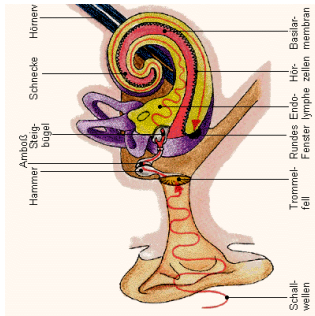


Gehörwahrnehmung: Agenda

- Schall und Schallausbreitung
- Aufbau des Gehörs
- Außenohr, Mittelohr, Innenohr
- Frequenzanalyse in der Cochlea
- Hörschwelle, Hörbereich, Lautheit
- Tonhöhe, Lautstärke, Klangfarbe
- Klangfarbe und Spektren
- Lärmempfindung, Schädigungen
- Sprache, Vokalartikulation



Motivation

wozu die Gehörwahrnehmung untersuchen?

- Forschung / Erkenntnisgewinn
- möglichst optimale Verfahren zur:
 - Aufzeichnung / Speicherung / Wiedergabe von Audio
 - so gut wie nötig - aber nicht besser (!)
- psychoakustische Verfahren, etwa MP3-Kompression
- bessere Algorithmen für Musik und Musikproduktion
- Raumsimulation für Spiele oder virtuelle Realität
- möglichst effiziente (low-level) Sprachverarbeitung
- medizinische Anwendungen, vor allem Hörgeräte

Schall

physikalischer Schallbegriff:

- alle elastischen Schwingungen eines Materials
- im allg. Fall longitudinal und transversal
- aber in Gasen (Luft) nur longitudinal (=Druckschwingungen)
- Schallgeschwindigkeit ca. 340 m/s in Luft
- Schallintensität lokale Druckänderung
 - Luftdruck auf Meereshöhe 1 bar = 10.000 Pa
 - Schmerzschwelle 1 mbar = 10 Pa
 - wahnehmbare Druckänderung 10E-9 bar = 0.00001 Pa
- Ohr hat optimale Empfindlichkeit
- weitere Steigerung: Wahrnehmen der Molekularbewegung

Schall: Leistung, dB

- Kenngrößen des Schallfelds: Druckamplitude P bzw. Intensität I
- Gesamtleistung: Integral der Intensität über eine Oberfläche
- Lautstärke (Intensität): dB
 - $L / dB := 10 \log_{10}(I / I_{ref})$
 - $P_{ref} := 20 \log_{10}(P / P_{ref})$
 - $P_{ref} := 0.00002 Pa$
- alternativ: dB als Maß für das Verhältnis (p_1/p_2)

20 dB	zehnfacher Schalldruck
6 dB	doppelter Schalldruck
-6 dB	halber Schalldruck

Quelle	Leistung (typ.)
Sprache	10 μW
Geige	1 mW
Lautsprecher	100 W

Schall: Phon

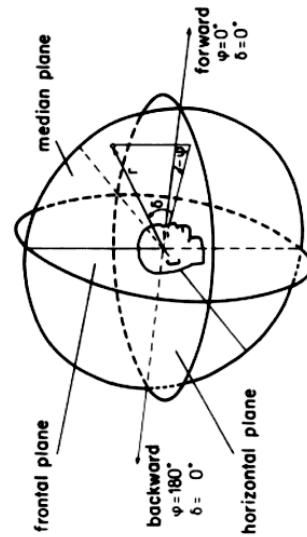
Phon: hörbezogene Einheit der Schallintensität
 Intensität eines gleichlaut empfundenen Referenztons
 logarithmische Skala notwendig (13 Größenordnungen)

$$I/\text{phon} := 10 * \log_{10}(I / I_0)$$

I_0 := Intensität der Hörschwelle bei 1000Hz
 := 10E-16 W/cm²

Quelle	Phon
Hörschwelle	0 (I = I ₀)
Flüstern	10 (I = 10 * I ₀)
Sprache	50 (I = 100000 * I ₀)
laute Radiomusik	80 (I = 100000000 * I ₀)
Disco,Preßlufthammer	130 (Schmerzschwelle)

Koordinatensystem



[Blauert]

Figure 1.4 A head-related system of coordinates used in auditory experiments; r is distance, φ is the azimuth, and δ is the elevation.

- sphärische Koordinaten sinnvoll
- bei Bedarf Umrechnung Welt/Kopf-Koordinaten

Schall: Nah- und Fernfeld

[Blauert]

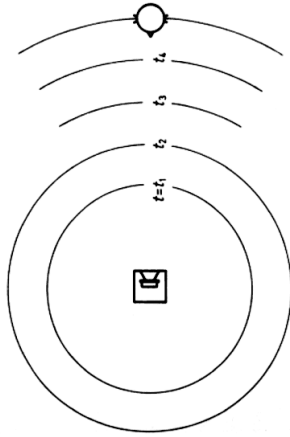


Figure 1.14 Any pulsating sound source that is small compared with the wavelength generates spherical waves at a moderate distance. As the distance is increased, the spherical waves become more and more plane.

- relative Größe / Entfernung von Quelle und Hörer

Schall: Nah- und Fernfeld

[Blauert]

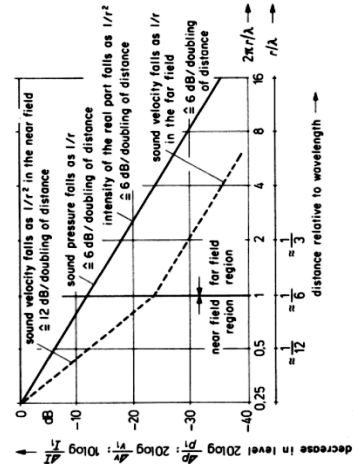


Figure 1.13 Pressure, velocity, and intensity of the sound field near to and distant from a spherical radiator of the zeroth order.

- typischer Abfall mit 6dB bzw. 12 dB / Entfernung

Schall: Beugung an einer Kugel

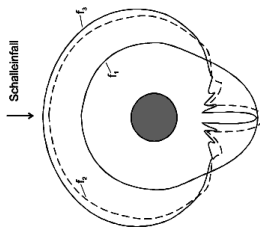


Abbildung 2.4: Beugungsbild bei Schallbeugung an einer Kugel für unterschiedliche Frequenzen, wobei $\theta_1 < \theta_2 < \theta_3 < \theta_4$ [Kollmeier]

Bei niedrigen Frequenzen stellt die Kugel fast kein Hindernis dar, während bei hohen Frequenzen der Grenzfali der **geometrischen Akustik** bereits angenähert ist, d. h. eine starke akustische Schattenwirkung und eine geringe seitliche Streuung. Interessanterweise entsteht immer auf der genau der Schalleinfallrichtung abgekehrten Seite („auf der akustischen Achse“) ein Schalldruckmaximum, das man sich nach dem Babinetischen

Raumwahrnehmung

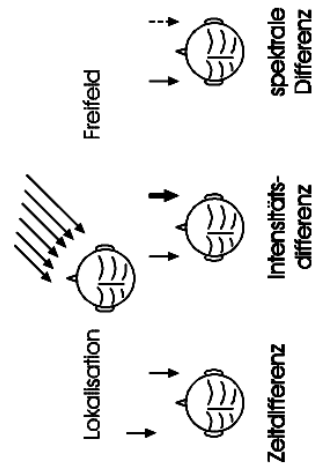


Abbildung 5.11: Physikalische Merkmale für die Lokalisation [Kollmeier]

- für 3D-Audio müssen alle Effekte beachtet werden
- zusätzlich Doppler-Effekt für bewegte Quellen

