbildung 37). An kritisch erscheinenden Stellen wurden zusätzliche Verweise angebracht (z.B. bei den Hörversuchen). Mit einem Klick auf den Hilfe-Link öffnet sich ein neues Browser-Fenster, das auf Erklärungen zu verschiedenen Bereichen und ein kleines Lexikon wichtiger Begriffe verweist und einen zusätzlichen eMail-Kontaktlink bietet.

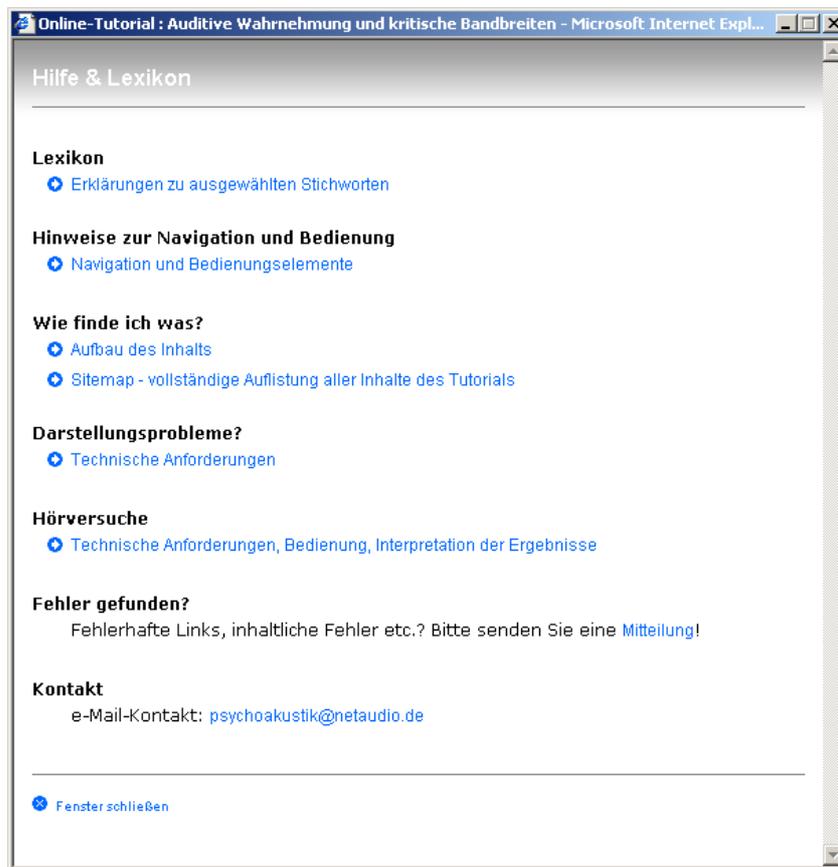


Abbildung 38: Inhalt der Hilfe-Funktion

Die Inhalte des Lexikons sind an zahlreichen Stellen direkt aus dem Text abrufbar und bietet kurze Erklärungen zu wichtigen Begriffen und ggf. Verweise auf ausführliche Erklärungen an anderer Stelle (vgl. Abbildung 39).

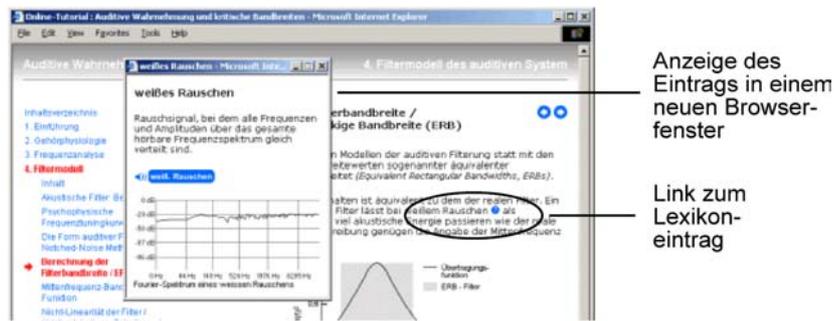


Abbildung 39: Aufruf der Lexikon-Funktion aus dem Text

12.4 Hörbeispiele und Hörversuche

12.4.1 Die verwendeten Audiosamples

Die Handhabung von Audiodaten durch Macromedia Flash schränkt die Möglichkeiten für die Entwicklung von Hörversuchen ein. Die Software bietet keine Möglichkeit, geeignete Testsignale mittels mathematischer Funktionen zu generieren. Daher mussten alle Testsignale extern erzeugt und als Samples in die Flash-Filme importiert werden. Sinustöne und weißes Rauschen wurden mit dem Signalgenerator von Steinberg Wavelab 4 erzeugt; die bandbegrenzten Rauschen mit der Auditory Toolbox, einer Sammlung von Matlab-Funktionen von Malcolm Slaney.¹⁶³ Die Verwendung von Samples bedeutet, dass die Versuche mit einer festgelegten Anzahl vorher definierter Signale auskommen müssen. Ihre Anzahl ist begrenzt, da Audiosamples große Datenmengen erzeugen. Es stellte sich jedoch heraus, dass die Hörversuche mit einer geringen Zahl geeigneter Samples auskommen. Eine Verringerung der Dateigröße der Hörversuche ließ sich erreichen, indem die *Sinustöne* von Macromedia Flash in hoher Qualität (128 kbps) MP3-codiert werden, was in diesem speziellen Fall (bei Sinustönen) keinen Einfluss auf die Klangqualität hat. Lediglich die *Rauschsignale* sind als WAV-Files mit 44 kHz Samplerate und 16 Bit-Auflösung in den Flash-Filmen enthalten. Dadurch ließ sich die Dateigröße in zwei Fällen stark reduzieren. Lediglich der Versuch zur effektiven Maskierungsbandbreite erreicht wegen der enthaltenen verschiedenen bandbegrenzten Rauschen eine Dateigröße von ca. 1 MB.

Aufgrund der technischen Voraussetzungen steht von vornherein fest, dass die Hörversuche auch nicht annähernd den Kriterien an wissenschaftliche psychoakustische Experimente genügen können. Angestrebt wird vielmehr, die ausgewählten Effekte *qualitativ* zu demonstrieren.

¹⁶³ Slaney (1993)

12.4.2 Zu Entwicklung und Test der Hörversuche

In der vorliegenden Form lieferten die Versuche bei den meisten Tests zufrieden stellende Ergebnisse. Der Hörversuch zur Rauigkeit erwies sich als relativ unkompliziert, was die Erzeugung des gewünschten Effekts anging. Die verschiedenen Hörempfindungen sind bei verschiedenen Arten der Klangwiedergabe deutlich wahrzunehmen, sowohl über Kopfhörer als auch Lautsprecher verschiedenster Qualität. Die Einschränkungen durch die Verwendung von Samples traten hier jedoch am deutlichsten zu Tage. Eine flexiblere und feinere Abstufung der Frequenzabstände oder fließende Frequenzübergänge wären wünschenswert, lassen sich aber mit Macromedia Flash bisher nicht realisieren.

