

Binaurales Hören und die Hörgeräte-Technologie der Zukunft

Markus Bodden

Zusammenfassung Im Vergleich zum einohrigen Hören eröffnet das beidohrige Hören eine Vielzahl von Vorteilen. Durch eine beidohrige Hörgeräte-Versorgung wird angestrebt, diese Vorteile auch Hörgeschädigten nutzbar zu machen. Um eine optimale Unterstützung der binauralen Resthörfähigkeiten Hörgeschädigter zu erreichen, müssen die dem binauralen Hören zugrundeliegenden Verarbeitungsmechanismen bekannt sein. Daher werden die Grundlagen des Richtungshörens, die das Gehör zur Bestimmung des Azimut, der Elevation und der Schätzung der Schallquellenentfernung anwendet, hier näher erläutert. Ausgehend von diesen Mechanismen werden die Probleme einer originalgetreuen elektroakustischen Übertragung bezüglich des räumlichen Hörens angesprochen und auf die beidohrige Hörgeräte-Versorgung von Hörgeschädigten übertragen. Ein Konzept für die zukünftige Hörgeräte-Versorgung wird vorgestellt, das eine optimale Unterstützung der binauralen Resthörfähigkeiten erlaubt. Dieses Konzept eines zentral gesteuerten binauralen Hörgerätes bietet den Vorteil, daß auch Hörschäden mit stark eingeschränktem oder völlig fehlendem binauralen Resthörvermögen besser versorgt werden können. Ein zentraler Prozessor, der den Kern dieses binauralen Hörgerätes bildet, könnte Teilbereiche des beidohrigen Hörens simulieren und so beispielsweise eine wirksame Störgeräusch-Unterdrückung durchführen. Das Forschungsmodell eines sogenannten »Cocktail-Party-Prozessors« wird beschrieben. Die Fähigkeit dieses Prozessors, die Verständlichkeit der durch sprachliche Störgeräusche gestörten Sprache zu verbessern, wurde anhand von Sprachverständlichkeits-Tests mit Hörgeschädigten nachgewiesen.

Schlüsselwörter: Beidohriges Hören;
Hörgerät;
binaurale Merkmale;
»Cocktail-Party-Prozessor«

Binaural Hearing and Advanced Hearing-Aid Technology

Summary *Binaural hearing offers a lot of advantages compared to monaural hearing. In order to support binaural abilities of hearing impaired persons it is necessary to understand the underlying processes of the binaural hearing system. This article summarizes some fundamentals of binaural hearing. For the mechanisms to estimate azimuths, elevations, and distances of sound sources head-related transfer functions play the essential role. They describe parameters encoding the direction of sound incidence that can be recorded at the eardrum of a listener. The binaural system decodes simultaneously binaural cues (interaural differences in time and intensity) and monaural cues (spectral cues). An important fact is that the interaural cues cannot be regarded independently. Their characteristic combinations are individual for each person. This may cause problems if signals have to be recorded and presented with an electroacoustic equipment. If the characteristics of the head-related transfer functions used in the recording equipment do not perfectly agree with those of the listener, front-back reversals and localization of sound sources inside of the head may occur. If, for example, due to an unsymmetric hearing loss the interaural intensity differences are changed and do not fit any more to the corresponding interaural time differences, a single sound source may split up into several hearing events confusing the listener. In order to avoid these problems, a strategy of a binaural hearing aid is proposed. This binaural hearing aid consists of two hearing aids that are controlled by a central processor. In a first step, this processor could simply adjust the amplifications of both hearing aids in order to control the interaural parameters. In cases where binaural hearing abilities are reduced, a central processor as a part of the hearing aid could simulate some of these binaural abilities. The aim is to implement a powerful speech enhancement method that is able to suppress interfering speech signals. Such a 'Cocktail-Party Processor' has been developed at the Ruhr-University. The fundamentals of the system are described. This system shall simulate binaural processes as being performed by the human auditory system. In intelligibility tests with hearing impaired persons the ability of the system to enhance the intelligibility of speech interfering with noise has been proven. The system can not yet be implemented into hearing aids, since the computational needs are too high and do not allow a sufficient miniaturization.*

Keywords: *binaural hearing;
hearing aid;
binaural cues;
Cocktail-Party Processor*

Einleitung

Im Vergleich zum einohrigen Hören eröffnet das beidohrige Hören eine Vielzahl von Vorteilen im täglichen Leben. Normalhörende halten diese Vorteile aber meist für so selbstverständlich, daß sie sich der Vorteile häufig nicht bewußt sind. Ein Grund hierfür mag darin begründet liegen, daß das Hörorgan im Gegensatz zum Sehorgan, mit dem es ansonsten einige Gemeinsamkeiten aufweist, nicht reflektorisch verschlossen werden kann. Ein weiterer, wesentlicher Unterschied zum Sehsinn besteht darin, daß das »Hörfeld« nicht wie das Sehfeld nach vorne beschränkt ist, sondern Information aus allen Richtungen ausgewertet wird. Die auffällige äußerliche Gemeinsamkeit zum Sehsinn besteht nun darin, daß das Gehör über zwei Empfangsorgane verfügt, was vergleichbar zum dreidimensionalen, räumlichen Sehen ein dreidimensionales räumliches Hören ermöglicht.

Das beidohrige Hören ermöglicht die folgenden Vorteile:

- die Positionen von Schallquellen können in einem dreidimensionalen Raum geortet werden. Dies wird in der Psychoakustik mit dem Begriff »Richtungshören« bezeichnet. Gerade aufgrund des eingeschränkten Sehfeldes kommt dem Richtungshören im Alltag eine entscheidende Bedeutung als Warnfunktion zu, da sich auf einen zubewegende Schallquellen eine Gefahr darstellen können, die eine entsprechende Reaktion erfordern.

- Nachhall, der durch Reflexionen von Schall an Wänden erzeugt wird, kann unterdrückt werden. Durch diesen Effekt wird die Sprachverständlichkeit auch in halligen Umgebungen aufrechterhalten.

- Störschall kann unterdrückt werden. Diese Störschall-Unterdrückung wird durch den sogenannten »Cocktail-Party-Effekt« beschrieben, der besagt, daß man sich in einer Situation, in der mehrere Menschen gleichzeitig reden, auf einen Sprecher konzentrieren und ihn verstehen kann. Diese Vorteile können jedoch nur bei intaktem binauralen Hörsystem erzielt werden. Die dem zugrundeliegenden Verarbeitungsmechanismen werden im folgenden Kapitel erläutert. Für eine ausführlichere Beschreibung des beidohrigen Hörens sei hier auf *Blauert* (1974, 1983) verwiesen.