Raumwahrnehmung

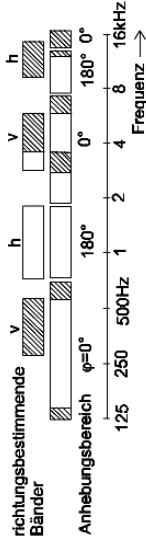


Abbildung 2.2: Richtungsbestimmende Bänder nach Blauert (aus Blauert, J.: Räumliches Hören. Nachschrift. S. Hirzel Verlag, Stuttgart 1985) Die erste Zeile gibt den Frequenzbereich an, bei dem in psychoakustischen Tests die Anwesenheit (schraffiert) oder Abwesenheit (nicht schraffiert) einer spektralen Anregung zu einer Lokalisation nach vorne („v“) bzw. hinten („h“) führt. Die zweite Spalte zeigt analoge Auswertungen aus Übertragungsfunktionen für Schalleintall von vorn ($\phi = 0^\circ$) bzw. von hinten ($\phi = 180^\circ$). [Kollmeier]

Raumwahrnehmung: Cone of confusion

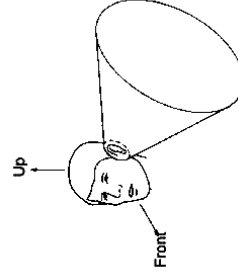
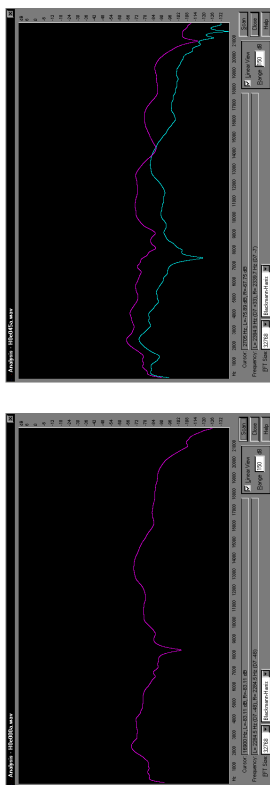


Abbildung 2.1: Cone of confusion, d.h. Fläche im dreidimensionalen Raum, auf der sämtliche Schalleintallrichtungen dieselbe interaurale Zeit- und Pegeldifferenz aufweisen (aus Moore, B. C. J.: Perceptual Consequences of Cochlear Damage. Oxford University Press, Oxford 1995) [Kollmeier]

Raumwahrnehmung: HRTF



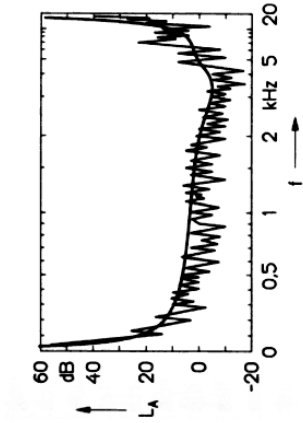
45 Grad seitlich

von vorne

Head-related transfer function

- individuelle Kopf (Außenohr)- Übertragungsfunktion
- mißt Intensität des Schalls am Trommelfell, links/rechts separat
- außerordentlich komplexe Winkel/Frequenzabhängigkeit
- Beispieldaten: sound.media.mit.edu/KEMAR.html

absolute Hörschwelle

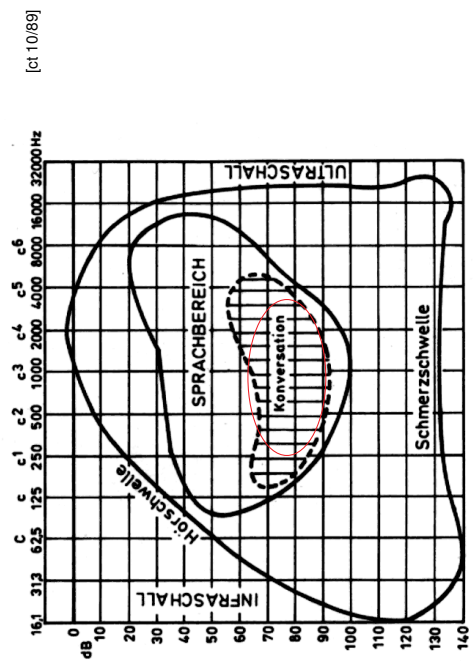


[Terhardt]

Abb. 9.11. Monaurale Absoluthörschwelle für Sinustöne. Glatte Kurve: Mittelwert des Schwellenpegels L_A von normalhörenden Personen. Gezackte Kurve: Mit der Methode des pendelnden Regelns (Békésy tracking) gemessen (25jähriger Mann, rechtes Ohr). Abszisse: Tonfrequenz, SPINC-skaliert; Ordinate: Schallpegel

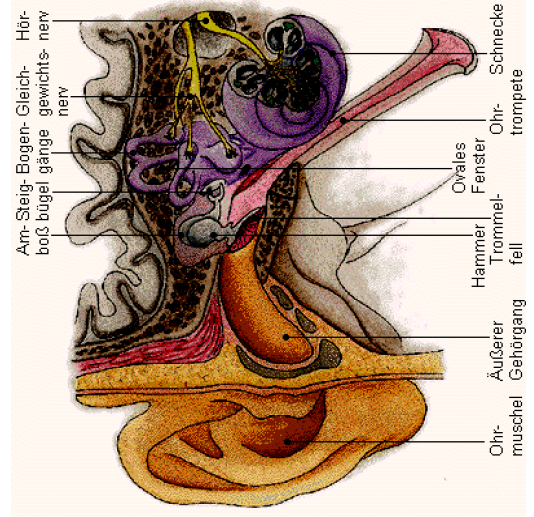
- geringe Empfindlichkeit bei tiefen und hohen Tönen
 - maximale Empfindlichkeit zwischen 2..5 KHz
- => Loudness-Kurven

Hörbereich und Sprachbereich



gute Sprachverständlichkeit im "Hauptsprachbereich" (ca. 300-3500 Hz)

Ohr: Schema



drei Untereinheiten:

- Außenohr
 - Mittelohr
 - Innenohr
- bzw.:
- Verstärkung, Delay
 - Impedanzwandlung
 - Frequenzanalyse

(glasklar Mensch 3D)

Gehör: Modell der Signalverarbeitung

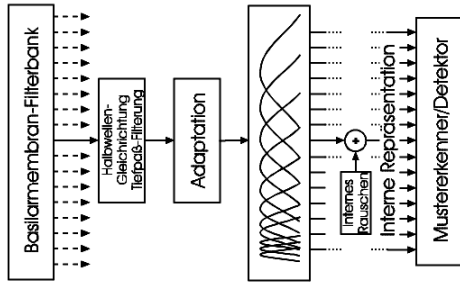
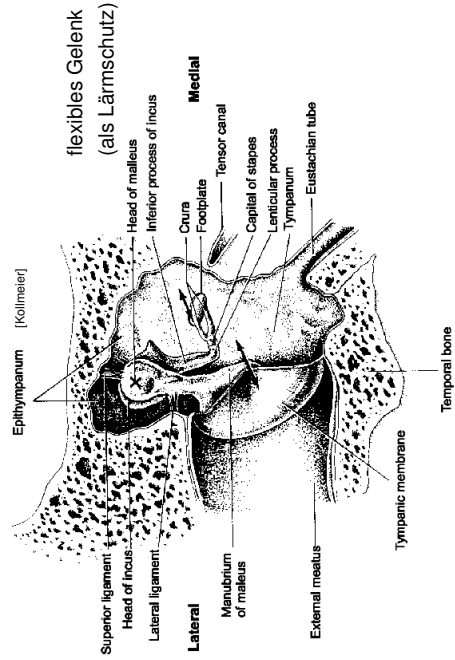


Abbildung 1.4: Modell der effektiven Signalverarbeitung nach Dau et al. (1987) [Kollmeier]