Während der Erprobung der Versuche zur Frequenzabhängigkeit der Mithörschwelle und zur effektiven Maskierungsbandbreite bestätigte sich dagegen, dass die technischen Bedingungen der Wiedergabe bei sensiblen Hörversuchen, wie sie die Feststellung von Schwellen darstellen, einen deutlichen Einfluss auf das Ergebnis haben. Kopfhörer sind unbedingte Voraussetzung. Gute, geschlossene Kopfhörer und eine gute Soundkarte sind empfehlenswert. Ähnliche Ergebnisse wie in Abbildung 43 konnten zwar z.T. auch mit billigen Walkman-Kopfhörern oder mit einer hörbar schlechten Soundkarte eines Laptops erzielt werden. Bei der Kombination von Wiedergabegeräten schlechter Qualität häuften sich allerdings die „misslungenen“ Ergebnisse.¹⁶⁴

Einen weit größeren Einfluss hatten unterschiedliche Anweisungen an die Testpersonen und die Erwartungen an den Test. Letzteres wurde zunächst nach zahlreichen Selbstversuchen in der Anfangsphase deutlich. Diese zeigten sich nach kurzer Zeit als sehr begrenzt aussagefähig, was die anschließenden ersten Tests mit anderen Versuchspersonen ergaben. Insgesamt haben etwa ein Dutzend Personen mit sehr unterschiedlichen Vorkenntnissen und nach unterschiedlichen Einweisungen die Hörversuche getestet. Darunter waren mehrere Personen, die aufgrund von Vorkenntnissen aus verschiedenen Gründen mit den Klängen der Testsignale vertraut waren (Musiker, Veranstaltungstechniker, Nachrichtentechniker) sowie mehrere Personen ohne Vorkenntnisse. Die Tests, die aufgrund des geringen Umfangs sicher nicht repräsentativ sein können, ergaben, dass sich die angestrebten Demonstrationseffekte leicht erzielen ließen, wenn ein gewisses Maß an Vertrautheit mit den

¹⁶⁴ Als qualitativ hochwertigste Ausstattung standen Studio-Kopfhörer von audio-technica (ATH-M40fs) und eine Soundkarte von Creative (Soundblaster Live) zur Verfügung. Weitere Test wurden u.a. mit einem Gericom-Laptop der untersten Preisklasse durchgeführt (hörbar schlechte Soundkarte, die das weiße Rauschen stark verzerrte) sowie mit verschiedenen qualitativ schlechteren Kopfhörern (u.a. Philips SBC HP 100, geschlossene Kopfhörer, um 15 Euro; Sony Walkman-Kopfhörer zum Tragen im Ohr, mitgeliefert bei einem Kassetten-Walkman um 50 Euro)

Testsignalen vorhanden war. Bei Versuchspersonen ohne Vorkenntnisse schwankten die Ergebnisse und waren offensichtlich sehr leicht durch unterschiedliche Einweisungen zu beeinflussen. Unabhängig von den Vorkenntnissen erwies sich als hinderlich, wenn die Versuchspersonen eine ausgeprägte Erwartungshaltung hatten oder den Versuchsaufbau und das Ziel sehr genau kannten. Besondere Schwierigkeiten bereitete die Definition des Testtons als deutlich hörbar oder nicht hörbar. Die Versuchspersonen mit Vorkenntnissen entschieden sich offensichtlich nach festen Kriterien entsprechend ihrer Vorstellung eines „deutlichen Sinustons“, während Versuchspersonen ohne Vorkenntnisse bereits schwache Veränderungen des Rauschens als Einsetzen des Testtons interpretierten.

12.4.3 Die Hörbeispiele

Neben den Hörversuchen wurden einige einfache Hörbeispiele eingefügt. Sie dienen lediglich zur Verdeutlichung der beschriebenen Sachverhalte. Sie sind als Flash-Filme realisiert und lassen sich direkt aus dem Text heraus mit einem einfachen Mausklick starten. Neben den Beispielen für die typischen Testsignale Sinuston, bandbegrenztetes Rauschen und weißes Rauschen enthält die Anwendung das Sample einer Straßenszene mit Stimmen, vorbeifahrenden Autos und Bahnen sowie Kleingeldgeklimmer, das das Prinzip der auditiven Objektwahrnehmung verdeutlichen soll. Dieses Sample wurde mit besonderen Stereomikrofonen¹⁶⁵ und einem MD-Recorder aufgenommen, um sie Szene möglichst originalgetreu wiedergeben zu können.



Abbildung 40: Einbettung des Hörbeispiels in die Seitendarstellung
Das Sample wird durch Anklicken unmittelbar gestartet.

¹⁶⁵ Soundman OKM. Diese Mikrofone werden im Ohr getragen, so dass die Wiedergabe per Kopfhörer einen sehr originalgetreuen Raumeindruck erzeugt.

12.4.4 Hörversuch: Rauigkeit

Dieser Versuch soll die Abhängigkeit des Klangeindrucks eines Zusammenklangs von zwei Sinustönen von deren Frequenzabstand demonstrieren (vgl. Abschnitt 6.6). Dazu werden ein Vergleichston mit einer festen Frequenz von 440 Hz und ein zweiter Sinuston mit verschiedenen Frequenzen erst nacheinander und dann gleichzeitig abgespielt. Es gibt die Möglichkeit, entweder die Frequenz des zweiten Tons zu wählen und dadurch einzelne Vergleiche für verschiedene Frequenzabstände manuell durchzuführen oder einen Demonstrationsmodus zu starten, in dem der Frequenzabstand automatisch Stufe für Stufe variiert wird. Die Frequenz der Töne wird auf der Anzeigeskala im unteren Bereich markiert. Dort sind – farblich gekennzeichnet – die Bereiche der verschiedenen Hörempfindungen erkennbar. So lässt sich die Änderung des Höreindrucks mit der Änderung des Frequenzabstands gut verfolgen.

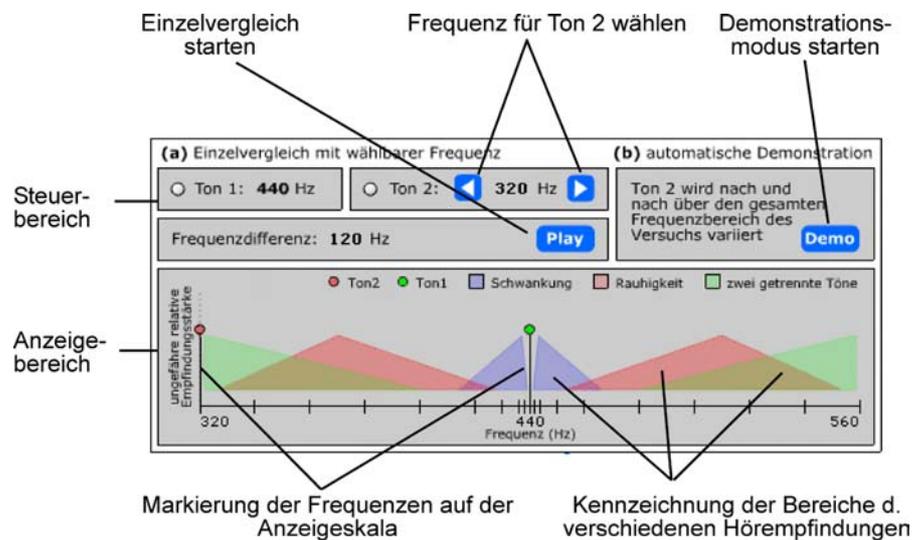


Abbildung 41: Hörversuch Rauigkeit

Der Versuch enthält 19 Sinustöne zwischen 320 und 560 Hertz. Der Vergleichston liegt mit 440 Hertz genau in der Mitte. Der maximale Frequenzabstand zwischen den beiden Tönen beträgt 120 Hertz, was ausreicht, um einen Zusammenklang zwei deutlich getrennter Töne wahrzunehmen. Der Abstand zwischen den Samples beträgt 20 Hertz. Im Bereich um 440 Hertz wurden geringere Frequenzabstände gewählt, da sich hier die Hörwahrnehmung mit der Frequenz rascher ändert.

12.4.5 Hörversuch: Frequenzabhängigkeit der Mithörschwelle

Dieser Versuch soll die Abhängigkeit der Mithörschwelle von der Frequenz verschiedener Sinustöne und somit die mit der Mittenfrequenz wachsende Bandbreite der auditiven Filter demonstrieren (5.3.2).