Beidohriges Hören

Die Fähigkeiten des räumlichen Hörens basieren auf einem Codier-Dekodier-System. Die

Codierung wird durch die Außenohr-Übertragungsfunktionen vorgenommen, die die richtungsabhängige Filterwirkung des Systems bestehend aus Ohrmuschel, Kopf und Oberkörper beschreiben. Die Dekodierung erfolgt anschließend auf unterschiedlichen, höheren Ebenen des Gehörs. Durch das beidohrigen Hören sind wir in der Lage, eine räumliche Analyse des uns umgebenden Schallfeldes durchzuführen. Die Verarbeitungsmechanismen, die zu einer Lokalisation von Hörereignissen führen, können in drei Kategorien eingeteilt werden:

- das Gehör muß eine Rechts-links-Unterscheidung durchführen und den Einfallswinkel des Schalls in der Horizontalebene schätzen (Azimut);

- das Gehör muß den Erhebungswinkel aus der Horizontalebene schätzen (Elevation). Eng hiermit verbunden ist die Vorne-hinten-Unterscheidung;
- das Gehör muß die Entfernung der Schallquelle schätzen.

Diese Verarbeitungsmechanismen stützen sich auf die Auswertung sowohl monauraler als auch binauraler Merkmale. Sie werden im folgenden eingehender betrachtet.

Die interauralen Unterschiede zur Schätzung des Azimut sind:

- interaurale Zeitdifferenzen. Aufgrund des Wegunterschiedes bei seitlicher Auslenkung erreicht der Schall das dem Zuhörer zugewandte Ohr zuerst;

- interaurale Pegeldifferenzen. Aufgrund von Beugungseffekten und Abschattung am Kopf ist der Pegel am weiter entfernt liegenden Ohr im Vergleich zum zugewandten Ohr reduziert.

Jede dieser interauralen Differenzen enthält codiert Information über den Schallquellen-Azimut. Analysiert man die Auswertung durch das Gehör, so stellt man jedoch fest, daß diese interauralen Parameter nicht etwa unabhängig voneinander ausgewertet werden, sondern daß erst die Kombination beider interauraler Parameter zu einer korrekten Schätzung des Azimut führt. Hinzu kommt, daß diese Kombinationen der interauralen Parameter interindividuell unterschiedlich ausgeprägt sind, was durch die unterschiedliche Kopfgeometrie und Abmessung erklärt wird. Das Gehör hat folglich in einem Lernprozeß die Ausprägung der interauralen Parameter als charakteristisches Merkmal für die entsprechende Schalleinfall-Richtung erfaßt. In Abb. 1 ist oben die interaurale Zeitdifferenz in Abhängigkeit vom Schallquellen-Azi-

mut aufgetragen, unten die interaurale Pegeldifferenz. Die interaurale Zeitdifferenz weist ein weitgehend frequenzunabhängiges Verhalten auf, während die interaurale Pegeldifferenz stark frequenzabhängig ist. Sie ist in Abb. 1 für die Frequenzen 200 Hz und 4 kHz aufgetragen. Es ist zu beobachten, daß die interaurale Pegeldifferenz bei hohen Frequenzen für Winkel größer 60 Grad wieder abnimmt, was durch Beugungseffekte hervorgerufen wird. Hierdurch läßt sich auch die zur Seite hin zunehmende Lokalisations-Unschärfe erklären, die bei breitbandigen Signalen von etwa 1 Grad für die Vorne-Richtung auf 4 Grad zunimmt.

Elevation und Vorne-hinten-Unterscheidung

Die Mechanismen zur Schätzung der Elevation sind bis heute noch nicht ausreichend geklärt. Eng mit der Schätzung der Elevation ist auch die Vorne-Hinten-Unterscheidung verbunden, da ähnliche Verarbeitungsstrategien vermutet werden (Hartung et al. [1993]). Man geht davon aus, daß im wesentlichen monaurale spektrale Merkmale der Außenohr-Übertragungsfunktionen ausgenutzt werden. Die Grundlage hierfür bildeten Untersuchungen von Blauert zum Richtungshören in der Medianebene (Blauert [1974]). Ausgehend von den sogenannten »richtungsbestimmenden Bändern«, innerhalb derer die Lokalisation nur von der Signalfrequenz abhängt, lassen sich folgende Richtungen zuordnen:

- 500 Hz und 4 kHz: vorne
- 1 kHz: hinten
- 8 kHz: oben

Es zeigt sich, daß die Außenohr-Übertragungsfunktionen aus den entsprechenden Richtungen diese richtungsbestimmenden Bänder betonen. Für eine korrekte Lokalisation kann diese Information nur dann ausgenutzt werden kann, wenn das Spektrum des Signals bekannt ist. Für den Bereich der Elevationsschätzung und Vorne-hinten-Unterscheidung ist jedoch zu berücksichtigen, daß weitere Informationskanäle ausgenutzt werden können. Es sind vor allem Kopfbewegungen zu beachten, die dominante sogenannte »dynamische Cues« liefern. Des weiteren steht in der Regel zusätzlich der Sehsinn zur Verfügung, so daß der Trend besteht, Quellen, die nicht gesehen werden können tendenziell nach hinten zu legen.

Schallquellenentfernung

Die Effekte, die zur Schätzung der Schallquellenentfernung eingesetzt werden, können in drei

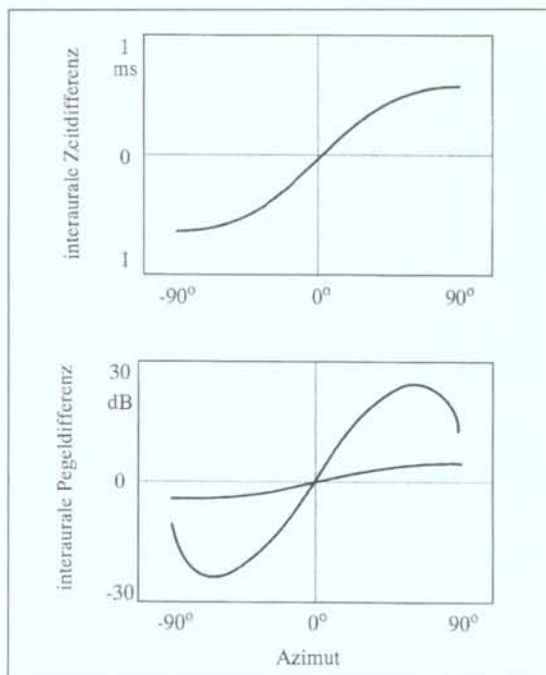


Abb. 1: Interaurale Zeit- und Pegeldifferenzen

Fig. 1: Interaural differences in time and intensity as a function of the sound source azimuth

Entfernungsbereiche unterteilt werden. Für Entfernungen kleiner als drei Meter ändern sich spektrale Eigenschaften der Außenohr-Übertragungsfunktionen, so daß bei Bekanntheit des Signals eine recht gute Entfernungsschätzung möglich ist. Bei Entfernungen zwischen drei und 15 Metern liefern die Außenohr-Übertragungsfunktionen keine charakteristischen Merkmale mehr. Hier hängt die geschätzte Entfernung im wesentlichen von der wahrgenommenen Lautstärke ab, da sich diese mit wachsender Entfernung verringert. Bei etwa 15 Metern Entfernung liegt der sogenannte »akustische Horizont«. Dies bedeutet, daß Hörereignisse ohne zusätzliche Information anderer Sinne nicht weiter als 15 Meter wahrgenommen werden.