Digitale Audioverarbeitung | WS 2000 | 18.205

Mittelohr



- Verstärkung (Impedanzwandlung), Übergang Luft nach Wasser
- Trommelfeld ca. 20fache Fläche des ovalen Fensters

Digitale Audioverarbeitung | WS 2000 | 18.205

Ohr: Schema

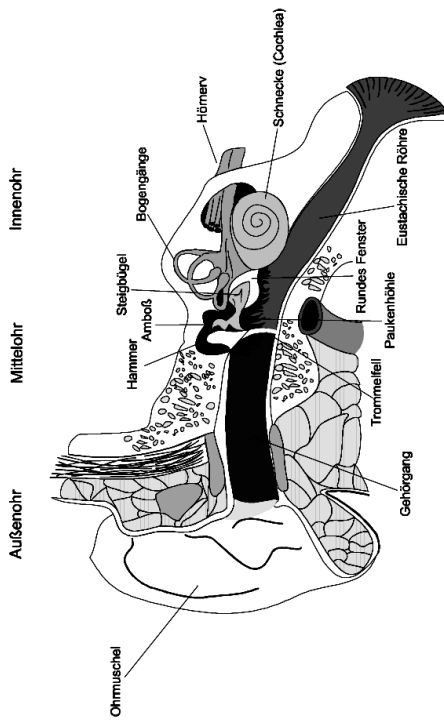


Abbildung 1.1: Übersicht über den Aufbau des Gehörs [Kollmeier]

Digitale Audioverarbeitung | WS 2000 | 18.205

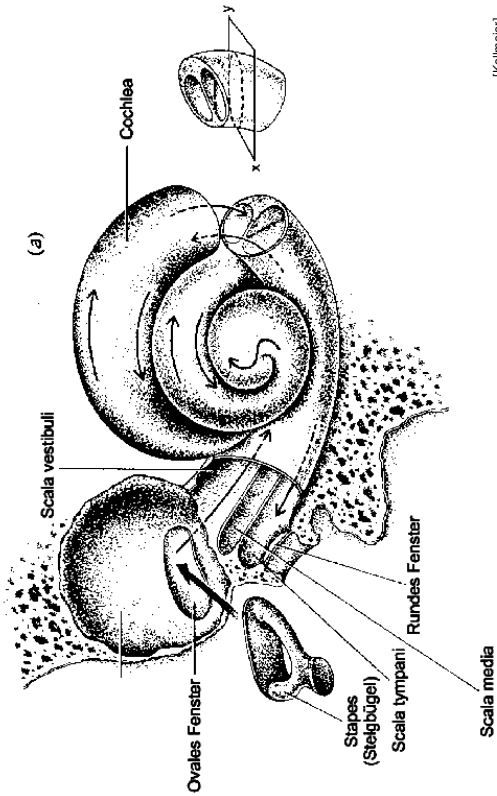
Gehör: Teile, Funktion, Störungen

Anteil des Gehörs	Funktion	Störung	Ätiologische Diagnosezeit
Außenohr	Richtungsabh., Filterung / Bündelung	Ohrmitbildung	klinisch
Mittelohr	Impedanztransformation Luft → Wasser	Schallleitungs-Schwachhörigkeit (z.B. Erguß, Otitis media)	Otoskopie
Innenohr	Umwandlung Wasserschall → Nerven- erregung, Frequenz-Orts-Transformation	Cochleäre Schwerhörigkeit (z.B. Lärm, Alter, ...)	Impedanzaudiometrie Stimmgabeltest Ton-Audiogramm Otoskopische Emissionen Überschwellige Tests
Hörnerv / Hirnstamm	Kodierung akustischer Information, Auswertung interauraler Unterschiede, Modulations-Kodierung	Retrocochleäre (neurale) Schwerhörigkeit, Lokalisationsstörung	Hirnstamm-Audiometrie Sprachaudiometrie Bildgebende Verfahren
Cortex	Sprachwahrnehmung, komplexe Verar- beitung	zentrale Hörstörungen (z.B. Aphasie)	Zentrale Sprachtests

[Kollmeier]

Digitale Audioverarbeitung | WS 2000 | 18.205

Innenohr: Aufbau und Cochlea



[Kollmeier]

Innenohr: cortisches Organ

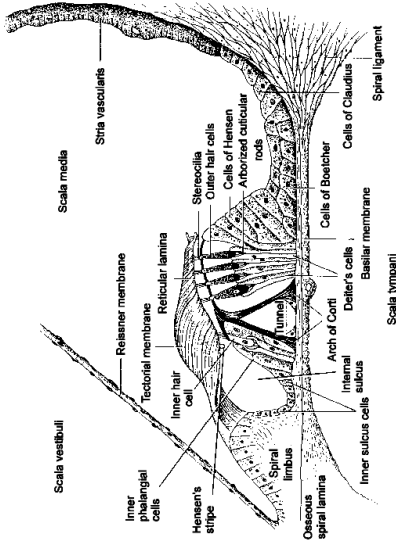
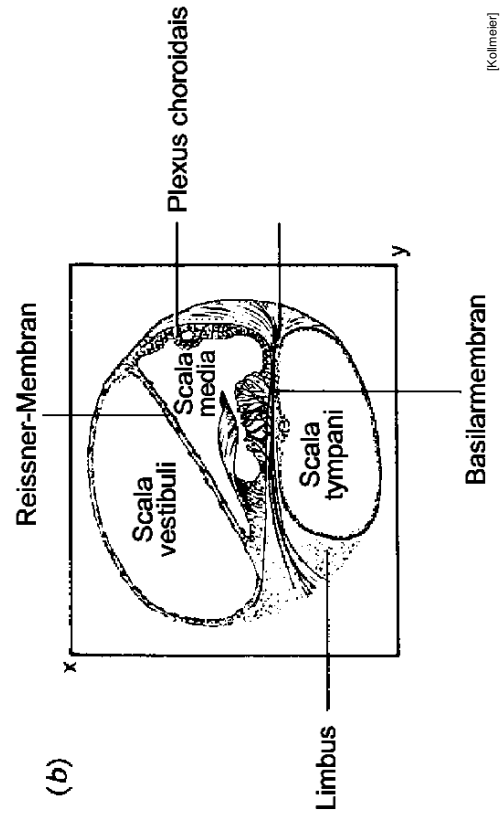


Abbildung 3.7: Schematische Darstellung des Cortischen Organs (aus Gulick, W. L., Gescheider, G. A., Frisina, R. D.: Hearing. Oxford University Press, Oxford 1989)

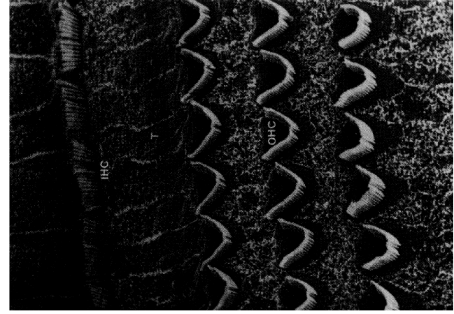
[Kollmeier]

Innenohr: cortisches Organ



[Kollmeier]

Innenohr: Haarzellen



innere Haarzellen
"Sensoren"

äußere Haarzellen
"Aktoren"

Abbildung 3.8: Detailansicht der Innenohr-Anatomie, IHC = Innere Haarzellen, OHC = Äußere Haarzellen (aus Dallos, P., Popper, A. N., Fay, R. R. (Hrsg.): The Cochlea, Springer-Verlag, New York 1986)