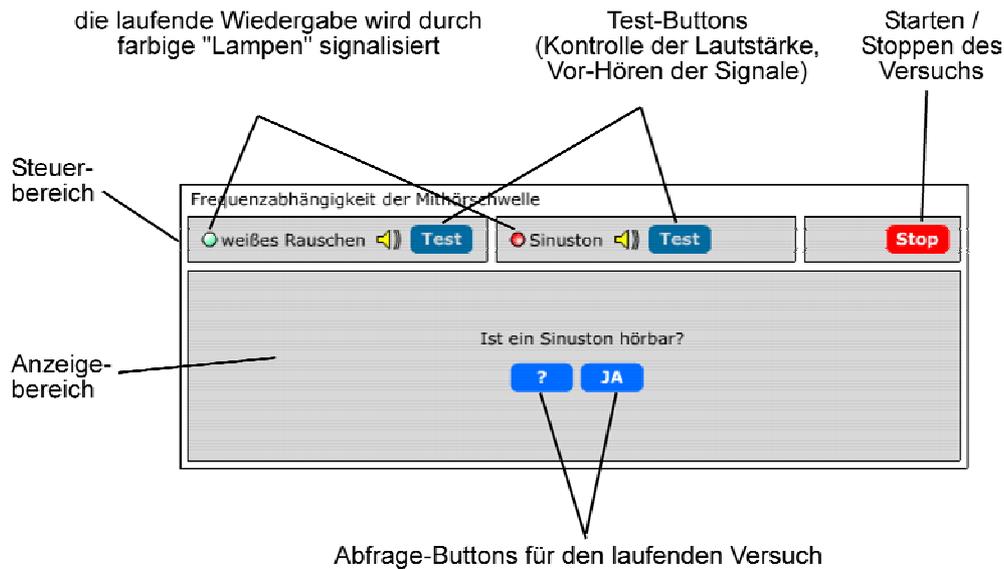


Abbildung 42: Hörversuch zur Frequenzabhängigkeit der Mithörschwelle

Darstellung der Oberfläche während des laufenden Versuchs. Die Test-Buttons sind Inaktiv (dunklerer Blau-Ton), die Abfrage-Buttons sind aktiv (heller Blau-Ton wie alle Navigationselemente der Anwendung).

Der Versuch enthält 13 Samples von Sinustönen verschiedener Frequenz (von 400 Hz bis 2800 Hz in 200 Hz-Schritten, MP3-codiert). Dieser Bereich wurde gewählt, weil dort die Hörschwelle weitgehend linear verläuft. Das Rauschsignal ist ein 2,5 s-Sample eines weißen Rauschens (44 kHz Sample-Rate, 16 Bit Auflösung). Die Bedienungsoberfläche bietet die Möglichkeit, die enthaltenen Samples mit den Test-Buttons abzuspielen.

Der Versuch läuft folgendermaßen ab: Zu Beginn wird allen Sinustönen die gleiche Anfangslautstärke zugewiesen. Bei Betätigung des Start-Buttons wird einer der Sinustöne durch eine Zufallsfunktion ausgewählt; das Rauschen und – mit einer Verzögerung von etwa 0,5 Sekunden – der gewählte Sinuston werden zwei Mal gleichzeitig abgespielt, danach wird der Ablauf angehalten. Die Versuchspersonen sind aufgefordert, auf den Button „JA“ – zu klicken, falls sie in dem Rauschen einen deutlichen Sinuston mit eindeutiger Tonhöhe wahrgenommen haben. Falls sie sich unsicher sind oder den Ton nur undeutlich oder gar nicht wahrgenommen haben, soll der Button „?“ angeklickt werden. Erst nach dem Anklicken eines Buttons wird der nächste Sinuston zufällig ausgewählt und die Prozedur wiederholt.

Wurde der abgespielte Sinuston gehört, wird die ihm zugewiesene Lautstärke um einen festen Wert verringert, so dass er beim nächsten Mal leiser abgespielt wird. War sich die Versuchsperson nicht sicher und hat den „?“ – Button gewählt, wird der Ton beim nächsten Mal noch einmal mit der gleichen Lautstärke abgespielt. Wird dieser Button beim gleichen Ton zum zweiten Mal angeklickt, wird er als nicht mehr deut-

lich wahrnehmbar betrachtet und programmintern gekennzeichnet. Die Lautstärke wird zwischengespeichert, und der Ton wird nicht mehr durch die Zufallsfunktion ausgewählt.

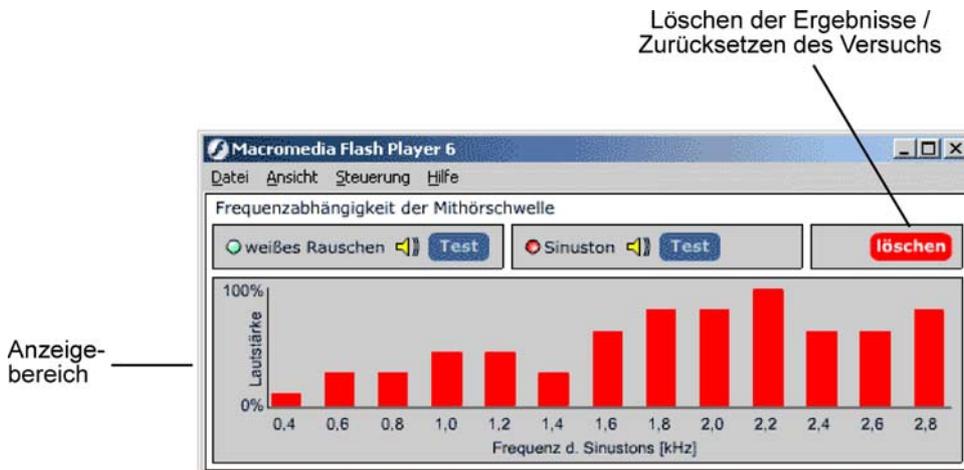


Abbildung 43: Auswertung des Versuchs Frequenzabhängigkeit der Mithörschwelle
 Die Auswertung zeigt die Lautstärken, bei denen die einzelnen Sinustöne nicht mehr wahrgenommen wurden. Im abgebildeten Beispiel wird deutlich, dass die Mithörschwelle bei höheren Frequenzen bei deutlich höheren Lautstärken lag.

Sobald alle Töne auf diese Art als nicht mehr wahrnehmbar gekennzeichnet sind, wird die Auswertung angezeigt. Die Lautstärken, bei denen die einzelnen Töne nicht mehr wahrnehmbar waren, werden als Balken angezeigt (Abbildung 43). Die Anzeige ist normiert: Der höchsten festgehaltenen Lautstärke wird die maximal darstellbare Balkenlänge zugewiesen. Die Länge der anderen Balken geben das Verhältnis dazu wieder.

12.4.6 Hörversuch: Effektive Maskierungsbandbreite

Dieser Versuch soll demonstrieren, dass zur Verdeckung eines Testtons nur die Energie eines Störschalls beiträgt, die in die Bandbreite des auditiven Filters um die Frequenz des Testtons fällt (vgl. Abschnitt 6.3). Der Versuch enthält das Sample eines Sinus-Testtons (Frequenz 2 kHz, Dauer 2 s, 50 ms Fade In/Out, Mono-WAV-File mit 22 kHz, 16 Bit) und 10 Samples von bandbegrenzten Rauschen mit verschiedener Bandbreite und gleichem RMS-Pegel (Dauer 2 s, Stereo-WAV-Files mit 22 kHz, 16 Bit). Die Mittenfrequenz der Rauschsignale liegt bei 2 kHz, die Bandbreiten sind 300 Hz bis 3000 Hz in 300 Hz-Schritten. Zu Beginn wird ein Array von gleichen Lautstärkewerten für den Sinuston erzeugt. Jede Bandbreite des Rauschens ist mit einem eigenen Lautstärkewert für den Sinuston verbunden.

