

ZUR THEORIE DER OPTIMALEN WIEDERGABE VON STEREOFONEN SIGNALEN ÜBER LAUTSPRECHER UND KOPFHÖRER

VON GÜNTHER THEILE¹

Manuskript eingegangen am 9. Juli 1981

Kopfbezogene Stereophonie

Zusammenfassung

Eng verknüpft mit der Weiterentwicklung der kopfbezogenen Stereophonie sind die theoretischen Vorstellungen über die Funktion des Gehörs beim räumlichen Hören erweitert worden. Sie lassen sich beschreiben mit Hilfe eines neuen Lokalisationsmodells, dem „Assoziationsmodell“. Es basiert auf der Hypothese, daß Assoziationsvorgänge ein Grundprinzip der sensorischen Reizverarbeitung darstellen. Danach resultiert die auditive räumliche Wahrnehmung grundsätzlich aus zwei unterschiedlichen Verarbeitungsprozessen. Jeder der beiden Verarbeitungsprozesse geschieht auf dem Wege einer assoziativ gesteuerten Musterselektion: Ein aktueller Reiz, welcher von einer Schallquelle herrührt, löst in der ersten Stufe eine Ortsassoziation und in der zweiten eine Gestaltassoziation aus. Beide Stufen bestimmen gemeinsam die Hörereigniseigenschaften.

Die Abhängigkeit der Ohrsignale vom Schallquellenort wird als Form der Codierung räumlicher Information aufgefaßt, deren Kenntnis eine Decodierung der räumlichen Information ermöglicht. Die Decodierung der räumlichen Information leistet die Ortsassoziationsstufe des Modells. Die Wirkung kann mit einem steuerbaren Filter beschrieben werden, dessen Übertragungsfunktion sich infolge der assoziativen Mustererkennung invers verhält zur ortsabhängigen Übertragungsfunktion des Außenohres.

Es zeigt sich, daß dieses Funktionsschema eine einheitliche Erklärung der wichtigen Phänomene des räumlichen Hörens zuläßt. Es ermöglicht eine einfache systemtheoretische Darstellung der gesamten „Übertragungsstrecke“ Schallereignis — Ohrsignal — Hörereignis. Insbesondere für die unterschiedlichen elektroakustischen Übertragungsverfahren läßt sich damit die Auswertung der jeweils wirksamen Ohrsignale leicht übersehen und abschätzen.

Für die optimale Wiedergabe raum- oder kopfbezogener Signale über Lautsprecher oder Kopfhörer ergeben sich verschiedene Konsequenzen. Speziell die Frage nach der Kompatibilität raum/kopfbezogener Signale hat zwei neue Erkenntnisse gebracht: Es läßt sich zeigen, daß

- die Freifeldentzerrung des Kunstkopfes eine im Prinzip falsche Entzerrung darstellt in Hinblick auf Lautsprecherwiedergabe;
- die (standardisierte) Freifeldentzerrung der Kopfhörer eine im Prinzip falsche Entzerrung darstellt in Hinblick auf Wiedergabe der gebräuchlichen intensitätsstereofonen Signale.

Es wird deshalb aus systemtheoretischen Gründen, die sich aus der Wirkungsweise unseres Gehörs ableiten, als „Universalschnittstelle“ zwischen Aufnahme- und Wiedergabeseite eine Entzerrung vorgeschlagen, die sich grundsätzlich auf das Diffusfeld bezieht.

Summary On the theory of the optimum reproduction of stereophonic signals by way of loudspeakers and headsets

In close relation to the continuing development of head-related stereophony, the theoretical conceptions regarding the function of the ear in the case of spatial hearing have been extended. They may be described by means of a new localisation model, the „association model“. This is based on the hypothesis that the association processes constitute a basic principle of the sensorial stimulation processing. According to this, the spatial auditive perception results basically from two different processes of assimilation. Each of the two assimilation processes occurs by way of an associatively orientated selection of samples: an actual stimulus, derived from a sound source, gives rise in the first stage to a localisation association and in the second stage to a form association. Both stages together determine the characteristics of the hearing event.

The dependence on the location of the sound source of the aural signal is adopted as the form of the coding of the spatial information, the knowledge of which makes possible the decoding of the spatial information. The decoding of the spatial information is dealt with in the localisation association stage of the model. The effect can be represented by means of an adjustable filter whose transfer function, because of the associative sample recognition, is the inverse of the location-dependent transfer function of the external ear.

It is shown that this function scheme makes possible a general explanation of the important phenomena of spatial hearing. It makes possible also a simple system-theoretical representation of the entire „transmission chain“: sound event — ear signal — hearing event. For the different electro-acoustical transmission methods, in particular, the evaluation of the effective ear signal in question may thus be easily observed and interpreted.