Cochlea und Nerven

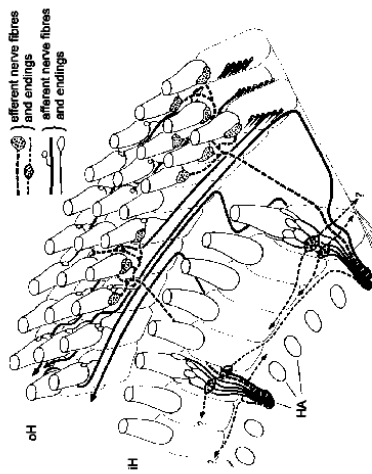


Abbildung 4.5: Schematische Ankopplung der Haarzellen an den Hörnerven (oH: äußere Haarzellen, iH: innere Haarzellen) (aus Luce, R.D.: Sound and Hearing. Lawrence Erlbaum Verlag, New Jersey 1993) [Kollmeier]

Digitale Audioverarbeitung | WS 2000 | 18.205

Basilarmembran

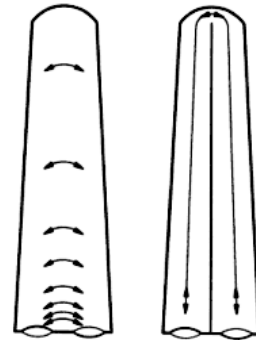


Abb. 9.9. Wenn die Basilarmembran nicht vorhanden wäre (obere Skizze, ausgestreckter Cochleakanal), würde die Flüssigkeitsbewegung sich bei Auslenkung des ovalen Fensters wegen der inneren Reibung auf den basalen Teil (links) konzentrieren. Wenn dagegen der Kanal über einen Teil seiner Länge durch eine elastische Membran unterteilt wird (unten), wird bei sehr langsamer Auslenkung die Strömung um die Membran herum gezwungen, während rasche Auslenkung wegen der Massenträgheit der bewegten Flüssigkeitssäule zur Querauslenkung der Membran führt, und zwar umso weiter links, je rascher die Auslenkung erfolgt [Terhardt]

- mechanische Lösung zur Frequenzanalyse

Digitale Audioverarbeitung | WS 2000 | 18.205

Cochlea: Halbwellengleichrichtung

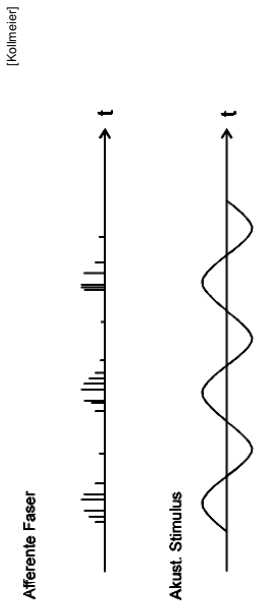


Abbildung 4.6: Halbwellengleichrichtung bei der Kodierung eines akustischen Stimulus in der Haarzelle bzw. im nachgeschalteten auditorischen Nerv (schematisch). Dargestellt ist die Aktivität des Hörnerven, die nur während der positiven Halbwellen des Eingangsstimulus auftritt [Kollmeier]

- innere Haarzellen reagieren nur auf Dehnung
- Nerven-Spikes nur für positive Halbwellen

Digitale Audioverarbeitung | WS 2000 | 18.205

Haarzellen: otoakustische Emission

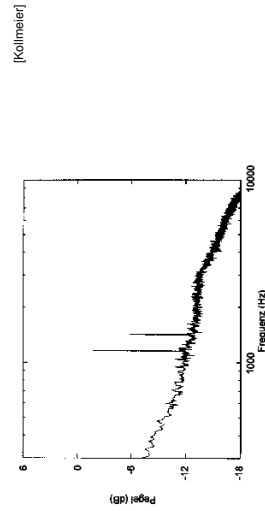


Abbildung 3.17: Gemitteltes Leistungsspektrum bei der Bestimmung einer spontanen otoakustischen Emission, die durch vereinzelte Spitzen im Spektrum gekennzeichnet ist. (aus Uppenkamp, S.: Grundlagen und Anwendung der otoakustischen Emissionen zur objektiven Hörprüfung. Dissertation, Universität Göttingen 1992) [Kollmeier]

- Haarzellen sind aktive Steilglieder
- harmlose spontane Emission, bis zum Tinnitus

Digitale Audioverarbeitung | WS 2000 | 18.205

Lautheit

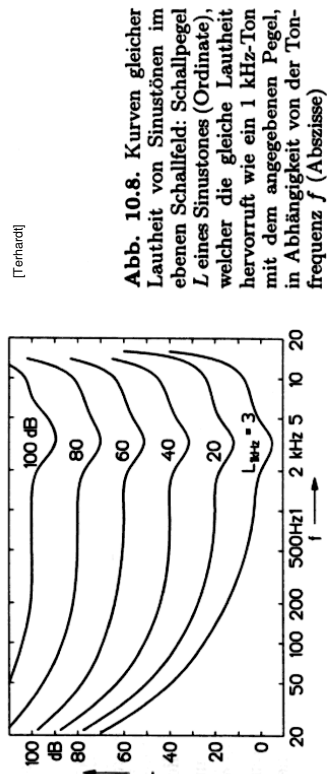


Abb. 10.8. Kurven gleicher Lautheit von Sinustönen im ebenen Schallfeld: Schallpegel *L* eines Sinustones (Ordinate), welcher die gleiche Lautheit hervorruft wie ein 1 kHz-Ton mit dem angegebenen Pegel, in Abhängigkeit von der Tonfrequenz *f* (Abszisse)

- bei leisen Tönen schwache Wahrnehmung von Höhen/Bässen
- Empfindlichkeitsmaximum bei ca. 3 KHz
- "loudness"-Regler am HiFi-Verstärker zur Kompensation

Tuningkurven: Schema

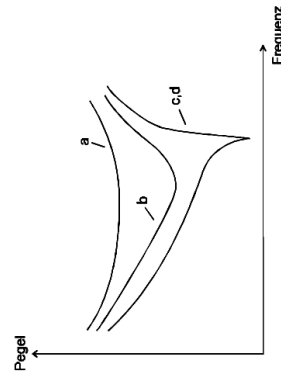


Abbildung 3.6: Schematischer Verlauf von „Tuning-Kurven“, d. h. des Pegels eines Sinustons, der zum Erreichen einer bestimmten Auslenkung der Basilar membran (bzw. zu einer bestimmten Erhöhung der Nervenfaser-Feuerrate gegenüber der Spontarrate) notwendig ist, als Funktion der Frequenz. Dargestellt sind schematisch die Ergebnisse für (a) optisch/stroboskopisch (v. Békésy), (b) Mißbauertechnik, (c) neuere Laser-Interferometrische Messungen, (d) neurophysiologisch

[Kollmeier]

Ohr und Hörbahn

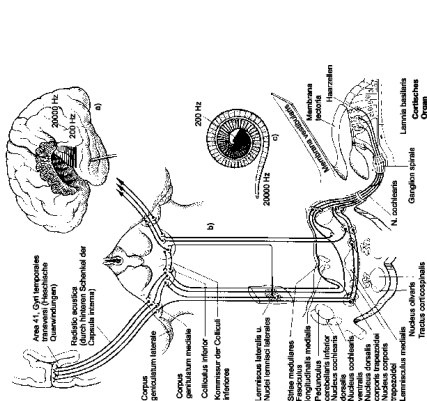


Abbildung 4.12: Übersicht über die Anatomie der Hörbahn (aus Daus, P.: Neurologisch-topische Diagnostik, Thieme Verlag, Stuttgart 1995).