Zusammenfassend kann man aussagen, daß das Gehör bezüglich der Schätzung des Azimuts, die sich auf die Auswertung interauraler Parameter stützt, im Vergleich zur Schätzung der Elevation und der Entfernung, die auf der Auswertung monauraler spektraler Parameter und auch der Bekanntheit des Signals beruht, leistungsfähiger erscheint und genauere Schätzungen ermöglicht.

Einohriges Hören

Die Lokalisation kann bezüglich des Azimuts anhand der interauralen Differenzen durchgeführt werden. Diese interauralen Differenzen stehen beim einohrigen Hören nicht zur Verfügung, so daß eine Ortung der Schallquelle nur anhand der spektralen Parameter der Außenohr-Übertragungsfunktion und durch Peilbewegungen des Kopfes erreicht werden kann. Da aber die interauralen Differenzen fehlen, wird sich das Hörereignis nicht am Ort der Schallquelle befinden. Es stellen sich in einem solchen Fall Hörereignisse ein, die sich im Kopf befinden können (sogenannte »Im-Kopf-Lokalisation«). Diese Hörereignisse werden sich in der Regel aufgrund der fehlenden räumlichen Information in der Nähe des Ohres oder direkt am Ohr befinden. Bei monauralem Hören ist eine Trennung konkurrierender Sprachsignale schlechter möglich als beim beidohrigen Hören, bei dem die zugehörigen Hörereignisse räumlich getrennt sind. Eine psychoakustische Quantifizierung dieses Unterschiedes wird im folgenden Kapitel dargestellt.

Binaurale Sprachverständlichkeits-Differenzen

Der Gewinn an Sprachverständlichkeit in stör-schallerfüllter Umgebung wird in der Psychoakustik anhand sogenannter Sprachverständlichkeits-Differenzen (ILD) gemessen. Abb. 2 zeigt den von *Platte* (1979) eingesetzten Versuchsaufbau. Bei allen Messungen wird eine Störquelle fest

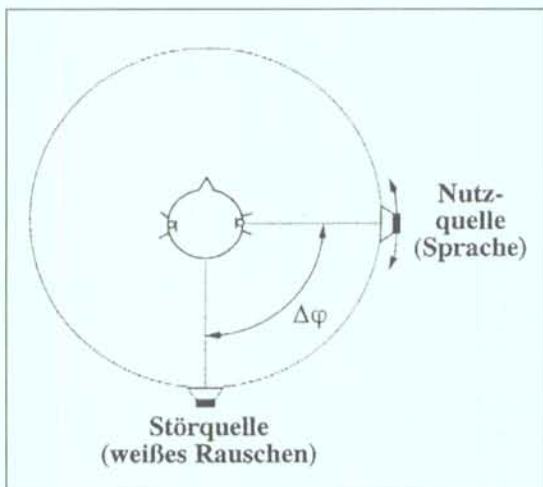


Abb. 2: Aufbau zur Messung der BILD

Fig. 2: Measurement of the BILD

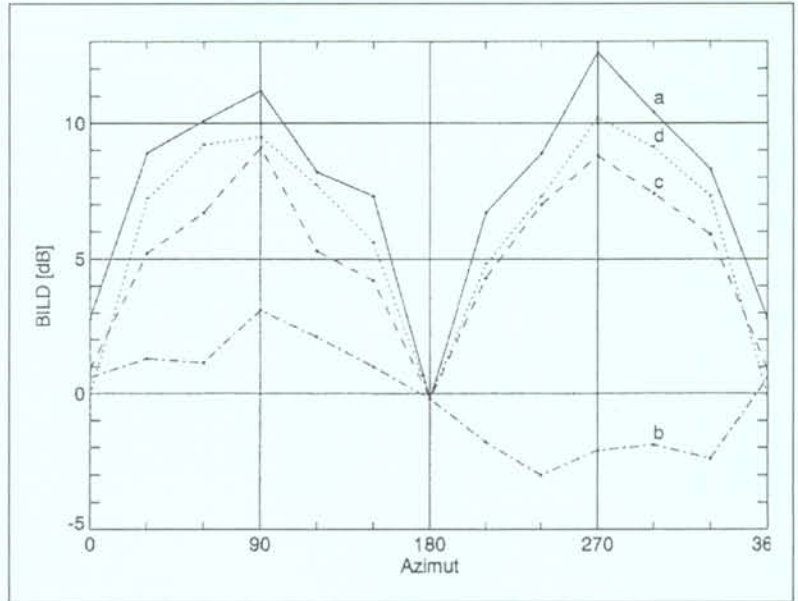
an einer Position hinter der Versuchsperson installiert, die in diesem Fall weißes Rauschen abstrahlt. Der Azimut der Nutzquelle wird nun in einer ersten Messung ebenfalls in die Position hinter der Versuchsperson gebracht, so daß keine räumliche Trennung zwischen Nutz- und Störquelle vorliegt. Der Pegel der Störquelle wird nun solange verändert, bis die gemessene Sprachverständlichkeit einen Wert von 50% annimmt. Diese Situation bildet die Referenzsituation für die folgenden Messungen, bei denen nun die Nutzquelle um den Kopf herumgeführt wird. An jeder Position wird der Pegel der Störquelle nun solange verändert, bis sich eine Sprachverständlichkeit von 50 Prozent einstellt. Der Pegelunterschied des Störsignals bezogen auf die Referenzsituation von hinten wird als ILD (Intelligibility Level Difference) bezeichnet und beschreibt den Nutzen, der aus der räumlichen Verteilung der Quellen gewonnen werden kann. Führt man nun diese Messungen zum einen monaural und zum anderen binaural aus, so kann man hieraus den Gewinn durch das beidohrige Hören, die BILD (Binaural ILD) ableiten. Sie ist als die Differenz der binaural und monaural gemessenen ILDs definiert. Abb. 3 zeigt die resultierenden monauralen (b) und binauralen (a/d/c) ILDs. Man erkennt, daß die BILD-Werte von bis zu 9 dB bei einem Azimut von 90 Grad annimmt.