Der Versuch läuft folgendermaßen ab: Bei Betätigung des Start-Buttons wird eines der Rauschsignale durch eine Zufallsfunktion ausgewählt; das Rauschen und der Sinuston werden zwei Mal gleichzeitig abgespielt, danach wird der Ablauf an-

gehalten. Die Versuchspersonen sind aufgefordert, auf den Button „JA“ – zu klicken, falls sie in dem Rauschen deutlich den Sinuston wahrgenommen haben. Falls sie sich unsicher sind oder den Ton nur undeutlich oder gar nicht wahrgenommen haben, soll der Button „?“ angeklickt werden. Erst nach dem Anklicken eines Buttons wird der nächste Sinuston zufällig ausgewählt und die Prozedur wiederholt.

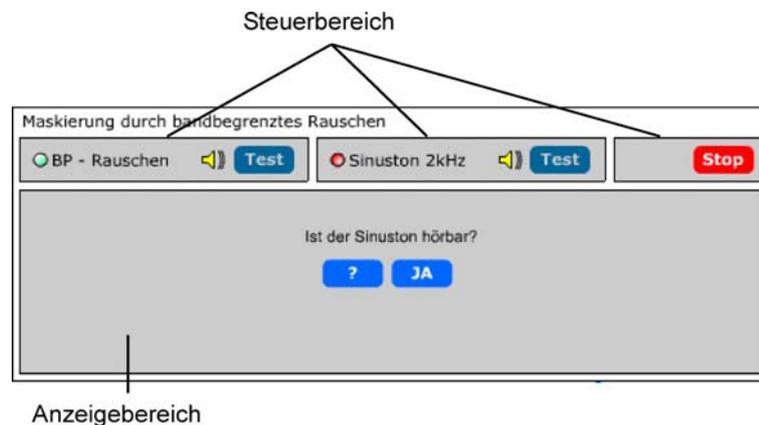


Abbildung 44: Hörversuch effektive Maskierungsbandbreite

Darstellung der Oberfläche während des laufenden Versuchs. Die Steuerung ist analog zum Versuch Frequenzabhängigkeit der Mithörschwelle (vgl. Abbildung 42)

Wurde der abgespielte Sinuston gehört, wird die ihm in Verbindung mit diesem Rauschsignal zugewiesene Lautstärke um einen festen Wert verringert, so dass er beim nächsten Mal leiser abgespielt wird. War sich die Versuchsperson nicht sicher und hat den „?“ – Button gewählt, wird der Ton beim nächsten Mal noch einmal mit der gleichen Lautstärke abgespielt. Wird dieser Button beim gleichen bandbegrenzten Rauschen zum zweiten Mal angeklickt, wird der Testton als nicht mehr deutlich wahrnehmbar betrachtet und das Rauschen programmintern gekennzeichnet. Die Lautstärke wird zwischengespeichert, und dieses Rauschsample wird nicht mehr durch die Zufallsfunktion ausgewählt.

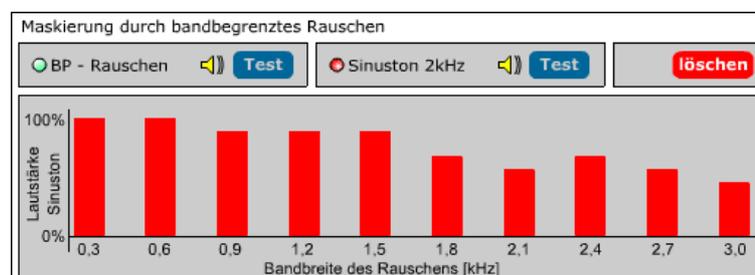


Abbildung 45: Auswertung des Hörversuchs zur effektiven Maskierungsbandbreite

Im abgebildeten Beispiel sinkt wie zu erwarten die Mithörschwelle für den Testton mit der Bandbreite des Rauschens.

Sobald alle Töne auf diese Art als nicht mehr wahrnehmbar gekennzeichnet sind, wird die Auswertung angezeigt. Die Lautstärken, bei denen die einzelnen Töne nicht mehr wahrnehmbar waren, werden als Balken angezeigt. Die Anzeige ist wie im Versuch zur Frequenzabhängigkeit der Mithörschwelle normiert (vgl. Abschnitt 12.4.5).

12.5 Einbindung ins WWW

Aufgrund der gewählten Software-Plattform kann die Anwendung sowohl lokal auf einem Rechner als auch auf einem Internet-Server gespeichert und über das Internet abgerufen werden. Auf jeden Fall werden die Inhalte in einem WWW-Browser wiedergegeben. Dies ermöglicht die einfache Anbindung an weitere im WWW erhältliche Ressourcen wie die in Abschnitt 11.2 aufgelisteten Anwendungen. Auf einige davon wird im vorliegenden Projekt verwiesen.

WWW-Link: [ars auditius](#)
Einführendes Online-Tutorial zu Akustik, Gehör und Psychoakustik von Martina Kremer, Bergische Universität Gesamthochschule Wuppertal

Abbildung 46: Link zu externer WWW-Ressource

Da zum Abruf eines externen Links eine bestehende Internet-Verbindung notwendig ist, die Inhalte externer Links nicht kontrollierbar sind und nicht zuletzt die Leistungen dritter nicht für das eigene Projekt vereinnahmt werden sollen, sind diese Links deutlich als externes Angebot gekennzeichnet (Abbildung 46).

13 Abschlussbetrachtung Teil B

Nach Abschluß des Projekts in der vorliegenden Form kann zusammenfassend festgestellt werden, dass mit den gewählten Mitteln eine Anwendung entwickelt werden konnte, die den aus der Analyse der Ausgangssituation entwickelten Anforderungen gut gerecht wird. In zwei wichtigen Bereichen mussten Kompromisse eingegangen werden, die Wünsche für Verbesserungen offenlassen: Dies betrifft die Hörversuche sowie an einigen Stellen die Verständlichkeit der Texte. Tests der Hörversuche – dies betrifft die Versuche zur Frequenzabhängigkeit der Mithörschwelle und zur effektiven Maskierungsbandbreite – mit mehreren Versuchspersonen ergaben z.T. Unsicherheiten, was die Anweisungen zur Durchführung angeht. Diese wurden daraufhin mehrfach verändert, ohne allerdings alle Unklarheiten ausräumen zu können. Dafür wären noch weitere Tests nötig. Zum Teil entstand auch während des Ablaufs der Versuche Unsicherheit durch das generelle Leiserwerden der Testtöne. Abhilfe

könnte eine Anlehnung des Ablaufs an die Methode des einpendelnden Regelns bieten. Dies konnte nicht mehr realisiert werden. Während der Aufbereitung der Inhalte für die Lehranwendung ergab sich darüber hinaus, dass die Anpassung des Textes von Teil A dieser Arbeit größeren Aufwand erforderte, als angenommen. Texte, die auf Papier gut verständlich sind, wirken auch in vereinfachter und gekürzter Form auf dem Bildschirm z.T. noch zu umfangreich und kompliziert. Hier musste daher ein Kompromiss gefunden werden zwischen optimaler Gestaltung und der angestrebten Abdeckung des gesamten Inhalts.

Dagegen erwiesen sich die gewählte Struktur und die Navigation den Inhalten als vollkommen angemessen. Die Anforderungen größtmöglicher Übersichtlichkeit und Benutzerfreundlichkeit können als voll erfüllt betrachtet werden. Darüber hinaus sind Struktur und Navigationssystem unabhängig vom speziellen Inhalt. Daher ist es denkbar, sie auch für die Präsentation anderer Inhalte, etwa eines Vorlesungsskripts, anzuwenden. Die Realisierung des Projekts auf HTML-Basis und die Einbettung von Hörbeispielen und Hörversuchen als Flash-Filme in den Text erfüllte ebenfalls den angestrebten Zweck. Die Anwendung ist weitestgehend plattformunabhängig einsetzbar und erfordert keine Downloads und zusätzliche Software.