[Kollmeier]

Inter-Spike Histogramme

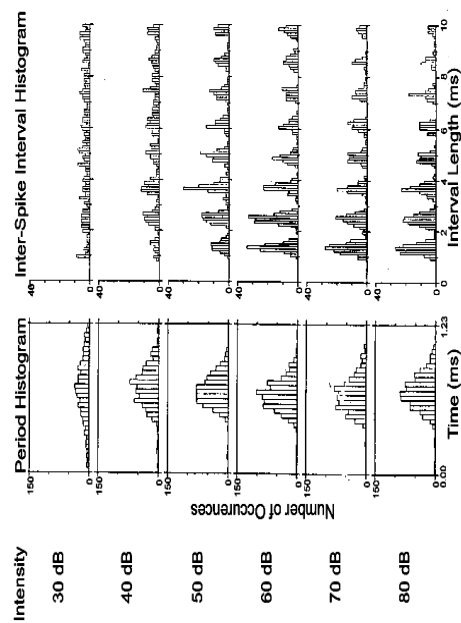
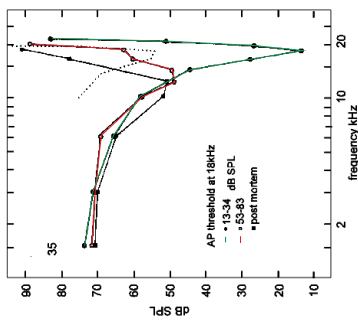


Abbildung 4.9: Periodenhistogramm und Inter-Spike-Intervall-Histogramm für eine Hörnervenfaser bei systematischer Erhöhung der Intensität eines sinusförmigen Reizes von ca. 800 Hz von 30 dB auf 80 dB.

[Kollmeier]

Tuningkurven: gemessen

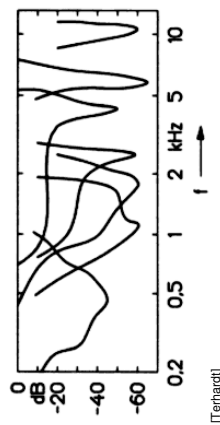


- hohe Güte
- keineswegs passiv
- aktive Regelung

Abbildung 4.7: Schematische Tuning-Kurve (Reizintensität an der Schwelle als Funktion der Frequenz des Stimulus) für zwei verschiedene sensitive Hörsinnesfasern sowie für den Zustand nach abgestellter Blutzufuhr (post mortem). (aus Moore, B. C. J.: Frequency Selectivity in Hearing. Academic Press, London 1986)

[Kollmeier]

Tuningkurven: Katzen



[Terhardt]

Abb. 9.10. Tuningkurven, gemessen an anästhesierten Katzen. Jede Kurve gehört zu einer anderen Faser des akustischen Nerven. Die Ordinate gibt denjenigen Schallpegel eines Sinustones an, welcher zur Erregung einer festen Anzahl von Aktionspotentialen pro Zeiteinheit erforderlich ist. Abszisse: Frequenz des Sinustones. Nach King [50]

- hohe Güte der Frequenzselektion
- "aktive" Frequenzbereiche der Fasern überlappen
- menschliches Gehör entsprechend

Maskierung

Mithörschwelle eines Sinustones bei Verdeckung durch Schmalbandrauschen. In Abb.11.4 sind drei Mithörschwellen von Sinustönen verdeckt durch Schmalbandrauschen der angegebenen Mittenfrequenzen dargestellt [1125]. Die Fourier-Spektren der maskierenden Schmalbandgeräusche sind durch Verwendung von Bandpassfiltern hoher Flankensteilheit auf Bandbreiten von ungefähr 100, 160 und 700 Hz beschränkt [CD 2].

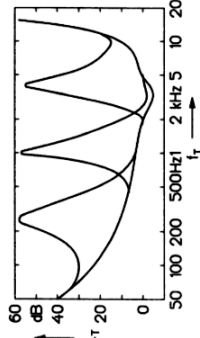
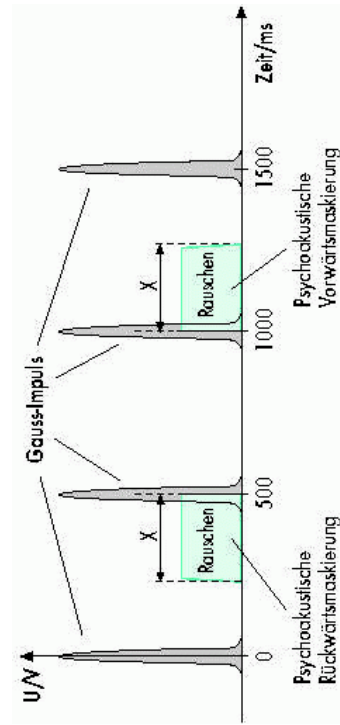


Abb. 11.4. Mithörschwellen von Sinustönen bei Verdeckung durch frequenzgruppenbreite Schmalbandrauschen mit den Mittenfrequenzen 250, 1000, 4000 Hz und dem Maskierpegel $L_M = 60$ dB. Nach [1125]

[Terhardt]

- Frequenzanalyse in der Cochlea nicht perfekt
- laute Töne verdecken benachbarte leisere
- Ausnutzung in psychoakustischen Modellen (MiniDisc, MP3, usw)

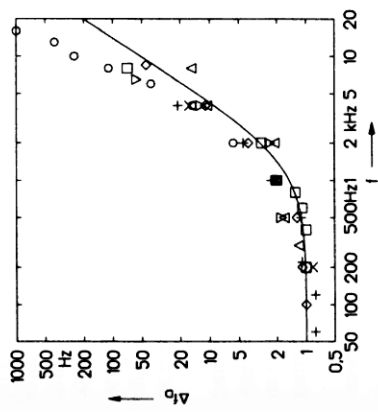
Maskierung: temporal



[cf]

- laute Töne verdecken spätere leisere Töne
- und frühere leisere Töne (!)
- Ausnutzung in psychoakustischen Modellen (MiniDisc, MP3, usw)

Frequenzauflösung



- bis 1 Hz Frequenzauflösung (bei geübten Personen)

Tuningkurven vs. Tonhöhe

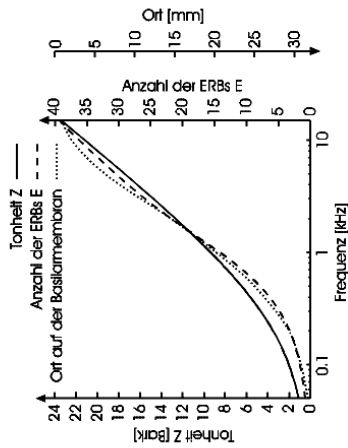


Abbildung 5.3: Mel-Skala, Bark-Skala, ERB-Skala

- wahrgenommene Tonhöhe korrespondiert mit Position
- passende Skalen

Frequenzgruppen

Tabelle 9.1: Frequenzgruppen nach Zwicker 1982

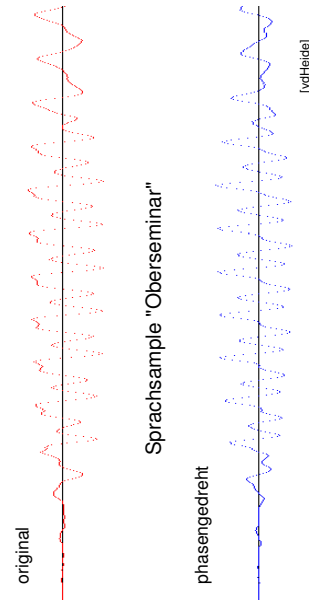
z/Bark	f_z /Hz	f_c /Hz	Δf_c /Hz	f_m /Hz
0	100	0	100	50
1	200	100	100	150
2	300	200	100	250
3	400	300	100	350
4	510	400	110	450
5	630	510	120	570
6	770	630	140	700
7	920	770	150	840
8	1080	920	160	1000
9	1270	1080	190	1170
10	1480	1270	210	1370
11	1720	1480	240	1600
12	1790	1720	280	1850
13	2000	1790	320	2150
14	2320	2000	380	2500
15	2700	2320	450	2900
16	3150	2700	550	3400
17	3700	3150	700	4000
18	4400	3700	900	4800
19	5300	4400	1100	5800
20	6400	5300	1300	7000
21	7700	6400	1800	8500
22	9500	7700	2500	10500
23	12000	9500	3500	13500
24	15500	12000	3500	13500

[Zwicker]

Tonhöhenwahrnehmung?