For the optimal reproduction of space- or head-related signals by way of loudspeakers or headsets, there are several consequences. The question of the compatibility of space/headrelated signals, in particular, has given rise to two new concepts; it is possible to show that:

- the free-field correction of the artificial head constitutes a correction that is, in principle, wrong as regards reproduction by loudspeakers;
- the (standardised) free-field correction of the headset constitutes a correction that is wrong in principle, as regards the reproduction of the usual intensity-stereophonic signals.

Therefore, for system-theoretical reasons that are based upon the functioning of our sense of hearing, a correction is proposed as a „universal interface“ between the recording and reproduction ends, which relates basically to the diffused field.

Sommaire Considérations théoriques sur la reproduction optimale de signaux stéréophoniques au moyen de haut-parleurs et d'écouteurs

Les notions théoriques relatives au fonctionnement de l'oreille dans le cas de l'écoute directionnelle ont été étendues, en relation étroite avec les développements en cours dans le domaine de la stéréophonie liée à la tête. On peut décrire ces propriétés par un nouveau modèle dit d'association, basé sur l'hypothèse que les associations constituent une des données fondamentales du traitement des stimulations sensorielles. Dans ces conditions, la perception auditive de l'espace résulte fondamentalement de deux processus d'assimilation différents. Chacun découle d'une sélection d'échantillons orientée associativement. Un stimulus produit par une source sonore donne tout d'abord lieu à une association de localisation, puis dans une seconde étape, à une association de forme et c'est la synthèse de ces deux phénomènes qui détermine les caractéristiques de la perception auditive.

La relation entre l'emplacement de la source sonore et le signal auditif constitue un codage de l'information spatiale dont le décodage permet la localisation. Cette opération correspond à l'étape d'association de

localisation. Son effet peut être représenté par un filtre variable dont la fonction de transfert est l'inverse de celle de l'oreille externe, du point de vue de la localisation, du fait de la reconnaissance associative des échantillons.

On montre que ce schéma de relation fonctionnelle permet de donner une explication générale à l'important phénomène de l'écoute directionnelle et aussi de donner une représentation théorique simple de l'ensemble de la "chaîne de transmission": phénomène sonore, signal auditif, phénomène auditif. La méthode permet notamment d'observer et d'interpréter facilement l'évaluation du signal auditif effectif pour différents systèmes de transmission électro-acoustiques. Il en résulte plusieurs conséquences pour la reproduction optimale par haut-parleurs ou par écouteurs de signaux liés à l'espace ou à la tête. La question de la compatibilité de ces signaux a notamment permis de dégager deux nouvelles notions et on peut démontrer que:

- la correction de champ libre de la tête artificielle est, par principe, erronée en ce qui concerne la reproduction au moyen de haut-parleurs;
- la correction d'espace libre normalisée appliquée aux écouteurs est erronée par principe, du point de vue de la reproduction des signaux en stéréophonie d'intensité normale.

On propose donc une correction appuyée sur des raisons théoriques basées sur le fonctionnement de l'ouïe humaine, ce qui permettrait de définir une "interface universelle" entre la prise de son et la reproduction. Cette correction porte principalement sur le champ diffus.

1. Einleitung

Sowohl die Erfahrungen, die man während der letzten 8 Jahre in den Rundfunkanstalten mit der Kunstkopfstereofonie gemacht hat [1, 2, 3], als auch die Erkenntnisse, die man in jüngster Zeit bei der Weiterentwicklung der Kunstkopfstereofonie gewonnen hat [4, 5, 6], zeigen, daß dieses Übertragungsverfahren eine echte Alternative darstellt zur bisher gebräuchlichen Intensitätsstereofonie.

Im Gegensatz zur intensitätsstereofonen Technik ist es das Ziel der Kunstkopftechnik, dem Zuhörer die Hörillusion zu verschaffen, daß er sich am Aufnahmeort befände und dort am akustischen Geschehen unmittelbar teilhätte. Darüber hinaus aber müssen die Kunstkopfsignale aus Kompatibilitätsgründen bei Lautsprecherwiedergabe gleichwertige Abbildungseigenschaften aufweisen wie entsprechende intensitätsstereofone Signale.

Es hat sich gezeigt, daß die Anforderungen bei der technischen Entwicklung einer rundfunktauglichen kopfbezogenen Stereofonie besonders vielfältige Probleme aufwerfen, Probleme, die überwiegend mit den wichtigen Phänomenen des räumlichen Hörens direkt zusammenhängen: Dazu gehören Phänomene der Lokalisation einer Einzelschallquelle, der Lokalisation im Schallfeld, welches von mehreren Schallquellen hervorgerufen wird (überlagertes Schallfeld), sowie Phänomene der Kopfhörerbeschallung (Lateralisation).