Die Möglichkeit einer Veröffentlichung der Anwendung im WWW lässt noch Wünsche und bisher nicht beantwortete Fragen offen. Eine konsequente Ausrichtung auf eine WWW-Einbindung würde interessante zusätzliche Möglichkeiten bieten, so z.B. die Integration eines Nutzerforums. Ein bedeutendes Hindernis für einen öffentlichen Zugang zu den Inhalten in ihrer jetzigen Form stellt jedoch die Verwendung von Abbildungen aus urheberrechtlich geschützten Werken Dritter dar. Bedingung wäre demnach eine umfassende Klärung aller rechtlichen Fragen im Zusammenhang mit einer WWW-Publikation, die den Umfang dieser Arbeit sprengen würden.

14 Fazit

Die Ausarbeitung von Teil A ergab eine unerwartete Vielschichtigkeit des Themas und offenbarte vor allem eine erstaunliche Unübersichtlichkeit der Literatur (vgl. Abschlussbetrachtung Teil A, Abschnitt 9). Daraus wurden angesichts des Ziels, zu diesem Thema eine Multimedia-Lehranwendung zu entwickeln, Folgerungen gezogen, die zu der vorliegenden Form der Ausarbeitung führten. Diese Folgerungen werden durch die Ergebnisse nicht in Frage gestellt. Sie bedeuteten jedoch die Notwendigkeit, das Thema in großer inhaltlicher Breite zu bearbeiten. Damit war ein inhaltlicher Umfang abzudecken, der bei der Umsetzung in die Multimedia-Anwendung im Rahmen dieser Magisterarbeit kaum noch zu bewältigen. Die entstandene Anwendung wird trotz der in der Abschlussbetrachtung zu Teil B (Abschnitt 13) genannten

Kritikpunkte und offenen Wünsche dennoch den formulierten Anforderungen insgesamt gerecht und wird als für den angestrebten Zweck einsetzbar betrachtet. Theoretische Aspekte des *Lehrens und Lernens* und die Berücksichtigung der *Didaktik* bei der Erstellung der einzelnen Inhalte konnten aus Gründen des Umfangs jedoch nicht mehr in einem Maße behandelt werden, wie es unter anderen Voraussetzungen bei der Entwicklung einer Multimedia-Lehranwendung notwendig gewesen wäre.

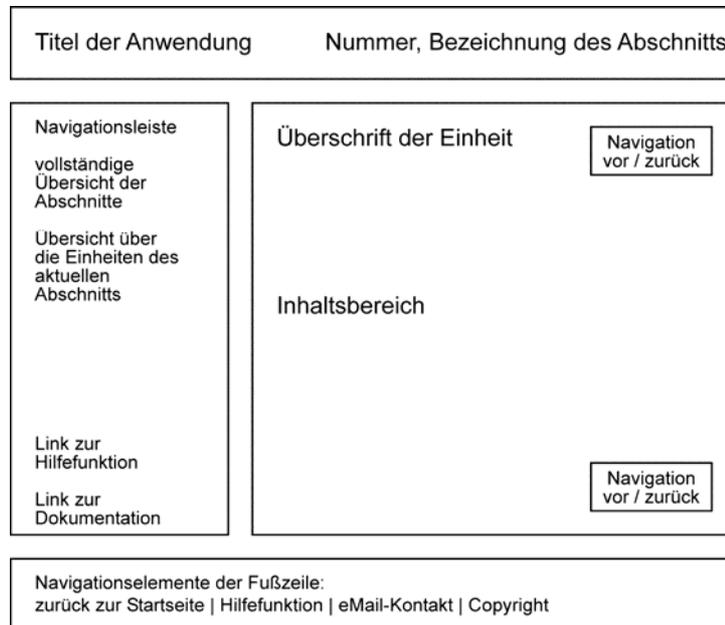
15 Literatur

- Allen, Jont B. (1995): Harvey Fletcher 1884-1981. In: Fletcher, Harvey (1995): Speech and hearing in communication. Woodbury: Acoustical Society of America. S. A1-A34 (ASA-Edition; Reprint der Originalausgabe von 1953).
- Burrows, Dennis L./S. Joseph Barry (1990): Electrophysical evidence for the critical band in humans: Middle-latency responses. In: Journal of the Acoustical Society of America 88(1), S. 180-184.
- Evans, Edward Frank (1992): Auditory processing of complex sounds: an overview. In: Phil. Trans. R. Soc. Lond. B 336, S. 295-306.
- Fletcher, Harvey/Wilden A. Munson (1933a): Loudness, its definition, measurement, and calculation. In: Journal of the Acoustical Society of America 5, S. 82-108.
- Fletcher, Harvey/Wilden A. Munson (1933b): Loudness, its definition, measurement, and calculation. Bell System Technical Journal xii (Oktober), S. 377-430.
- Fletcher, Harvey/Wilden A. Munson (1937): Relation between loudness and masking. In: Journal of the Acoustical Society of America 9, S.1-10.
- Fletcher, Harvey (1938a): Loudness, masking and their relation to the hearing process and the problem of noise measurement. In: Journal of the Acoustical Society of America 9, April, S. 275-293.
- Fletcher, Harvey (1938b): The mechanism of hearing as revealed through experiments on the masking effect of thermal noise. In: Proceedings of the National Academy of Science 24, S. 265-274.
- Fletcher, Harvey (1940): Auditory Patterns. In: Reviews of Modern Physics 12, S. 47-65.
- Fletcher, Harvey (1995): Speech and hearing in communication. Woodbury: Acoustical Society of America (ASA-Edition; Reprint der Originalausgabe von 1953).
- Goldstein, E. Bruce (1989): Sensation and Perception (3rd Edition). Pacific Grove (CA): Brooks/Cole.
- Goldstein, E. Bruce (1997): Wahrnehmungspsychologie. Heidelberg; Berlin; Oxford: Spektrum Akad. Verl.
- Greenwood, Donald D. (1990): A cochlear frequency-position function for several species – 29 years later. In: Journal of the Acoustical Society of America 87(6), S. 2592-2605.
- Hall, Joseph W./John H. Grose/Lee Mendoza (1995): Across-Channel Processes in Masking. In: Moore, Brian C.J. (Hrsg.): Hearing. San Diego: Academic Press.
- Hartmann, William M. (1997): Signals, Sound, and Perception. Woodbury (New York): American Institute of Physics.
- Hellbrück, Jürgen (1993): Hören. Physiologie, Psychologie und Pathologie. Göttingen: Hogrefe
- Holzinger, Andreas (2001): Basiswissen Multimedia. Band 3: Design. Würzburg: Vogel.
- Issing, Ludwig J./Paul Klimsa (1997): Informationen und Lernen mit Multimedia. 2. Auflage. Weinheim, Psychologie Verlags Union.
- Luce, R. Duncan (1993): Sound & Hearing. A conceptual introduction. Hillsdale (New Jersey): Lawrence Erlbaum Associates.
- Møller, Aage R. (2000): Hearing. Its Physiology and Pathophysiology. San Diego: Academic Press.
- Moore, Brian C.J./ Michael J. Shailer (1992): Modulation discrimination interference and auditory grouping. In: Phil. Trans. R. Soc. Lond. B 336, S. 339-346.
- Moore, Brian C.J. (1995): Hearing. San Diego: Academic Press
- Moore, Brian C.J. (1997): An introduction to the psychology of hearing. London: Academic Press.
- Palmer, Alan R. (1995): Neural Signal Processing. In: Moore, Brian C.J. (Hrsg.): Hearing. San Diego: Academic Press. S. 75-121.
- Pickles, James O. (1988): An introduction to the physiology of hearing – 2nd ed. London: Academic Press.
- Plomp, Reinier (2002): The Intelligent Ear. On the Nature of Sound Perception. Mahwah (New Jersey); London: Lawrence Erlbaum Associates.