- Messungen
- vgl. Tuningkurven
- tiefe Frequenzen: ca. 100 Hz Bandbreite
- hohe Frequenzen: log. Breite der Gruppen => "Bark" Frequenzskala
- Ausgangspunkt für psychoakustische Modelle

Phasen: Hilbertfilter



- Ohr unempfindlich gegen Phaseninformation
- verändertes Signal klingt einzeln (vollkommen) gleich
- erst bei Interferenzen mehrerer Signale hörbar
- Phasengang vieler Systeme (Telefon!) extrem schlecht

Klangfarben

akustisches Signal mit vielen Teiltönen:

- keine Wahrnehmung aller Einzelöne
- sondern Zuordnung zu Quellen/Instrumenten
- tiefste Frequenz bestimmt die wahrgenommene Tonhöhe
- Obertonspektrum identifiziert die Quelle
- mit Eigenschaft "Klangfarbe"
- Klangfarbe auch stark vom "attack" des Klangs abhängig (=> Synthesizer mit kurzen Samples plus Schleifen)
- entsprechend automatische "Sprecheridentifikation"
- komplexe Zusammenhänge für Akkordwahrnehmung

Klangfarben: Spektrogramm



[Terhardt]

Abb. 1.9. Beispiel eines Teilkonzertmusters. Ausschnitt einer mehrstimmigen Blasmusikdarbietung. Die hauptsächlich beteiligten Instrumente sind Trompete, Posaune, Klarinette und Tuba. Die instantane Teilton-Amplitude wird durch die Strichstärke angegeben. Ordinate Bark-skaliert. Nach [969]

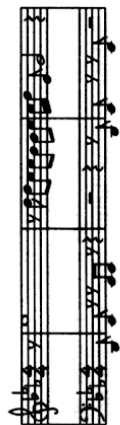


Abb. 1.10. Transkription der Trompeten- und der Tuba-Stimme des Musikbeispiels, dessen Teilkonzertmuster in Abb. 1.9 dargestellt ist

Klangfarben: Beispiele

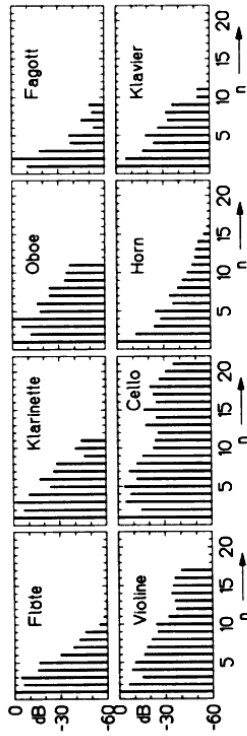
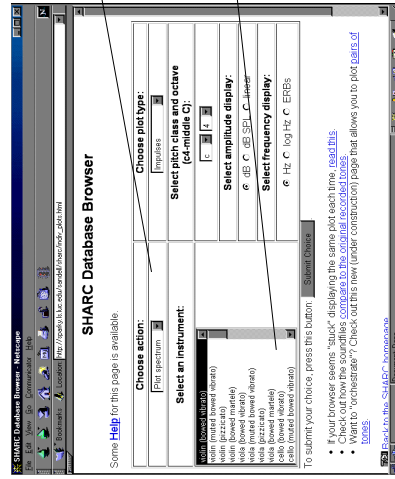


Abb. 8.12. Teiltonspektren verschiedener Musikinstrumente schematisch

[Terhardt]

- Streicher mit reichem Obertonspektrum
- Flöten mit sehr reinen Tönen (insb. Blockflöte)
- Klavier stark anschlag- und tonhöhenabhängig
- aber auch "Attack-Phase" maßgeblich für Klangwahrnehmung

Klangfarben: SHARC-Datenbank

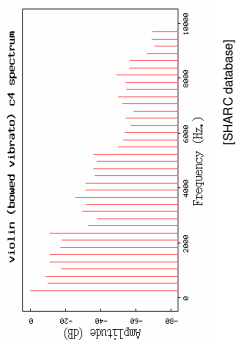
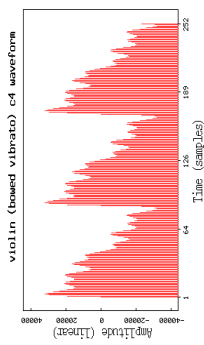


Samples oder zugehörige Spektren

viele Instrumente diverse Tonhöhen Spieltechniken

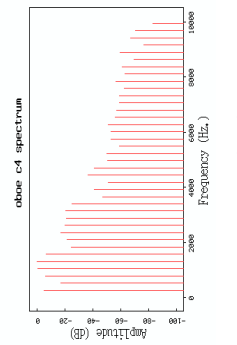
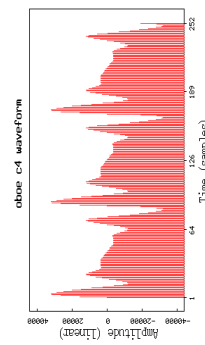
- www.sparky.is.luc.edu/sandell/sharc/indiv_plots.html
- schönes Web-Interface zu Samples und Spektren
- kein Zusammenhang mit den "SHARC" Signalprozessoren

Klangfarben: Violine



- Sägezahnform: reiches Obertonspektrum
- auch viele gerade (angenehme) Obertöne
- typisch für Streichinstrumente

Klangfarben: Oboe



- irreguläre Wellenform
- außergewöhnlich reiches Obertonspektrum
- andere Blasinstrumente mit weit weniger Obertönen
- z. B. Blockflöten fast reine Sinusklänge (plus Anblasgeräusche)

Sprache

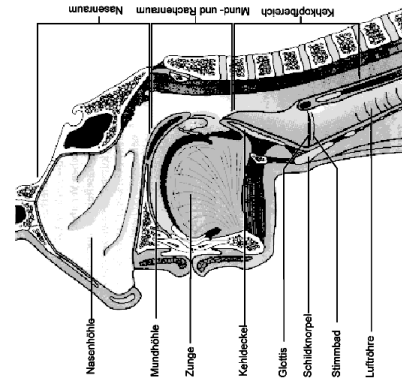
"niedere" Aspekte:

- Sprachsignale, Spektren, Formanten, ...
- Trennung stimmhafte / stimmlose Laute
- Synthese / Analyse von Sprachsignalen
- low-bitrate Kodierung (z.B. für Mobiltelefone)

"höhere" Sprachverarbeitung:

- NATS (von Hahn, Menzel)
- WSV (Habel, Freksa)

Sprache: Vokaltrakt



- Glottis, Stimmband
- Klang zusätzlich durch Zungen-/Lippenstellung
- typische Spektren: Mann / Frau / Kind normal / erregt

Abbildung 6.1: Vokaltrakt im Querschnitt (schematisch) (aus Kießling, J., Kollmeier, B., Diller, G.: Versorgung und Rehabilitation mit Hörgeräten. Thieme Verlag, Stuttgart 1997) [Kollmeier]

Sprache: Spektrogramm "Durst"