Ferner hat sich gezeigt, daß die bisherigen theoretischen Vorstellungen über die Funktion unseres Gehörs beim räumlichen Hören nicht ausgereicht haben, um die komplexen Zusammenhänge und Gesetzmäßigkeiten für die Konzeption eines optimalen stereofonen Übertragungsverfahrens zu nutzen. Zwar sind die vielen Einzelphänomene – oft sehr ausführlich – experimentell untersucht worden, und für den überwiegenden Teil der bekannten Gesetzmäßigkeiten gibt es Modellvorstellungen, die das Gehör in den verschiedenen Funktionen zu beschreiben suchen. Doch die Mehrheit dieser Modelle sind nachrichtentechnische Modelle, welche nur innerhalb ihrer eng begrenzten Gültigkeitsbereiche Aussagen machen über die Beziehungen zwischen bestimmten Ohrsignalmerkmalen und Hörereigniseigenschaften. Es ist wegen der eng begrenzten Gültigkeitsbereiche

aber zweifelhaft, ob sie damit überhaupt Informationen liefern, die das Verständnis über die Funktion des gesamten Auswertungsprozesses beim Hören stereofoner Signale verbessern. Welchen Sinn hat beispielsweise ein Modell nur für das Richtungshören bei raumbezogener Stereofonie, gültig nur im freien Schallfeld, nur für die Horizontalebene, nicht für breitbandige Lautsprechersignale und nicht für die Hörereignisentfernung (vergleiche Wendt [7])?

Dem heute gebräuchlichen intensitätsstereofonen Übertragungsverfahren lag bisher keineswegs eine allgemeingültige Theorie zur „Bildung von Hörereignissen zwischen den Lautsprechern“ zugrunde. Vielmehr hat man die Möglichkeit, durch geeignete Wahl der gegenseitigen Pegeldifferenzen der Lautsprecher-signale die Hörereignisrichtung im Winkelbereich der Lautsprecher willkürlich festzulegen, ganz pragmatisch ausgemessen und für die stereofone Übertragung im Hörfunk genutzt. Die Intensitätsstereofonie hat sich ziemlich unabhängig von entsprechenden wissenschaftlichen Untersuchungen entwickelt; sie muß im wesentlichen als technisches Resultat der elektroakustischen Übertragungspraxis angesehen werden.

Für die bisherige Entwicklung der Kunstkopfstereofonie dagegen waren wissenschaftliche Erkenntnisse über die Funktion des Gehörs bei der Lokalisation entscheidend. Dies betraf das Verständnis über die Auswertung der Ohrsignale hinsichtlich Hörereignisrichtung, -entfernung und -ausdehnung, insbesondere Vorstellungen über die Ursachen der Im-Kopf-Ortung, der Vorne-Hinten-Vertauschung, der Hörereigniselevation. Die entsprechenden Theorien dazu stehen in guter Übereinstimmung mit experimentellen Ergebnissen, die auch im Zuge der Kunstkopfforschung gewonnen wurden.

Doch vor allem das bedeutsame Problem der Kompatibilität raum/kopfbezogener Signale hat Fragen aufgeworfen, die sich nicht mehr mit den bekannten speziellen Theorien, jede mit eng begrenztem Gültigkeitsbereich, beantworten lassen. Vielmehr müssen für diesen Problemkomplex gerade die Zusammenhänge unter den verschiedenen Teilaspekten der Lokalisation verstanden werden. Der in dieser Beziehung wichtige Teilaspekt liegt vor, wenn mehrere Schallquellen einem gemeinsamen Hörereignis zugeordnet sind, so daß Hörereignisort und Schallquellenorte nicht übereinstimmen, beispielsweise bei Intensitätsstereofonie. Aber es hat sich erwiesen, daß keine der bisher bekannten Theorien zur Intensitäts-

¹ Dr.-Ing. Günther Theile ist Leiter des Arbeitsbereiches Systeme der Ton-Aufnahme/Wiedergabe im Institut für Rundfunktechnik, München.

stereofonie übertragbar ist auf die Lokalisation einer Einzelschallquelle [8]. Ebenso gibt es keine allgemeine Theorie zur Lokalisation, welche die Hörereigniseigenschaften auch dann einheitlich erklärt, wenn mehr als eine Schallquelle die Beschaffenheit der Ohrsignale bestimmt.