- Scharf, Bertram (1970): Critical Bands. In: Tobias, Jerry V.(Hrsg.): Foundations of Modern Auditory Theory, Vol. 1. New York; London: Academic Press.
- Schmidt, Robert F. (1993) (Hrsg.): Neuro- und Sinnesphysiologie. Berlin: Springer.
- Schmidt, Robert F./Hans-Georg Schaible (2000): Neuro- und Sinnesphysiologie. Berlin, Heidelberg, New York: Springer.
- Schwier, Richard A./Earl R. Misanchuk (1993): Interactive Multimedia Instruction. Englewood Cliffs: Educational Technology Publications.
- Steinmetz, Ralf (1998): Multimedia-Technologie. Grundlagen, Komponenten und Systeme. Berlin: Springer.
- Stevens, Kenneth N. (1998): Acoustic Phonetics. Cambridge (MA); London: MIT Press.
- Terhardt, Ernst (1998): Akustische Kommunikation. Berlin: Springer.
- Thissen, Frank (2001): Screen-Design Handbuch. Effektiv informieren und kommunizieren mit Multimedia. Berlin: Springer.
- Warren, Richard M. (1999): Auditory perception: A new analysis and synthesis. Cambridge: Cambridge University Press.
- Zenner, H.P. (2000): Hören. In: Schmidt, Robert F./Hans-Georg Schaible (Hrsg.): Neuro- und Sinnesphysiologie. Berlin, Heidelberg, New York: Springer, S. 317-342
- Zwicker, Eberhard/G. Flottorp/S.S. Stevens (1957): Critical bandwidth in loudness summation. In: Journal of the Acoustical Society of America 29, S. 548-557.
- Zwicker, Eberhard (1961): Subdivision of the audible frequency range into critical bands (Frequenzgruppen). In: Journal of the Acoustical Society of America 33, S. 248.
- Zwicker, Eberhard/Fastl, Hugo (1999): Psychoacoustics. Facts and Models. Berlin: Springer.

16 Anhang

16.1 Designvorgaben: Bildschirmmaske



Bildschirmmaske für alle inhaltlichen Einheiten

Die Bildschirmmaske ist die verbindliche Vorgabe - mit festen Maßen - für die Anordnung der Inhalts-, Navigations- und Gestaltungselemente in allen inhaltlichen Einheiten. Realisiert wurde die Anordnung mittels eines HTML – Tabellengerüsts, das als Vorlage in Macromedia Dreamweaver benutzt wurde.

16.2 Designvorgaben: Text / Typographie

Die benutzten Schriftarten Arial und Verdana sind serifenlos, um eine gute Lesbarkeit am Bildschirm zu gewährleisten. Die Textausrichtung ist grundsätzlich linksbündig.

Bezeichnung für Textelemente und deren - Ableitungen	Schriftart	Schriftgröße	Farbe (RGB-Wert)	Schriftschnitt / weitere Attribute
Textkörper - Standarddarstellung	Verdana	12 Pixel	Schwarz #000000	normal
- hervorgehobene Begriffe	Verdana	12 Pixel	Schwarz #000000	fett
Überschrift - für Kernkomponente	Verdana	14 Pixel	Schwarz #000000	fett

- im Text	Verdana	12 Pixel	Schwarz #000000	fett
Navigation: - Standarddarstellung	Arial	12 Pixel	Blau #0066FF	normal / nicht unterstrichen
- aktiver Link	Arial	12 Pixel	Rot #FF0000	fett / nicht unterstrichen
- kleine Darstellung	Arial	10 Pixel	Blau #0066FF	normal / nicht unterstrichen
Bildunterschrift	Arial	10 Pixel	Schwarz #000000	normal
- Numerierung	Arial	10 Pixel	Schwarz #000000	normal

16.3 Designvorgaben: Navigationselemente

Alle anwählbaren Navigationselemente – Schaltflächen und Links – sind im gleichen Farbton gehalten (Blau mit RGB-Wert #0066FF). Der Link zur aktuell aufgerufenen Seite ist in der Navigation mit Rot (RGB-Wert #FF0000) gekennzeichnet. Navigationselemente sind immer an der gleichen Stelle angeordnet. Gleiche Elemente erfüllen immer gleiche Funktionen (Prinzip der Konsistenz). Symbole sind eindeutig gestaltet und am üblichen Gebrauch orientiert (🔍 für die Hilfefunktion, ✖ für Fenster schließen, ⏪ Pfeile für vor/zurück und Verweise, 🔊 für Audio).

16.4 Designvorgaben: Abbildungen

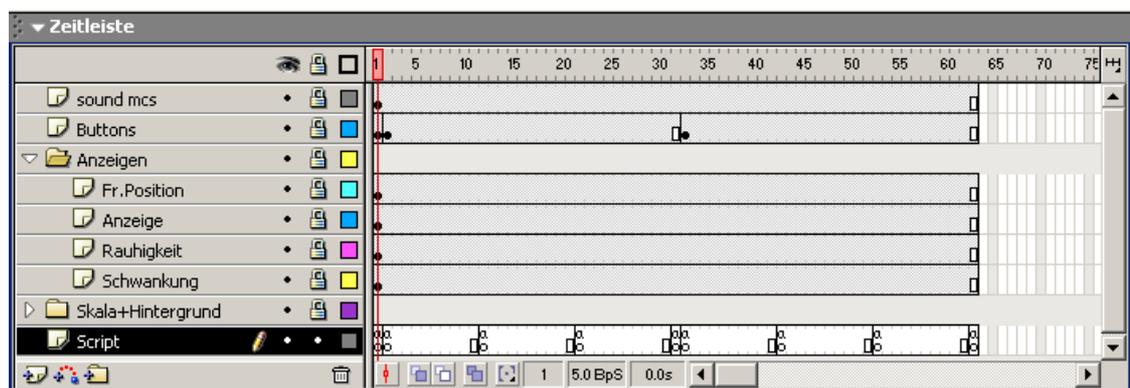
Abbildungen sind auf Übersichtlichkeit ausgerichtet. Diagramme werden in Strichzeichnungen mit starken Kontrasten gehalten. Abbildungen sollen jeweils nur eine Kernaussage beinhalten. Die Sprache der Beschriftungen ist Deutsch. Schriftart ist auch in Grafiken die serifenlose Type Arial. Die Dateigrößen sind so klein wie möglich gehalten. Die ALT-Tags im HTML-Quellcode enthalten sinnvolle Beschreibungen der Grafiken, um das Auslesen durch Suchmaschinen zu erleichtern. Die Abbildungen sind in den Abschnitten fortlaufend nummeriert. In den Literaturangaben zum Abschnitt sind Verweise auf die Quellen angegeben.

16.5 Designvorgaben: Text / Inhalt und Stil

Komplizierte Formulierungen und verschachtelte Sätze sollen vermieden werden. Inhaltliche Kernaussagen sollen als Absätze gekennzeichnet werden. Wo immer es möglich ist, werden deutsche Übersetzungen von Fachbegriffen vorgezogen.

16.6 Programmstruktur der Hörversuche

Macromedia Flash ist im Grunde genommen ein zeitleistenbasiertes Animationsprogramm. Erst im Laufe verschiedener Versionen wurde es erweitert und durch die Scriptsprache ActionScript zu einer flexiblen, leistungsfähigen Entwicklungssoftware zur Realisierung verschiedenster interaktiver Internet-Anwendungen. Dieser Ursprung ist dafür verantwortlich, dass die Struktur von Flash-Anwendungen schwer dokumentierbar ist.



Die Zeitleiste der Entwicklungs-Datei *rauhigkeit fla*. ActionScript-Elemente können einzelnen Bildern der Zeitleiste und verschiedenen Arten von Objekten zugewiesen werden, die wiederum eine Zeitleiste besitzen können.