Originalgetreue elektroakustische räumliche Übertragung

Will man einem Hörgeschädigten die Vorteile des beidohrigen Hörens durch eine entsprechende Hörgeräte-Versorgung ermöglichen, so muß man zunächst beachten, welchen Einfluß die dafür notwendige elektroakustische Übertragungsanlage in Form von Mikrofonen, Verstärker und Hörer auf die Reproduzierung räumlicher Eigenschaften nehmen kann. Hierzu sei zunächst der Optimalfall einer originalgetreuen akustischen Übertragung, die Kunstkopf-Technologie, betrachtet. Wie man in Abb. 4 erkennt, liegt nun eine Aufnahmesituation vor, in der das Schallfeld mit zwei in den Ohrnachbildungen eines Kunstkopfes angebrachten Mikrofonen aufgezeichnet wird. Das Ziel dieser Technologie besteht darin, daß sich in der Wiedergabesituation bei einem Zuhörer, dem die Signale über Kopfhörer dargeboten werden, dieselben räumlichen Hörereignisse einstellen, als wenn er sich selbst an der Aufnahmeposition befinden würde. Stimmen die monauralen und binauralen

Abb 3: BILD nach Platte (1979); a/d/c: binaural, b: monaural

Fig 3: BILD (data taken from Platte [1979]); a/d/c: binaural, b: monaural



Außenohr-Eigenschaften des bei der Aufnahme eingesetzten Kunstkopfes mit dem des Zuhörers recht gut überein, so werden die Hörereignis-Positionen auch recht gut mit den Schallquellen-Positionen übereinstimmen. Zu beachten ist jedoch, daß sowohl Information über Kopfbewegungen als auch über den Sehsinn nicht vorliegen, so daß die Leistungsfähigkeit zur Lokalisation bezüglich der Elevation und der Entfernungsabschätzung eingeschränkt sein kann. Ein generelles Problem der Kunstkopf-Technologie besteht aber darin, daß die Außenohr-Übertragungsfunktionen, wie bereits zuvor besprochen, interindividuell unterschiedlich ausgeprägt sind. Das hat zur Folge, daß die Außenohr-Parameter des Zuhörers durchaus von denen des Kunstkopfes abweichen können. Aus diesem Grunde treten recht häufig zwei unterschiedliche Probleme auf:

- Vorne-hinten-Vertauschungen;
- Im-Kopf-Lokalisationen; aufgrund der elektroakustischen Übertragung kann ein Teil der räumlichen Information verloren gehen. Gut reproduziert wird jedoch in der Regel die Wahrnehmung der Schallquellen-Azimute, so daß eine räumliche Trennung der Hörereignisse aufrechterhalten bleibt.

Originalgetreue räumliche Übertragung mit Hörgeräten

Eine originalgetreue räumliche Übertragung ist mit einer einohrigen Hörgeräteversorgung nicht

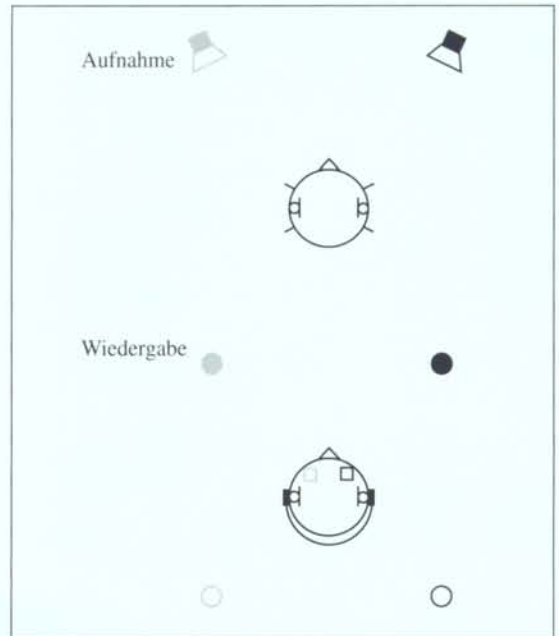


Abb. 4: Kunstkopf-Technologie. Die geschlossenen Kreise ● symbolisieren die Positionen der Hörereignisse bei originalgetreuer Übertragung, offene Kreise ○ Vorne-Hinten-Vertauschungen und offene Quadrate □ Im-Kopf-Lokalisation

Fig. 4: Dummy-head-technology. Filled circles ● symbolize correct localization, open circles ○ front-back-reversals, squares □ localisation inside of the head

möglich. Jedoch selbst bei der heute üblichen Kunstkopf-Technologie können Probleme bezüglich einer originalgetreuen räumlichen Abbildung auftreten. Ein Teil dieser Probleme kann jedoch vermieden werden, wenn bei einer beidohrigen Hörgeräteversorgung die Mikrophone in den Gehörgängen des Hörgeschädigten angebracht werden, da in diesem Fall seine eigenen, individuellen Außenohr-Eigenschaften berücksichtigt werden (Abb. 5). Es ist jedoch zu beachten, daß eine beidohrige Versorgung durch zwei unabhängig voneinander arbeitende Hörgeräte realisiert wird. Dies hat zur Folge, daß die getrennte Verarbeitung dieser Hörgeräte gerade auch die interauralen Differenzen beeinflussen und verändern kann, insbesondere wenn ein asymmetrischer Hörverlust vorliegt und Geräte mit einer AGC eingesetzt werden. Hierdurch kann nun der Fall eintreten, daß die interaurale Pegeldifferenz verändert und sogar pegelabhängig werden kann. Das Gehör kann zwar über einen längeren Zeitraum die veränderten Kombinationen der interauralen Parameter neu erlernen, sich jedoch nicht auf pegelabhängige Kombinationen einstellen. Die Auswirkungen können nun je nach Grad der Abweichung von den

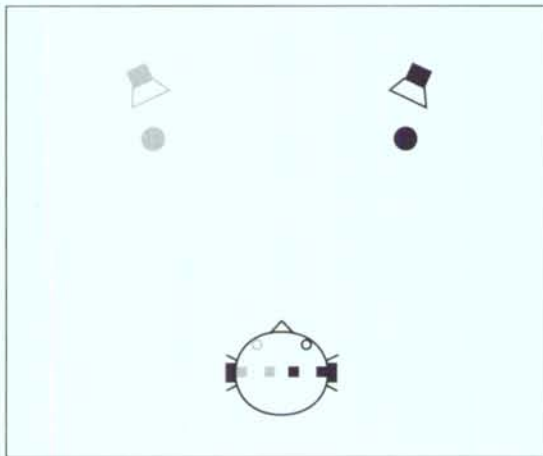


Abb. 5: Beidohrige Hörgeräteversorgung. Die geschlossenen Kreise ● markieren die Positionen der Hörereignisse bei korrektem Richtungshören, offene Kreise ○ Im-Kopf-Lokalisation und Quadrate □ den Zerfall in mehrere Hörereignisse

Fig. 5: Binaural hearing-aid fitting. Filled circles ● symbolize correct localization, open circles ○ localisation inside of the head; squares □ splitting up into multiple auditory events



Abb.6: Vor- und Nachteile der ein- und beidohrigen Hörgeräte-Versorgung

Fig. 6: Advantages and disadvantages of monaural and binaural hearing-aid fittings

gewohnten Kombinationen unterschiedliche Ausmaße annehmen:

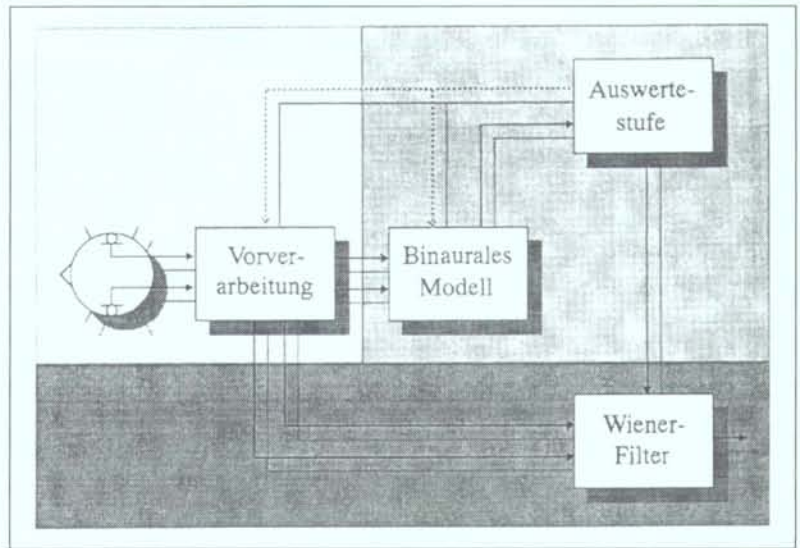
- Es können wiederum Vorne-hinten-Vertauschungen auftreten, die jedoch durch die Unterstützung des Sehsinnes reduziert werden können;
- es kann wiederum eine Im-Kopf-Lokalisation auftreten;
- bei größeren Abweichungen kann ein Zerfall der Hörereignisse auftreten. In diesem Fall erzeugt eine Schallquelle mehrere Hörereignisse, so daß die räumliche Information eher zu Irritationen führen kann und sich gegenüber einer monauralen Darbietung als nachteilig erweisen kann.

Konsequenzen für die Hörgeräteversorgung

Aus dem bisher Dargestellten ergeben sich Vor- und Nachteile der ein- und beidohrigen Hörgeräteversorgung (Abb. 6). Die Nachteile der heutigen ein- und auch der beidohrigen Hörgeräteversorgung können durch ein sogenanntes »binaurales Hörgerät« erheblich reduziert werden. Ein solches Hörgerät unterscheidet sich von der beidohrigen Hörgeräte-Versorgung dadurch, daß die beiden Hörgeräte nicht mehr unabhängig voneinander, sondern zentral gesteuert arbeiten. Der hierzu nötige zentrale Prozessor kann, falls es die Baugröße erlaubt, durchaus

Abb. 7: Aufbau des Systems zur Modellierung des Cocktail-Party-Effektes

Fig. 7: Structure of the system to model the Cocktail-Party-Effect



in einem der beiden Hörgeräte untergebracht werden. In Hinblick auf die zukünftige Weiterentwicklung von Hörgeräten bietet diese Strategie jedoch noch einen weiteren Vorteil. Die bisher behandelten Strategien unterstützen die Kommunikationsfähigkeiten von Hörgeschädigten nur dann, wenn sie über ein ausreichendes binaurales Resthörvermögen verfügen, also aus einer räumlichen Verteilung von Nutz- und Störschall noch Vorteile ziehen können. Diese binauralen Fähigkeiten sind bei Schwerhörigen jedoch häufig stark reduziert, so daß auch eine optimale beidohrige Versorgung keine ausreichende Sprachverständlichkeit im Störschall ermöglicht. In diesen Fällen müßte ein spezielles Hörgerät dem Hörgeschädigten die Fähigkeiten zur binauralen Störgeräusch-Unterdrückung eröffnen. Die Strategie des binauralen Hörgerätes sieht vor, daß in dem zentralen Prozessor Algorithmen integriert werden, die diese Aufgaben übernehmen. An der Ruhr-Universität in Bochum wurde ein »Cocktail-Party-Prozessor« als Forschungsmodell entwickelt, der im folgenden kurz vorgestellt wird.

Ein Cocktail-Party-Prozessor

Der von uns entwickelte Cocktail-Party-Prozessor verfolgt das Ziel, die binaurale Verarbeitung des menschlichen Gehörs weitestgehend nachzuvollziehen. Abb. 7 stellt das Konzept des Systems dar, für dessen genauere Beschreibung auf die Literatur verwiesen sei (Bodden [1992]). Das System enthält Bausteine, die ausgehend von einer Aufnahme der Schallsignale mit ei-

nem Kunstkopf oder zwei Ohreinsteck-Mikrofonen zunächst die Verarbeitung peripherer Stufen von Mittel- und Innenohr in Form einer konzentrierten Vorverarbeitung simulieren. Diese Vorverarbeitung besteht im wesentlichen aus einer 24kanaligen Filterbank, die eine Aufteilung der Schallsignale in die Frequenzgruppen entsprechend Zwicker und Feldtkeller (1967) durchführt. Nachgeschaltet folgt ein binauraler Prozessor, dessen Aufgabe darin besteht, die durch die Außenohren in die Signale eingepprägten interauralen Differenzen zu analysieren. Hier sollen nur kurz einige Merkmale dieses Prozessors vorgestellt werden, eine detaillierte Darstellung findet man bei Gaik (1990). Der binaurale Prozessor basiert auf einer interauralen Kreuzkorrelation der von der Vorverarbeitung gelieferten Signale. Eine wesentliche Erweiterung bildet die Integration von Inhibitionsmechanismen, die in der neuronalen Verarbeitung häufig beobachtet werden. Aufgrund dieser Verarbeitung kann der Prozessor sowohl interaurale Zeit- als auch Pegeldifferenzen analysieren. In einer speziell überwachten Lernphase adaptiert sich der Prozessor an die individuellen Ausprägungen der bei der Aufnahme eingesetzten Außenohr-Übertragungsfunktionen. Die Integration monauraler Prozessoren erlaubt schließlich auch die Berücksichtigung rein monauraler Hörereignisse. Die Ausgangsmuster des binauralen Prozessors stellen Simulationen neuronaler Erregungen dar. Abb. 8 zeigt beispielhaft ein solches neuronales Erregungsmuster für einen festen Zeitpunkt. Die Kreuzkorrelations-Achse