Ein Ansatz für eine allgemeine Theorie zur Lokalisation ergibt sich, wenn die Lokalisation konsequent als Folge eines Wahrnehmungsprozesses angesehen wird, der allein aufgrund der Hörerfahrung möglich ist. Die Hörerfahrung ergibt sich aus den natürlichen Hörbedingungen. Und unter natürlichen Bedingungen erfolgt das Hören vorwiegend im überlagerten Schallfeld. Störschallquellen aller Art, sehr häufig Reflexionen am Boden oder an den Begrenzungsflächen eines Raumes (Spiegelschallquellen), aber auch eine Vielzahl gleichwertiger Schallquellen, beispielsweise verschiedene Sprecher in der Nähe eines Hörers, erzeugen normalerweise Ohrsignale, die in keiner Weise mit dem Elementarfall einer Einzelschallquelle übereinstimmen. Aber das menschliche Gehör ist in der Lage, unter diesen Schallfeldbedingungen „einen sehr sinnvollen Auswahl-, Ordnungs- und Gliederungsprozeß durchzuführen“ [9], so daß trotz der Überlagerung die entsprechenden Hörereigniszustände kommen können. Bekannte Hörphänomene in diesem Zusammenhang sind das „Gesetz der ersten Wellenfront“ [10] und der „Cocktailparty-Effekt“ [11, 12].

In der vorliegenden Arbeit wird ein Modell für die Lokalisation im überlagerten Schallfeld beschrieben. Es basiert auf dem Ansatz, daß die Lokalisation über einen „Reizmustervergleich zwischen aktuellen Reizen und erlernten Reizmustern“ [9] erfolgt, daß also aufgrund der Hörerfahrung die Hörereigniszustände zustande kommen. Versteht man die Abhängigkeit der Ohrsignale vom Schallquellenort als einen Mechanismus zur Codierung räumlicher Information, so kann die Kenntnis dieser Ohrsignalabhängigkeit als Schlüssel zur Decodierung der räumlichen Information aufgefaßt werden. Im überlagerten Schallfeld stellt sich damit die Lokalisation als Prozeß zur gleichzeitigen Decodierung verschiedener räumlicher Informationen dar. Er gelingt ganz, teilweise oder gar nicht, abhängig von der Anzahl und den Eigenschaften der beteiligten Schallereignisse.

Es zeigt sich, daß die gleichzeitige Decodierung verschiedener räumlicher Informationen im überlagerten Schallfeld auch dann noch gelingen kann, wenn das Schallfeld von den beiden Lautsprechern, die stereofone Signale abstrahlen, hervorgerufen wird [8]. Damit ist das Lokalisationsmodell grundsätzlich geeignet für eine Vorhersage des Hörereignisortes, sowohl für den Fall einer Realschallquelle als auch für den Fall der „Phantomschallquelle“ bei Stereofonie über Lautsprecher.

Die Kompatibilität raum/kopfbezogener Signale war aber nicht nur hinsichtlich des Hörereignisortes zu klären. Besonders problematisch hat sich die Frage der Hörereignisklangfarbe erwiesen. Es hat bisher kein Lokalisationsmodell gegeben, welches die Wirkung der ausgeprägten Richtcharakteristik des Außenohres (bzw. des Kunstkopfes) auf das Leistungsdichtespektrum des Ohrsignals (bzw. des Kunstkopfsignals) im Hinblick auf die Klangfarbe des Hörereig-

nisses berücksichtigt. In der Literatur ist das Klangfarbenphänomen unter dem Aspekt der wirksamen Übertragungsfunktion des Außenohres überhaupt erst in jüngster Zeit beschrieben worden [6, 8, 13]; entsprechende systematische Hörversuche stehen noch aus.

Welchen Einfluß hat der Prozeß der Decodierung der räumlichen Information auf die Klangfarbe des Hörereignisses? Welche Ohrsignalanteile diskriminiert das Gehör für den Ort und/oder für die Klangfarbe des Hörereignisses? Welche Schallereignismerkmale gewährleisten dies?

Es ist zweckmäßig, an dieser Stelle zunächst den Begriff „Schallereignis“ präzise zu definieren. Bisher ist unklar, wie zwei Schallereignisse überhaupt voneinander unterschieden werden sollen. Betrachtet man entsprechend Blauert [15] ausschließlich „die physikalische Seite des Hörvorgangs“, so kann damit allein die Wirkung von Kopf und Außenohren im Schallfeld gemeint sein. Zwei verschiedene, simultane Schallereignisse sind dann aber nur infolge verschiedener räumlicher Merkmale vorhanden (Orte der Schallquellen, Ausbreitungsrichtungen der Schallwellen usw.). Diese Betrachtung erscheint im überlagerten Schallfeld deshalb nicht sinnvoll, weil der Hörvorgang hier wesentlich durch die Diskriminierbarkeit einzelner Signalanteile gekennzeichnet ist: Zwei räumlich getrennte Lautsprecher z. B. können – abhängig von den Eigenschaften der Sendesignale – entweder zwei simultane Hörereignisse an verschiedenen Orten oder ein einziges Hörereignis an einem dritten Ort erzeugen; weiterhin kann ein einziger Lautsprecher, der zwei „unterschiedliche Sendesignale“ abstrahlt, zwei simultane Hörereignisse erzeugen. Die Auswertungsprozesse im Gehör, die zur Bildung des Hörereignisortes und der Hörereignisgestalt führen, bestimmen stets gemeinsam die Hörereigniseigenschaften; sie beeinflussen sich scheinbar gegenseitig. Die physikalische Seite der Lokalisation ist erst mit Schallereignissen beschreibbar, die sich nicht allein hinsichtlich räumlicher Merkmale gegeneinander abgrenzen lassen. Im folgenden gilt deshalb für den Begriff folgende Definition:

Ein **Schallereignis** ist derjenige Schallanteil, der von einer Einzelschallquelle herrührt und der die Eigenschaft des zugeordneten Hörereignisses hinsichtlich Ort oder Gestalt bestimmt oder beeinflußt.

Wird unter Hörereignis nach Blauert [16] allgemein die „zeitlich, räumlich und eigenschaftlich“ bestimmte auditive Wahrnehmung verstanden, so ergibt sich damit für den Begriff „Lokalisation“ eine Definition, die auch für das überlagerte Schallfeld ausreicht:

Die **Lokalisation** ist das Zuordnungsgesetz zwischen dem Ort eines Hörereignisses außerhalb des Kopfes und bestimmten Merkmalen eines oder mehrerer Schallereignisse.

Das neue Lokalisationsmodell wird „Assoziationsmodell“ genannt. Es entspricht der Auffassung, daß die Wirkungsweise des Gehörs beim räumlichen Hören weniger in abgegrenzten Teilbereichen, sondern sinnvoller „ganzheitlich“ untersucht und verstanden werden kann. Es vermeidet bewußt das in der Nach-

richtentechnik übliche Bestreben, beobachtete Verknüpfungen zwischen Schall- und Hörereignismerkmalen sogleich als Gesetzmäßigkeit bzw. mit Hilfe eines Funktionsmodells zu formulieren (woraus sich die Notwendigkeit zur präzisen Abgrenzung ergibt). Vielmehr trägt es der Erkenntnis Rechnung, daß gerade die Grenzen der Gültigkeitsbereiche sowie mögliche Zusammenhänge einzelner Hörphänomene von besonderem Interesse sein müssen, ebenso verstärkt die Erkenntnisse aus anderen Wissenschaftszweigen, beispielsweise der Neurophysiologie und der Wahrnehmungsforschung. In dieser Hinsicht kann die Anwendung der assoziativen Speichertechnik und Signalverarbeitung neue Möglichkeiten schaffen.

2. Das Assoziationsmodell

Das Assoziationsmodell basiert auf der Hypothese, daß Assoziationsvorgänge ein Grundprinzip der sensorischen Reizverarbeitung darstellen. Im sensorischen System entspricht der hypothetische Assoziationsvorgang der Funktion eines hochwirksamen Filters zur Informationsreduktion, welche zwischen peripherer Rezeption und bewußter Wahrnehmung stattfindet. Gespeicherte Assoziationsmuster, die aufgrund früherer Erfahrungen geprägt werden, bewirken einerseits eine sinnvolle Informationsselektion. Andererseits ermöglichen sie, trotz reduzierter Information die streng determinierten Reizkonfigurationen der Umwelt mit ausreichender Genauigkeit zu erkennen [17, 18].

Können wegen mangelnder Erfahrung, fehlender Adaption oder aus organischen Gründen keine Assoziationen gebildet werden, entstehen anstelle bewußter Wahrnehmungen nur bedeutungslose Empfindungen. Man kennt diese Erscheinung in der Neurophysiologie. Eine Ausschaltung der sogenannten Assoziationsfelder im sensorischen System hat Wahrnehmungsstörungen zur Folge, und zwar fehlt die Fähigkeit, die Sinneseindrücke, die empfunden werden, zu erkennen (Agnosie). Agnosien können sich beim Menschen für verschiedene Sinnesmodalitäten entwickeln. Wird z. B. das akustische Assoziationsfeld (im linken Temporallappen) zerstört, so geht unter anderem das Sprachverständnis verloren. Obwohl der Betroffene noch hören kann, bleibt ihm die Bedeutung des Signals verborgen. Es ist allerdings nicht bekannt, ob sich ebenso die Lokalisationsfähigkeit unterbinden läßt (ein derartiger physiologischer Nachweis einer „Ortsassoziation“ wäre sehr aufschlußreich).