Elemente der Scriptsprache können verschiedensten Bildern, Ebenen und Objekten zugewiesen werden. Daher kann an dieser Stelle kein sinnvoller „Programmcode“ dokumentiert werden. Stattdessen wird die logische Struktur der Programmfunktionen skizziert. Die Original-Entwicklungsdateien, in denen auch die einzelnen Elemente des Programmcodes eingesehen werden können, sind auf der dieser Arbeit beigefügten CD-ROM enthalten.

16.6.1 Rauhigkeit

Frequenzstufe: Maximalwert vorgegeben durch die Anzahl der Samples

Nr.	Name der Funktion	Anweisungen	Verzweigung
1	Ausgangszustand herstellen		2, 3
1.1		Frequenzstufe Ton 2 = niedrigste	
2	Einzelvergleichsmodus		2.1, 2.2, 4
2.1	Frequenzauswahl		
2.1.1		Falls Button „Frequenzstufe niedriger“ gedrückt: Überprüfe, ob niedrigere Stufe vorhanden. Falls ja: Frequenzstufe Ton	2.1.2, 2.2, 4

		2 = nächstniedrigere	
2.1.2		Falls Button „Frequenzstufe höher“ gedrückt: Überprüfe, ob höhere Stufe vorhanden. Falls ja: Frequenzstufe Ton 2 = nächsthöhere	2.1.1, 2.2, 4
2.2	gewählten Vergleich abspielen		2.1, 4
		Falls Button „Play“ gedrückt, gehe zu 3	
3	Töne wiedergeben (Frequenzstufe)		Stop-Button
	3.1	Spieler Ton 1	
	3.2	Spieler Ton 2 (Frequenzstufe)	
	3.3	Spieler Ton 1 + 2 (Frequenzstufe) gleichzeitig	
4	Demo-Modus starten		
	4.1	Frequenzstufe Ton 2 = niedrigste	
	4.2	gehe zu 3	
	4.3	Überprüfe, ob höhere Stufe vorhanden. Falls ja: Frequenzstufe Ton 2 = nächsthöhere, gehe zu 4.2 Falls nein: gehe zu 1	

16.6.2 Effektive Maskierungsbandbreite

Vorgegeben sind: Zahl der Bandbreitestufen (Zahl der Samples verschiedener Bandbegrenzter Rauschen), Anfangslautstärke, Lautstärkestufe

Nr.	Name der Funktion	Anweisungen	Verzweigung
0	Ausgangszustand herstellen		1, 3, 5
	0.1	Für alle Bandbreitestufen: Hörschwelle = Anfangslautstärke	
	0.2	Für alle Bandbreitestufen: markiere als auswählbar	
	0.3	Für alle Bandbreitestufen: unsicher (Bandbreitestufe) = nein	
1	Testfunktion Sinuston		3,5
	1.1	Lautstärke = Anfangslautstärke	
	1.2	gehe zu 2	
2	Sinuston abspielen (Lautstärke)		
	2.1	Sinuston (Lautstärke) wiedergeben	
3	Testfunktion bandbegrenztes Rauschen		
	3.1	Bandbreitestufe = niedrigste	

3.2		gehe zu 4 (Bandbreitestufe)	
3.4		Falls keine höhere Bandbreitestufe vorhanden: gehe zu 0 Sonst: Bandbreitestufe = nächsthöhere, gehe zu 3.2	0, 3.2
4	bandbegrenzt tes Rauschen abspielen (Band- breitestufe)		
4.1		bandbegrenzt tes Rauschen (Band- breitestufe) abspielen	
5	Hörversuch		1,3
5.1		gehe zu 6	
5.2		Bandbreitestufe = Zufallszahl	
5.3		gehe zu 4 (Bandbreitestufe)	
5.4		Lautstärke = Hörschwelle (Bandbrei- testufe)	
5.5		gehe zu 2 (Lautstärke)	
5.6		wiederhole 5.2-5.5	
5.7		Programm anhalten	
5.8		Falls Button „JA“ gedrückt: Hör- schwelle (Bandbreitestufe) = Hör- schwelle (Bandbreitestufe)- Lautstärkestufe, gehe zu 5.1	5.1
5.9		Falls Button „?“ gedrückt: [Falls unsicher (Bandbreitestufe) = nein: markiere als unsicher, gehe zu 5.1 Falls unsicher (Bandbreitestufe) = ja: [speichere Hörschwelle (Bandbreite- stufe), getestete Stufen = getestete Stufen + 1, Falls getestete Stufen = Anzahl der Bandbreitestufen: gehe zu 7 Sonst: markiere Bandbreitestufe als nicht mehr auswählbar, gehe zu 5.1]]	5.1, 7
6	Zufallszahl auswählen		
6.1		Zufallszahl = wähle zufällig aus den auswählbaren Bandbreitestufen	
7	Auswertung (Hörschwel- len)		
7.1		ermittle Max.Schwellenwert = höch- ste gespeicherte Hörschwelle	
7.2		Für jede Bandbreitestufe: zeichne vertikalen Balken mit Länge Hörschwelle (Bandbreitestufe) * [Max.Schwellenwert / Hörschwelle (Bandbreitestufe)]	
7.3		Falls Button „löschen“ gedrückt: lösche Anzeige, gehe zu 0	0

16.6.3 Frequenzabhängigkeit der Mithörschwelle

Vorgegeben sind: Zahl der Frequenzstufen (Zahl der Samples verschiedener Sinustöne), Anfangslautstärke, Lautstärkestufe

Nr.	Name der Funktion	Anweisungen	Verzweigung
0	Ausgangszustand herstellen		1, 3, 5
	0.1	Für alle Frequenzstufen: Hörschwelle = Anfangslautstärke	
	0.2	Für alle Frequenzstufen: markiere als auswählbar	
	0.3	Für alle Frequenzstufen: unsicher (Frequenzstufe) = nein	
1	Testfunktion Rauschen		3,5
	1.2	gehe zu 2	
2	Rauschen abspielen		
	2.1	weißes Rauschen wiedergeben	
3	Testfunktion Sinustöne		
	3.1	Frequenzstufe = niedrigste	
	3.2	gehe zu 4 (Frequenzstufe, Anfangslautstärke)	
	3.4	Falls keine höhere Frequenzstufe vorhanden: gehe zu 0 Sonst: Frequenzstufe = nächsthöhere, gehe zu 3.2	0, 3.2
4	Sinuston abspielen (Frequenzstufe, Lautstärke)		
	4.1	Sinuston (Frequenzstufe, Lautstärke) abspielen	
5	Hörversuch		1,3
	5.1	gehe zu 6	
	5.2	Frequenzstufe = Zufallszahl	
	5.3	gehe zu 2	
	5.4	Lautstärke = Hörschwelle (Frequenzstufe)	
	5.5	gehe zu 4 (Frequenzstufe, Lautstärke)	
	5.6	wiederhole 5.2-5.5	
	5.7	Programm anhalten	
	5.8	Falls Button „JA“ gedrückt: Hörschwelle (Frequenzstufe) = Hörschwelle (Frequenzstufe)-Lautstärkestufe, gehe zu 5.1	5.1
	5.9	Falls Button „?“ gedrückt: [Falls unsicher (Frequenzstufe) = nein: markiere als unsicher, gehe zu 5.1 Falls unsicher (Frequenzstufe) = ja: [speichere Hörschwelle (Frequenz-	5.1, 7

stufe), getestete Stufen = getestete Stufen + 1,
 Falls getestete Stufen = Anzahl der Frequenzstufen: gehe zu 7
 Sonst: markiere Frequenzstufe als nicht mehr auswählbar, gehe zu 5.1]]

- | | | |
|---|--|--|
| 6 | Zufallszahl auswählen
6.1 | Zufallszahl = wähle zufällig aus den auswählbaren Frequenzstufen |
| 7 | Auswertung (Hörschwellen)
7.1

7.2

7.3 | ermittle Max.Schwellenwert = höchste gespeicherte Hörschwelle
Für jede Frequenzstufe:
zeichne vertikalen Balken mit Länge Hörschwelle (Frequenzstufe) * [Max.Schwellenwert / Hörschwelle (Frequenzstufe)]
Falls Button „löschen“ gedrückt: 0
lösche Anzeige, gehe zu 0 |