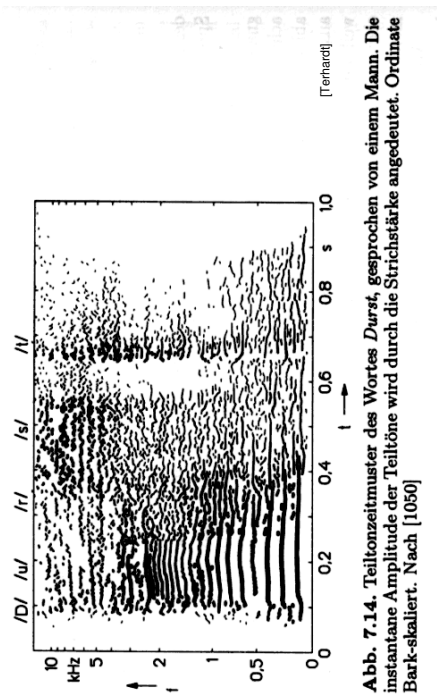


Abb. 7.14. Teilzeitmuster des Wortes *Durst*, gesprochen von einem Mann. Die instantane Amplitude der Teiltöne wird durch die Strichstärke angedeutet. Ordinate Bark-skaliert. Nach [1050]

- außerordentlich komplex: Problem Spracherkennung

Sprache: Spektrogramm

Formanten, d. h. die Frequenzbereiche, bei denen der Vokaltrakt eine besonders hohe Verstärkung des akustischen Signals bewirkt. Die Formanten stellen sich damit als Spitzen im Spektrum dar und sind charakteristisch für den jeweils artikulierte Vokal (s. unten).

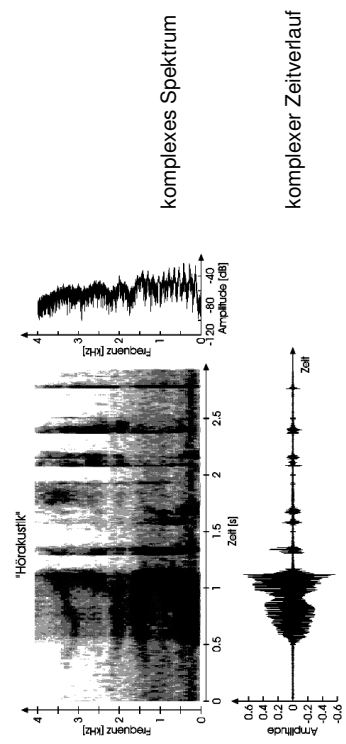


Abbildung 6.2: Zeit-Signal (unten), Spektrum (rechts) und Spektrogramm (links oben) des Wortes „Hörakustik“ [Kollmeier]

Sprache: Spektren, Formanten

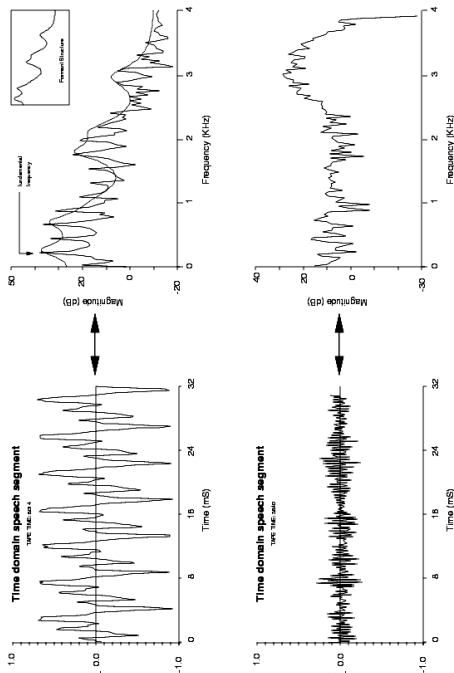


Fig. 1. Voiced and Unvoiced Segments and their short-time spectra. [Spanias]

Sprache: Artikulation

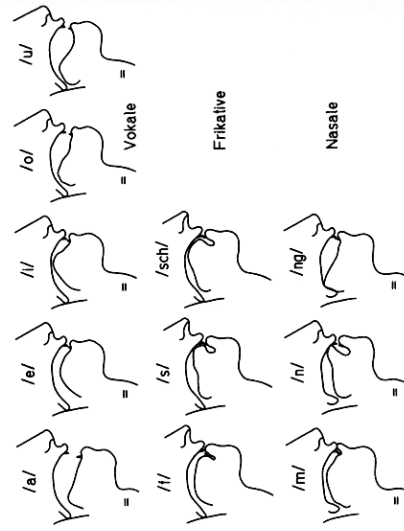


Abb. 7.1. Schematische Darstellung der Artikulation einiger Sprachlaute

- Zungen- und Lippenstellung maßgeblich

[Terhardt]

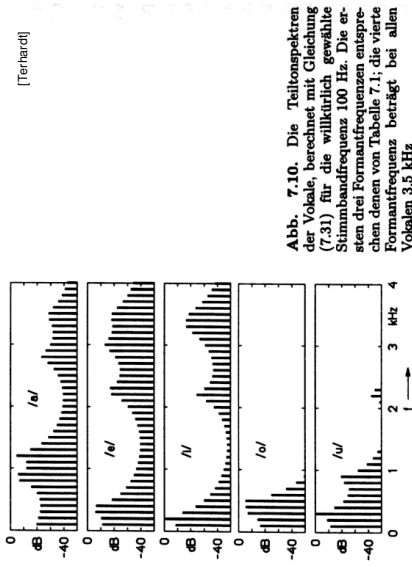
Sprache: Vokaltrakt

• **Stimmband + Resonanzrohr**
 • **Legen der Formanten:**
 Wellenlänge: $L/4 = 17\text{cm}$
 $F_1 = c/L = 340\text{m/s} / 68\text{cm} = 500\text{ Hz}$
 $F_2 = 1500\text{ Hz}$
 $F_3 = 2500\text{ Hz}$
 $F_4 = 3500\text{ Hz}$

Sprache: Konsonanten bdg

/b/-Schwa
 /d/-Schwa
 /g/-Schwa

Sprache: Teiltonspektren der Vokale



- Spektren der Vokale, Formanten deutlich sichtbar

Sprache: Vokaldreieck

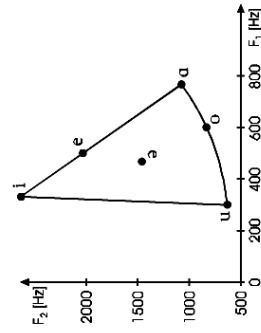


Abbildung 6.1.4: Vokaldreieck. Aufgetragen ist die Frequenz des ersten Formanten F_1 (Abszisse) und des zweiten Formanten F_2 (Ordinate) für unterschiedliche Vokale (schematisch) [Kollmeier]

- typische Frequenzen / Formanten der Vokale
- verhältnismäßig einfache Erkennung von stimmhaften Lauten

Lärm: Pegelbereiche

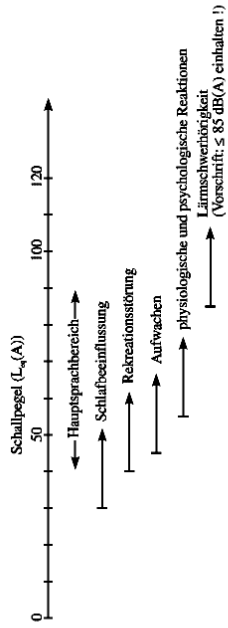
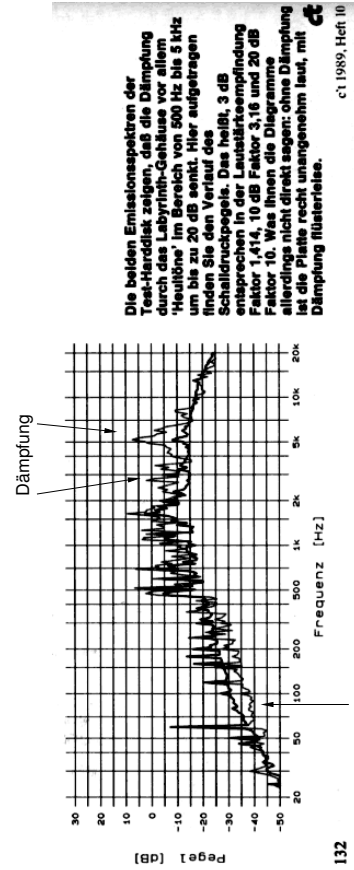