Die Bedeutung von Assoziationsprozessen für bewußte Wahrnehmungen führt auf eine sinnvolle Möglichkeit, die Begriffe „Empfindung“ und „Wahrnehmung“ terminologisch zu trennen:

Empfindungen sind die durch Reize ausgelösten, nicht weiter gliederbaren Sinnesereignisse, die sich gegenüber Lernvorgängen und bewußten oder unbewußten Interpretationen invariant verhalten; sie kommen nicht durch Assoziationen zustande.

Wahrnehmungen sind die durch Reize ausgelösten Sinnesereignisse, die infolge der Sinneserfahrung der Außenwelt zugeordnet auftreten und die deshalb durch Lernvorgänge und bewußte oder unbewußte Interpretationen beeinflusst werden können; sie kommen durch Assoziationen zustande.

Aus dieser Abgrenzung folgt für die Entwicklung von Perzeptionsmodellen, daß grundsätzlich geklärt sein sollte, ob ein Empfindungs- oder Wahrnehmungsmodell den Prozeß beschreiben soll bzw. kann.

Das Grundproblem bei der Untersuchung von Wahrnehmungsvorgängen ist damit die erforderliche Abgrenzung. Hier liegen Fehlerquellen, weil ohne Kenntnis des gesamten Zusammenhanges nicht entschieden werden kann, ob oder mit welcher Einschränkung aus dem Wahrnehmungsprozeß bestimmte Teilbereiche herausgelöst beschrieben werden können.

Es ist mit Vorbehalt wohl möglich, das räumliche Hören abgetrennt von der visuellen Wahrnehmung zu betrachten. Doch wie läßt sich begründen, das Richtungshören abgetrennt vom Entfernungshören zu betrachten? Richtung und Entfernung sind nur Koordinaten des Hörereignisortes; es gibt keine Richtung ohne Entfernung. Ein Modell für das Richtungshören ist kein Wahrnehmungsmodell, es beschreibt deshalb nicht zwangsläufig eine spezielle Funktion des Gehörs beim räumlichen Hören. Ebenso fragwürdig ist die Untersuchung der Lokalisation mit Tönen: Wie könnte das Gehör die Entfernung einer Schallquelle, die einen reinen Ton abstrahlt, bestimmen?

Bei dem zu beschreibenden Lokalisationsmodell wird davon ausgegangen, daß Richtung und Entfernung untrennbare Größen der räumlichen Wahrnehmung sind. Auffällig ist, daß auf dem Gebiet der visuellen Wahrnehmung viele echte Wahrnehmungsmodelle bekannt sind, dagegen für das Hören kaum. Das ist aus neurophysiologischer und informationstheoretischer Sicht nicht einzusehen, denn hier bestehen für das Auge und das Ohr, in ihrer Eigenschaft als Nachrichtenempfänger, keine prinzipiellen Unterschiede. Aber es ist aus zwei Gründen verständlich: Erstens besteht für technische Anwendungen der visuellen Wahrnehmungsmodelle (Datenreduktion durch geeignete Quellencodierung für die Bildübertragung, beispielsweise Nachbildung von Mustererkennungsprinzipien mit Hilfe von Klassifikationsprozessen oder adaptiven Filtern) aus kommerziellen Gründen ein wesentlich größeres Interesse. Zweitens sind Mustererkennungs- und Assoziationsprinzipien beim Sehen offensichtlicher vorhanden als beim Hören.

Das Assoziationsmodell ist durch eine **zweidimensionale Reizverarbeitung** wesentlich gekennzeichnet. Die erste Verarbeitungsstufe wird „Ortsassoziationsstufe“ genannt. Sie ist infolge der Hörererfahrung in der Lage, durch Bildung von Ortsassoziationen solche Ohrsignale zu selektieren, welche ausschließlich durch die Wirkung von Kopf und Ohrmuscheln im (überlagerten) Schallfeld miteinander verkoppelt sind. Im folgenden werden diese Ohrsignalanteile als „Lokalisationsreize“ bezeichnet:

Hinreichend breitbandige Ohrsignale oder Ohrsignalanteile an den Trommelfellen oder beiden Ohren bilden zusammen einen **Lokalisationsreiz**, wenn sie sich hinsichtlich der Zeit- und der spektralen Merkmale einem einzigen Schallereignisort zuordnen lassen.

Die Ortsassoziationsstufe ist im Modell einer zweiten, höher gelagerten zentralen Verarbeitungsstufe

vorgeschaltet. Diese sogenannte „Gestaltassoziationsstufe“ enthält die Prozesse, welche die eigenschaftlichen Merkmale des Hörereignisses mit Ausnahme der räumlichen Merkmale bestimmen.