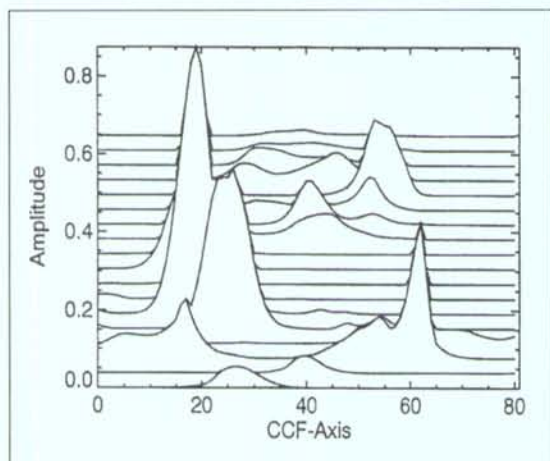
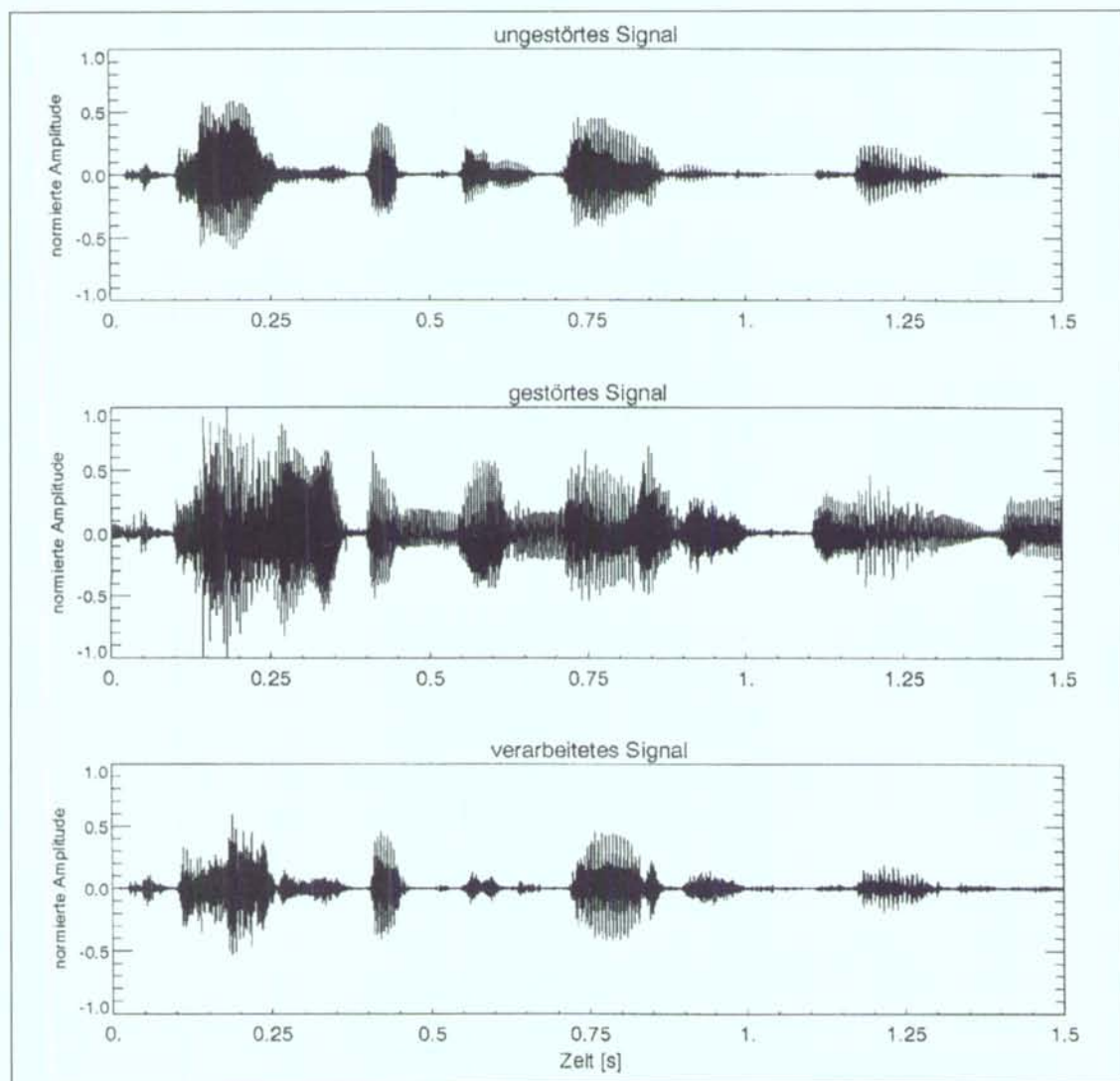


Abb. 8: Neuronales Erregungsmuster für eine Zwei-Sprecher-Situation. Jede Kurve stellt die neuronale Erregung in einer Frequenzgruppe dar. Die Korrelationsachse beschreibt die seitliche Auslenkung aus der Vorne-Richtung (Achsenabschnitt 40). Werte kleiner 40 stehen für eine Einfallsrichtung von links, Werte größer 40 für eine Einfallsrichtung von rechts

Fig. 8: Neural excitation pattern for one moment in time, two speakers. Each curve represents the neural excitation in one critical band. The correlation axis describes the direction of sound incidence. A value of less than 40 represents signals from the left, values above 40 represent signals from the right



beschreibt die seitliche Auslenkung, wobei der Achsenabschnitt 40 einer Schall-Einfallrichtung von vorne entspricht. Jede Kurve stellt die neuronale Erregung in einer Frequenzgruppe dar, wobei die Frequenz nach hinten ansteigt. Das Erregungsmuster repräsentiert eine Situation, in der zwei Sprecher gleichzeitig aktiv waren: Ein Sprecher befand sich links (Achsenabschnitt 20), der andere Sprecher rechts (Achsenabschnitt 60). Der Vorteil der Analyse der Signale mit dem binauralen Prozessor wird anhand dieses Musters deutlich: Die räumliche Verteilung der Schallquellen ist dem Muster zu entnehmen. Diese Information wird nun von der übergeordneten Auswertestufe analysiert. Sie kann mit Hilfe der neuronalen Erregungsmuster den Signal-Störabstand in jeder Frequenzgruppe zeitabhängig schätzen. Mit dieser Information wird dann die Übertragungsfunktion eines Wiener-Filter berechnet, mit dessen Hilfe ein Schätzer für das Nutzsignal aus den gestörten Eingangssignalen gewonnen werden kann. Dieses Wiener-Filter stellt letztendlich eine Art 24kanaligen Equalizer dar, in welchem Frequenzbänder mit viel Störenergie gegenüber den anderen Frequenzbändern abgeschwächt werden. Abb. 9 zeigt die Wirkungsweise des Cocktail-Party-Prozessors anhand eines Beispiels. Die obere Kurve stellt das als Nutzsignal definierte Sprachsignal dar, im Bild darunter ist die Überlagerung mit einem konkurrierenden Sprachsignal aufgetragen. Das verarbeitete Signal ist im unteren Bild dargestellt. Man erkennt, daß das Störsignal nahezu vollständig unterdrückt wird, aber das Nutzsignal auch verzerrt wird. Die Wirksamkeit der Störschall-Unterdrückung somit auch damit verbundenen Verzerrungen des Nutzsignals können über entsprechende Parameter eingestellt werden. Eine Verbesserung des Signal-Störabstandes, die mit dem Verfahren erzielt werden kann, muß jedoch nicht unbedingt eine Verbesserung der Sprachverständlichkeit nach sich ziehen. Aus diesem Grunde wurden Sprachverständlichkeits-Tests mit Hörgeschädigten durchgeführt.