Abbildung 3.12: Pegelbereiche, für die unterschiedliche Schallwirkungen beim Menschen auftreten [Kollmeier]

- Lärmpegel an Straßen selten unter 50 dB(A) ...
- laufender Computer:
- Nadeldrucker: 70+ dB(A)

Lärm: Festplatte normal und gedämpft



Die beiden Emissionsspektren der Test-Harddisk zeigen, daß die Dämpfung durch das Laubring-Gehäuse vor allem 'Heulöne' im Bereich von 500 Hz bis 5 kHz um bis zu 20 dB senkt. Hier aufgetragen finden Sie den Verlauf des Schalldruckspekts. Das heißt, 3 dB entsprechen in der Lautstärkeempfindung Faktor 1,414, 10 dB Faktor 3,16 und 20 dB Faktor 10. Was Ihnen die Diagramme allerdings nicht direkt sagen: ohne Dämpfung ist die Platte recht unangenehm laut, mit Dämpfung flüsterleise.

c1 1989, Heft 10

hier sogar lauter

- Schalldämpfung ist nicht einfach...

Sprache: Vokalkategorien

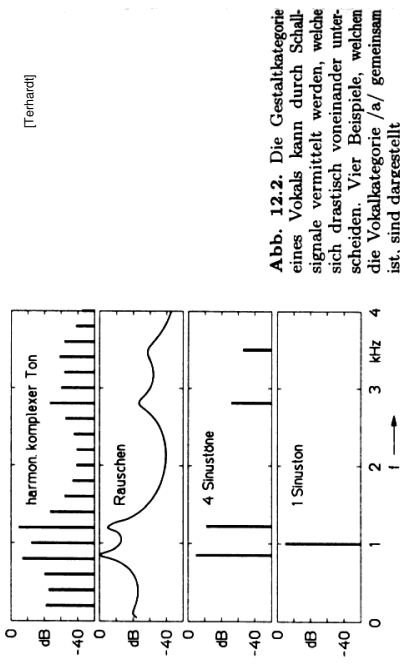
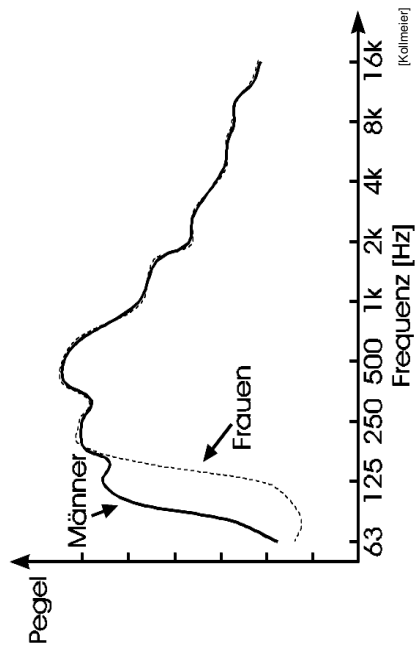


Abb. 12.2. Die Gestaltkategorie eines Vokals kann durch Schallsignale vermittelt werden, welche sich drastisch voneinander unterscheiden. Vier Beispiele, welchen die Vokalkategorie /a/ gemeinsam ist, sind dargestellt

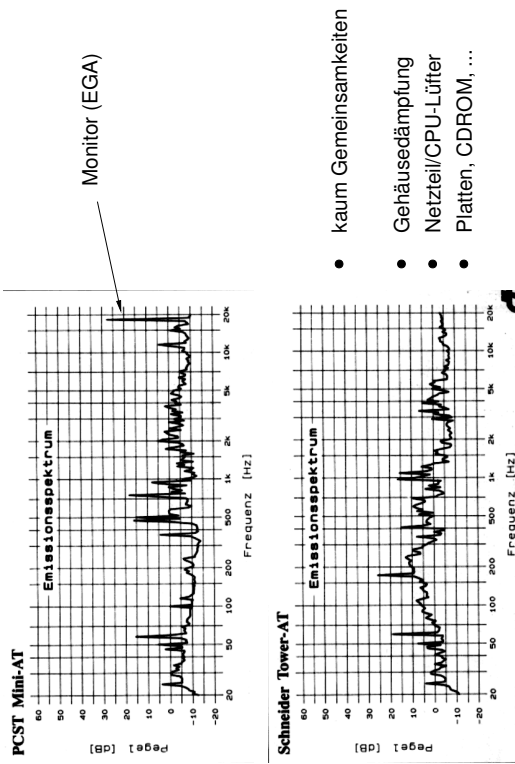
- Wahrnehmung: jeweils als Vokal /a/ => perceptual speech coding

Sprache: Frequenzbereiche



- Frequenzspektrum oberhalb 200 Hz identisch
- Männer mit tieferen Frequenzen zw. 60 und 250 Hz

Lärm: Emissionsspektren, 1989



Digitale Audioverarbeitung | WS 2000 | 18.205

Lärm: Schädigungen

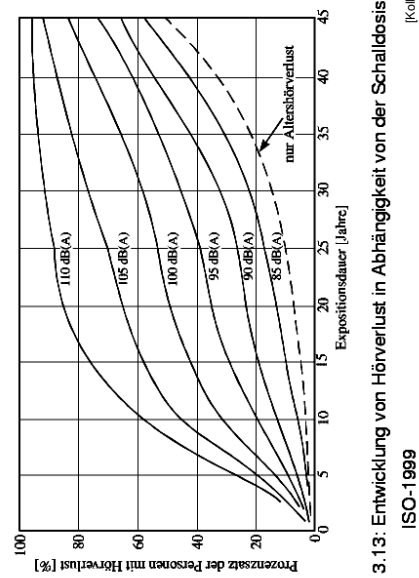
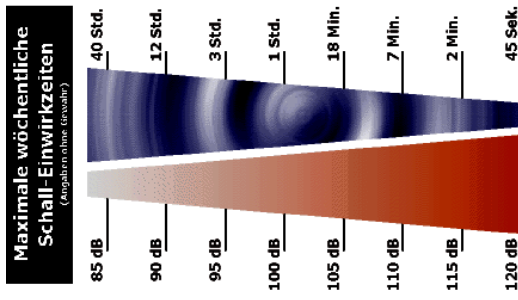


Abbildung 3.13: Entwicklung von Hörverlust in Abhängigkeit von der Schallintensität gemäß ISO-1999

- bei höheren Pegeln Schädigen innerhalb von Sekunden

Digitale Audioverarbeitung | WS 2000 | 18.205

Lärm: Schädigung



Digitale Audioverarbeitung | WS 2000 | 18.205

Gehör: Zusammenfassung

Gehör außerordentlich leistungsfähig:

- überdeckt 13 Größenordnungen Schalldruck
 - fast optimale Empfindlichkeit
 - Hörbereich etwa 20Hz .. 20 KHz
 - Frequenzanalyse "mechanisch" in der Cochlea
 - komplexe Verarbeitung/Erkennung im Gehirn
 - Wahrnehmung Klang vs. Geräusch
- insbesondere zur Raumwahrnehmung:
- Zeit- und Intensitätsdifferenzen der Signale
 - Kopfübertragungsfunktion
 - Auswertung von Echos

=> entsprechende Modelle für effiziente Signalverarbeitung

Digitale Audioverarbeitung | WS 2000 | 18.205