Die konsequente Unterscheidung der beiden Verarbeitungsstufen entspricht völlig den beiden elementaren Bereichen der Hörerfahrung: Die empfangenen Ohrsignale sind zurückzuführen auf die beiden voneinander unabhängigen, stets paarweise auftretenden Schallquelleneigenschaften „Ort“ und „Signal“. Demzufolge sind im Modell die auftretenden Hörereignisse zurückzuführen auf die Wirkung einer ortsbestimmenden und einer gestaltbestimmenden Verarbeitungsstufe.

Beide Stufen müssen durchlaufen werden, damit der Reiz zu einer Wahrnehmung führt. Beide Prozesse bestimmen also stets gemeinsam die Eigenschaften eines Hörereignisses; auch die gestaltbestimmende Verarbeitungsstufe ist damit ein Element des **Lokalisationsmodells**. Man erkennt, daß beispielsweise der oben erwähnte Zusammenhang zwischen Hörereignisort und Hörereignisklangfarbe prinzipiell vom Assoziationsmodell erfaßt wird.

2.1. Die Lokalisationsreizselektion

Die Ortsassoziationsstufe des Modells hat die Eigenschaft, einen empfangenen Reiz vorzugsweise als Lokalisationsreiz zu deuten, das heißt sie vergleicht spontan den aktuellen Reiz mit Reizmustern, welche infolge der Erfahrung bestimmten Hörereignisorten zugeordnet sind. Nur ein Lokalisationsreiz führt zur Lokalisation. Er liegt vor, wenn die Ohrsignalmerkmale hinsichtlich Zeit und Spektrum mit der Erfahrung vereinbar sind. Damit ist das Gehör in der Lage, aus einer Summe von Signalen die für einen Schallquellenort charakteristischen Signalanteile zusammenzufassen und geschlossen weiterzuleiten (Lokalisationsreizselektion). „Geschlossen weiterleiten“ heißt dabei: Die Lokalisationsreizselektion läßt sich als Fusionsprozeß interpretieren, welcher die binauralen Signalanteile einer bestimmten Schallquelle untrennbar vereint und mit der umcodierten, zur Wahrnehmung führenden räumlichen Information versieht. Die Lokalisationsreizselektion wirkt im überlagerten Schallfeld als Filter zur gleichzeitigen Diskrimination einzelner Sendesignale.

Diese Funktion der hypothetischen Ortsassoziationsstufe läßt sich wahrscheinlich nicht mit Operatoren nachbilden, die nach unserem heutigen Wissen physiologisch möglich sind. Vielmehr sollen die Leistungsmerkmale dieser Verarbeitungsstufe – soweit wie möglich – mit Mitteln der linearen Systemtheorie beschrieben werden, und in diesem Rahmen auch nur für maximal zwei Schallereignisse beliebiger räumlicher Konstellation. Als grobes Schema wird ein steuerbares Filter angenommen, dessen Übertragungsfunktion sich invers zur ortsabhängigen Übertragungsfunktion des Außenohres verhält. Die Bestimmung der momentan wirksamen Übertragungsfunktion geschieht durch Mustererkennungsprozesse.

2.1.1. Die räumliche Decodierung

Das Gehör besteht aus einem Auswertesystem mit zwei Eingangskanälen. Den Eingangskanälen ist ein gemeinsames lineares Netzwerk vorgeschaltet, dessen Übertragungsfunktion M sich aus der Wirkung

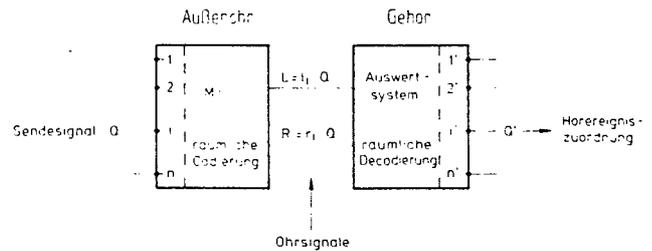


Bild 1
Das räumliche Übertragungssystem des Gehörs

des Kopfes und der Außenohren im Schallfeld ergibt. M ist eine Funktion des Schallquellenortes i , welche die Ohrsignale definiert miteinander verknüpft („räumliche Codierung“). Liegt am Eingang i der Matrix $M(i)$ ein Sendesignal Q an, so treten an den Eingängen des Auswertesystems die verknüpften Signale der Form $l_i \cdot Q = L$ und $r_i \cdot Q = R$ auf (l_i, r_i = akustische Übertragungsfaktoren bei Beschallung vom Ort i für das linke Ohr bzw. rechte Ohr); nach erfolgter räumlicher Decodierung erscheint die Antwort Q' am zugeordneten Ausgang i' (**Bild 1**).