Abb. 9: Nutzsignal (oben), mit einem Störsprecher überlagertes Signal (mitte) und mit dem Cocktail-Party-Prozessor verarbeitetes Signal (unten)

Fig. 9: Desired signal (upper curve), interfered signal (middle) and processed signal (lower curve)

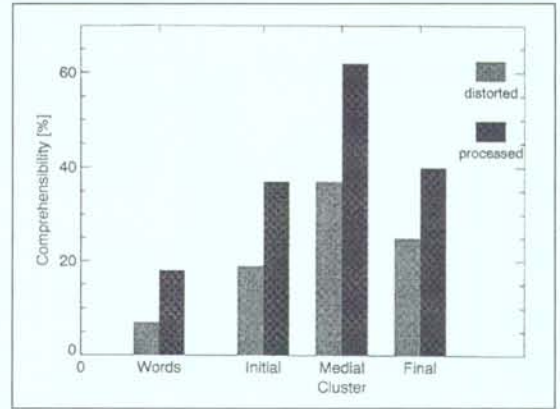


Abb. 10: Ergebnisse des Worttestes. Helle Balken zeigen die Ergebnisse für die gestörten Signale, dunkle Balken nach der Verarbeitung. Links sind die Ergebnisse für die gesamten Wörter, rechts für die einzelnen Cluster aufgetragen

Fig. 10: Results of the word test. Light bars represent results for the interfered signals, dark bars for the processed signals. Left: complete words; right: different clusters

Da insbesondere die Veränderung der Sprachverständlichkeit für das Störsignal Sprache untersucht werden sollte, konnte kein Standardtest eingesetzt werden, da diese nur sprachähnliche Störsignale einsetzen. Diese sprachähnlichen Signale werden entweder aus sprachsimuliertem Rauschen oder aber durch häufiges Überlagern von Sprachsignalen gebildet (Kollmeier, 1990), wodurch in jedem Fall das Störsignal keine Bedeutung mehr trägt. Aus diesem Grunde wurden zwei Tests entwickelt, bei denen Nutz- und Störsignal aus Sprache bestehen. Der erste Test wurde als Worttest konzipiert, für den 100 sinnleere Einsilber, bestehend aus Konsonant-Vokal-Konsonant-Clustern (CVC), generiert wurden. Die von Jekosch et al. (1991) vorgeschlagene Methode zur Messung der Sprachverständlichkeit mit sinnleerem Material weist den Vorteil auf, daß die Ergebnisse des Tests das tatsächlich erkannte Sprachmaterial darstellen. Bei Tests mit sinnbehaftetem Material fließt zusätzlich noch das Weltwissen der Versuchspersonen ein, die häufig nicht erkannte Wort- oder Satzteile aus dem Kontext »erraten« können. Die aufgrund einer Statistik über die Auftretenshäufigkeit der unterschiedlichen Cluster in der deutschen Sprache erzeugten Einsilber wurden von zwei männlichen Sprechern in einem reflexionsarmen Raum vorgelesen. Anschlies-

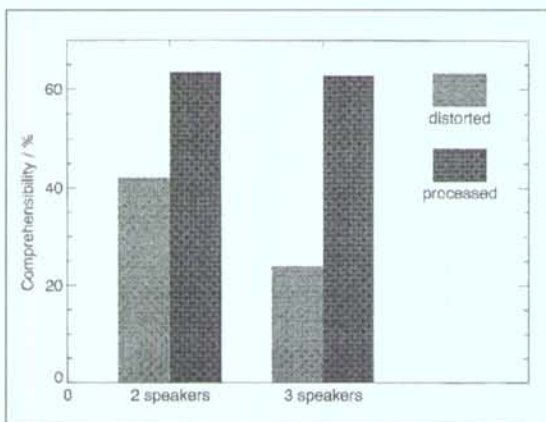


Abb. 11: Ergebnisse des Satztests. Helle Balken zeigen die Ergebnisse für die gestörten Signale, dunkle Balken nach der Verarbeitung. Links sind die Ergebnisse für eine Zwei-Sprecher-Situation, rechts für eine Drei-Sprecher-Situation aufgetragen.

Fig. 11: Results of the sentence test. Light bars represent results for the interfered signals, dark bars for the processed signals. Left: two-speaker situation; right: three-speaker situation

send wurden den Signalen durch Faltung mit Außenohr-Impulsantworten die Einfallrichtungen aus ± 30 Grad zugeordnet. Die synchrone Addition der Signale bei einem Signal-Störabstand von 0 dB führte zu den gestörten, binauralen Signalen, die eine Testreihe bildeten.

Die verarbeiteten Signale wurden dichotisch in der anderen Meßreihe dargeboten. Jeder der Sprecher fungierte 50 mal als Nutz- und 50 mal als Störsprecher, wobei jedesmal der Sprecher von rechts als Nutzsprecher definiert war. An dem Test nahmen 5 Hörgeschädigte teil, denen die Signale per Kopfhörer bei Benutzung ihrer eigenen Hörgeräte dargeboten wurden. Alle Versuchspersonen waren beidohrig versorgt. Die Ergebnisse sind in Abb. 10 dargestellt, wobei die Erkennbarkeiten der kompletten Wörter (links) sowie der einzelnen Cluster (rechts) aufgetragen sind. Die hellen Balken markieren die Erkennbarkeit vor, die dunklen Balken nach der Verarbeitung. Es ist zu erkennen, daß für alle Cluster eine signifikante Verbesserung der Erkennbarkeit durch die Vorverarbeitung erzielt wird. Hervorzuheben ist, daß auch die Erkennbarkeit der Konsonant-Cluster durch die Vorverarbeitung erhöht wird. Die relativ geringe Gesamterkennbarkeit ist auf die Versuchsbedingungen zurückzuführen: Da beide Sprecher

über eine ähnliche Sprechcharakteristik und Grundfrequenz verfügen, fällt eine Trennung relativ schwer. Es ist weiterhin anzumerken, daß den Hörgeschädigten keinerlei Gewöhnungszeit an die verarbeiteten Signale gewährt wurde. Durch die verfahrensspezifische Beeinflussung der Klangqualität der Signale kann bei einer entsprechenden Gewöhnung mit noch besseren Ergebnissen gerechnet werden.