16.7 Inhalte, Verzeichnisse und Dateinamen der Anwendung

Inhaltsverzeichnis	startseite.htm		
Projektdokumentation	/dokumentation		
	Über dieses Projekt	inhalt.htm	
	Literatur A - M	literatur.htm	
	Literatur P-Z	literatur2.htm	
Hilfe	/hilfe		
	Hilfe & Lexikon	inhalt.htm	
	Lexikon	/lexikon	
	Bandpassrauschen		bandpassrauschen.htm
	Bark-Skala		bark-skala.htm
	Basilarmembran		basilarmembran.htm
	Cochlea (Hörschnecke)		cochlea.htm
	ERB (Equivalent Rectangular Bandwidth)		erb.htm
	Filter		filter.htm
	Harvey Fletcher		fletcher.htm
	Frequenzselektivität		frequenzselektivitaet.htm
	Haarzellen		haarzellen.htm
	Inhalt		inhalt.htm
	Lautheit (empfundene Lautstärke)		lautheit.htm
	Linearität		linearitaet.htm
	Maskierung		maskierung.htm
	Mithörschwelle		mithoerschwelle.htm
	Brian C.J. Moore		moore.htm
	MP3		mp3.htm
	natürliche Hörsituation		natuerliche_hoersituation.htm

	Nervenfasern (Neuronen)	nervenfasern.htm
	Psychoakustik	psychoakustik.htm
	Sinuston / Sinusschwingung	sinuston.htm
	Spektrum	spektrum.htm
	typische Teststimuli	teststimuli.htm
	weißes Rauschen	weissesrauschen.htm
	Hinweise zur Navigation und Bedienung	navigation.htm
	Wie finde ich was? Aufbau des Inhalts	wiefinden.htm
	Sitemap - vollständige Auflistung aller Inhalte	sitemap.htm
	Technische Anforderungen	technische_anforderungen.htm
	Hörversuche: Technische Anforderungen, Bedienung, Interpretation der Ergebnisse	hoerversuche.htm
Übersicht der Hörbeispiele	audio.htm	
1. Einführung	/einfuehrung	
	Zum Thema	inhalt.htm
	Gliederung des Inhalts	gliederung.htm
	Benötigte Vorkenntnisse	vorkenntnisse.htm
	Anmerkungen zur Psychoakustik	psychoakustik.htm
	Akustische Größen und Empfindungsgrößen	groessen.htm
2. Grundlagen der Gehörphysiologie	/physiologie	
	Inhalt	inhalt.htm
	Das auditive System: Begriffe	begriffe.htm
	Das Außenohr	aussenohr.htm
	Das Mittelohr	mittelohr.htm
	Das Innenohr: Die Hörschnecke	innenohr.htm
	Basilarmembran und Cortisches Organ	basilarmembran.htm
	Auslenkung der Basilarmembran	auslenkung.htm
	Wanderwelle der Basilarmembran	wanderwelle.htm
	Stimulationsmechanismus der Haarzellen	stimulation.htm
	Die Funktion der Haarzellen	haarzellen.htm
	Zusammenfassung	zusammenfassung.htm
	Literatur + Quellen	literatur.htm
3. Frequenzanalyse durch das Gehör	/frequenzanalyse	
	Inhalt	inhalt.htm
	Frequenzselektivität des Gehörs	frequenzselektivitaet_gehoer.htm
	Frequenzselektivität der Basilarmembran	frequenzselektivitaet_bm.htm
	Die Nicht-Linearität der Basilarmembranschwingung	nichtlinearitaet.htm
	Frequenzselektivität der Hörnervenfasern	ftc.htm
	Tonotope Organisation des	tonotopie.htm

	Hörnervensystems	
	Phasenkopplung neuronaler Entladungen	phasenkopplung.htm
	Frequenz und neuronaler Code	neuronalercode.htm
	Örtliche Codierung	ortscodierung.htm
	Zeitliche Codierung	zeitcodierung.htm
	Auditive Vorverarbeitung in der Cochlea	vorverarbeitung.htm
	Zusammenfassung	zusammenfassung.htm
	Literatur + Quellen	literatur.htm
4. Das Filtermodell des auditiven Systems	/filter	
	Inhalt	inhalt.htm
	Akustische Filter: Begriffe	filterbegriffe.htm
	Psychophysische Frequenz-tuningkurven	ptc.htm
	Die Form auditiver Filter / Die Notched-Noise Methode	filterform.htm
	Berechnung der Filterbandbreite / Äquivalente rechteckige Bandbreite (ERB)	bandbreite.htm
	Mittelfrequenz-Bandbreite-Funktion	mittenfrequenz.htm
	Nicht-Linearität der Filter / Abhängigkeit von Schallpegel	nichtlinearitaet.htm
	Effekte nicht-simultaner Maskierung	nichtsimultan.htm
	Schematische Darstellung des Filter-Modells	filtermodell.htm
	Zusammenfassung	zusammenfassung.htm
	Literatur + Quellen	literatur.htm
5. Kritische Bandbreiten und Maskierung	/kritische_bandbreiten	
	Inhalt	inhalt.htm
	Maskierungseffekte	maskierungseffekte.htm
	Kritische Bandbreiten	kritische_bandbreiten.htm
	Effektive Maskierungsbandbreite	effektivemaskierungsbandbreite.htm
	Hörversuch: Effektive Maskierungsbandbreite	hoerversuch_maskierungsbandbreite.htm
	Frequenzabhängigkeit der Mithörschwelle	mithoerschwelle.htm
	Hörversuch: Frequenzabhängigkeit der Mithörschwelle	hoerversuch_mithoerschwelle.htm
	Lautheit und kritische Bandbreite	lautheitkritische_bandbreite.htm
	Absolute Hörschwelle und kritische Bandbreite	hoerschwellekritische_bandbreite.htm
	Rauhigkeit und Schwankungsstärke	rauhigkeit.htm
	Hörversuch: Rauhigkeit und Schwankungsstärke	hoerversuch_rauhigkeit.htm
	Neuronale Erregungsmuster	neuronaleerregungsmuster.htm
	Maskierung und neuronale Erregung	maskierungneuronale_erregung.htm

	Lautheit und neuronale Erregung	lautheitneuronale erregung.htm
	Zusammenfassung	zusammenfassung.htm
	Literatur + Quellen	literatur.htm
6. Filterübergreifende Verarbeitung und Grenzen des Filtermodells	/grenzen	
	Inhalt	inhalt.htm
	Das Energiespektrum-Modell	energiespektrum modell.htm
	Profilanalyse	profilanalyse.htm
	Comodulation Masking Release	cmr.htm
	Modulation Discrimination Interference	mdi.htm
	Auditive Objektwahrnehmung (Auditory Grouping)	auditorygrouping.htm
	Zusammenfassung	zusammenfassung.htm
	Literatur + Quellen	literatur.htm
7. Anwendungen des Filtermodells	/anwendungen	
	Inhalt	inhalt.htm
	Mathematische Filtermodelle	filtermodelle.htm
	Hörgeräte und Cochlea-Implantate	implantate.htm
	Perzeptive digitale Audiocodierung	audiocodierung.htm
	Links + Downloads	links.htm
	Literatur + Quellen	literatur.htm

17 Anlage: CD-ROM

Die CD-ROM enthält die vollständige Multimedia-Anwendung. Sie lässt sich durch Aufruf der Datei `start.htm` im Ordner `Psychoakustik` starten. Der Ordner `flash` enthält die Entwicklungsdateien der Hörbeispiele und Hörversuche, die sich mit Macromedia Flash MX öffnen lassen. Die Datei `liesmich.txt` enthält u.a. Hinweise zu den technischen Anforderungen.