Der zweite Test wurde als Satztest konzipiert, wobei die Sätze nun aus sinnbehafteten Worten grammatikalisch richtig zusammengesetzt waren, aber keine semantische Bedeutung enthielten (Beispiel: »Das Land schneit in das hohle Geld«). Jeweils zehn dieser Sätze wurden von drei Sprechern im reflexionsarmen Raum vorgelesen und anschließend binaural gemischt, wobei eine Nutzsprecherin vorne (0 Grad), eine Störsprecherin seitlich links (-45 Grad) und ein Störsprecher seitlich rechts (+30 Grad) angeordnet wurden. Alternativ wurde auch eine Zwei-Sprecher-Situation mit einer Nutzsprecherin von rechts (+30 Grad) und einem Störsprecher von links (-30 Grad) ausgewählt. Die Ergebnisse von Tests für vier Hörgeschädigte sind Abb. 11 zu entnehmen. Es ist in beiden Situationen eine deutliche Verbesserung der Erkennbarkeit zu beobachten.

Schlußfolgerung

Die beschriebenen Untersuchungen haben gezeigt, daß mit dem von uns entwickelten Verfahren eine Verbesserung der Verständlichkeit in Cocktail-Party-Situationen erzielt werden kann. Probleme bereitet heute jedoch noch die nötige Miniaturisierung der Algorithmen, so daß ein Einsatz des oben beschriebenen Cocktail-Party-Prozessors in Hörgeräten derzeit noch nicht möglich ist. Das vorgestellte Konzept des binauralen Hörgerätes ermöglicht jedoch ein schrittweises Vorgehen, indem zunächst einfachere Algorithmen implementiert werden, die zum Beispiel als ersten Schritt die Verstärkungs-Kennlinien beider Hörgeräte in der Art zentral steuern, daß die interauralen Pegeldifferenzen durch die Hörgeräteversorgung nicht verändert werden. Hierdurch könnte eine optimale Unterstützung der binauralen Resthörbarkeit erzielt werden. Aufgrund der rasch fortschreitenden Entwicklung immer leistungsfähigerer und kleinerer Computerchips kann dann erhofft werden, innerhalb eines Zeitraums von etwa zehn Jahren auch komplexere Algorithmen in den zentralen Prozessor integrieren zu können.

Literatur/References

- Blauert, J.:* »Räumliches Hören«. Hirzel Verlag, Stuttgart, 1973.
- Blauert, J.:* »Spatial Hearing«. MIT-Press, Cambridge, Mass, 1984.
- Bodden, M.:* »Binaurale Signalverarbeitung: Modellierung der Richtungserkennung und des Cocktail-Party-Effektes«. Fortschr.-Ber.; VDI-Reihe 17, Nr. 85, VDI-Verlag, Düsseldorf, 1992.
- Gaik, W.:* »Untersuchungen zur binauralen Verarbeitung kopfbezogener Signale«. Fortschr.-Ber., VDI-Reihe 17, Nr. 63, VDI-Verlag, Düsseldorf, 1990.
- Hartung, K., Miyoshi, M., Bodden, M., Blauert, J.:* »Merkmale zur Vorne-hinten-Unterscheidung unterhalb von 2 kHz«. Fortschritte der Akustik; DAGA '93; DPG-GmbH, Bad Honnef, 1993 (im Druck).
- Jekosch, U.:* »A weighted intelligibility measure for speech assessment«. Proceedings of the 1990 Int. Conf. on Spoken Lang. Proc., Kobe, Japan, Nov. 1990, 973-976.
- Jekosch, U., Belhoula, K., Hegehofer, Th., Mariniak, A.:* »Verwendung von sinnleeren Wörtern zur Sprachgütebeurteilung«. Fortschritte der Akustik; DAGA '91; DPG-GmbH, Bad Honnef, 1991, 901-904.
- Kollmeier, B.:* »Meßmethodik, Modellierung und Verbesserung der Verständlichkeit von Sprache«. Habilitationsschrift, Universität Göttingen, 1990.
- Lim, J. S.:* »Speech Enhancement«. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1983.
- Platte:* »Zur Bedeutung der Außenohr-Übertragungseigenschaften für den Nachrichtenempfänger menschliches Gehör«. Dissertation, Technische Hochschule Aachen, 1979.
- Zwicker, E., Feldtkeller:* »Das Ohr als Nachrichtenempfänger«. Hirzel Verlag, Stuttgart, 1967.



Johann Wolfgang Goethe-Universität Frankfurt am Main

An der Johann Wolfgang Goethe-Universität Frankfurt am Main ist unter den Einstellungs Voraussetzungen des § 39 a des Gesetzes über die Universitäten des Landes Hessen (HUG) im Fachbereich Humanmedizin im Zentrum der Hals-Nasen-Ohrenheilkunde in der Abteilung für Allgemeine Hals-Nasen-Ohrenheilkunde die

Professur (C 3) für Medizinische Akustik

zu besetzen. Mit der Professur ist die Vertretung dieses Faches in Forschung, Lehre und medizinischer Diagnostik verbunden. Die Bewerberinnen und Bewerber müssen entweder promovierte Naturwissenschaftler, Ingenieurwissenschaftler oder Mediziner sein. Nach § 39 a HUG wird für die Besetzung von Professuren in der Regel die Habilitation vorausgesetzt. In Ausnahmefällen kann diese durch den Nachweis gleichwertiger wissenschaftlicher Leistungen ersetzt werden. Langjährige Erfahrung mit sämtlichen medizinisch-akustischen Untersuchungsverfahren einschließlich otoakustischer Emissionen ist erwünscht. Von den Bewerberinnen und Bewerbern wird im Falle ihrer Berufung erwartet, daß sie ihr Fach in Kooperation mit den Professoren des Zentrums vertreten.

Schwerbehinderte werden im Rahmen der geltenden gesetzlichen Bestimmungen bevorzugt behandelt. Die Johann Wolfgang Goethe-Universität Frankfurt am Main strebt die Erhöhung des Anteils von Frauen am wissenschaftlichen Personal an und fordert daher Frauen nachdrücklich auf, sich zu bewerben. Ein Merkblatt zur besonderen Gestaltung des Schriftenverzeichnisses, wie es dem hessischen Ministerium für Wissenschaft und Kunst vorzulegen ist, versendet der Dekan des Fachbereiches Humanmedizin, Theodor-Stern-Kai 7, auf Anforderung. Einen Textauszug des § 39 a HUG hält der Präsident zur Versendung bereit. Bewerbungen sind innerhalb von sechs Wochen nach Erscheinen dieser Anzeige mit den üblichen Unterlagen an den Präsidenten der Johann Wolfgang Goethe-Universität Frankfurt am Main, Postfach 11 19 32, 6000 Frankfurt am Main 11, zu richten.

Postleitzahlen ab 1. Juli 1993: Universität, Postfach: 60054, Universität, Senckenberganlage 31 (Paketpost): 60325