Man kann nun zeigen [19], daß unter bestimmten Bedingungen eine inverse Matrix M^{-1} existiert, welche die Verknüpfung der Ohrsignale, die infolge der Matrix M auftritt, wieder vollständig aufhebt. Es ist also im Gehör ein Auswertesystem zur räumlichen Decodierung denkbar, das beispielsweise auf adaptivem Wege die Übertragungsfunktion M^{-1} annimmt. Hierzu sind lediglich Kenntnisse der interauralen Übertragungsfunktion notwendig. Damit ist zunächst auf einfachem Wege der systemtheoretische Nachweis möglich, daß im überlagerten Schallfeld prinzipiell zwei Schallquellen gleichzeitig diskriminierbar sind, sofern über hinreichend verschiedene elektroakustische Übertragungsfaktoren übertragen wird. Das Gehör kann theoretisch die Signalanteile trennen, die von dem Lautsprecher A und dem Lautsprecher B einer stereofonen Anordnung herrühren.

Zur genaueren Betrachtung der räumlichen Decodierung gibt **Bild 2** die Strukturen der zueinander inversen Netzwerke wieder. Die Schallquellen erzeugen an den Trommelfellen die Summensignale

$$L = l_1 \cdot A + l_2 \cdot B \quad (1)$$

bzw.

$$R = r_2 \cdot B + r_1 \cdot A, \quad (2)$$

wobei $l_{1,2}$ und $r_{1,2}$ die ortsabhängigen elektroakustischen Übertragungsfaktoren für das linke bzw. rechte Ohr darstellen (Quelle A befindet sich am Ort 1,

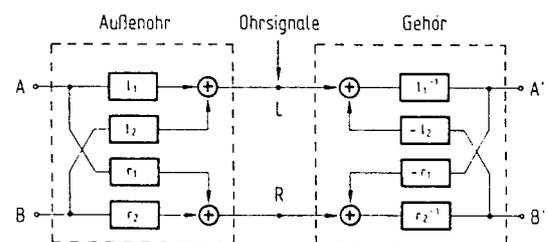


Bild 2
Räumliche Codierung und Decodierung bei zwei Schallquellen

Quelle B am Ort 2). Das Außenohr wirkt als Matrix M , die dem Signal der Quelle A bzw. B die räumliche Codierung aufprägt. Die räumliche Decodierung im Gehör geschieht nach erfolgter Mustererkennung in der angepaßten Matrix M^{-1} , deren inverse Übertragungsfunktion mit den Gleichungen

$$A' = \frac{r_2 \cdot L - l_2 \cdot R}{l_1 \cdot r_2 - r_1 \cdot l_2}, \quad (3)$$

$$B' = \frac{l_1 \cdot R - r_1 \cdot L}{l_1 \cdot r_2 - r_1 \cdot l_2} \quad (4)$$

beschrieben werden kann. Durch Einsetzen der Gleichungen (1) und (2) in (3) bzw. (4) ist die räumliche Decodierung nachgewiesen:

$$A' = \frac{r_2 \cdot l_1 \cdot A + r_2 \cdot l_2 \cdot B - l_2 \cdot r_2 \cdot B - l_2 \cdot r_1 \cdot A}{l_1 \cdot r_2 - r_1 \cdot l_2} = A, \quad (5)$$

$$B' = \frac{l_1 \cdot r_2 \cdot B + l_1 \cdot r_1 \cdot A - r_1 \cdot l_1 \cdot A - r_1 \cdot l_2 \cdot B}{l_1 \cdot r_2 - r_1 \cdot l_2} = B. \quad (6)$$

Die räumliche Decodierung kann demnach mit Hilfe eines linearen Netzwerkes erfolgen, das sich in der geschriebenen Form „an den Schallquellenort angepaßt“ verhält. Es ist an dieser Stelle ohne Belang, ob die Anpassung mit einem variablen Netzwerk oder mit variablen Bewertungsfaktoren an den Ausgängen einer Bank von Netzwerken nachgebildet wird.

Das Wesen der Lokalisation liegt offenbar in einem Erkennungsprozeß, der die Information für die optimale Anpassung liefern kann und damit erst die Voraussetzung schafft für eine Isolation bestimmter Signalmuster.

Das Wesen der Lokalisation soll deshalb in einem Erkennungsprozeß liegen, der zur Selektion führt. Ganz entsprechend enthält das Funktionsschema der Ortsassoziationsstufe einen Baustein „Mustererkennung“, der die Information für die optimale Filteranpassung liefert. Der Mustererkennungsprozeß schafft die Voraussetzung für eine Isolation bestimmter Signalmuster.

2.1.2. Die assoziative Mustererkennung

Eine wichtige Arbeitshypothese zur Entwicklung entsprechender Modelle für die Mustererkennung stellt die Annahme dar, daß bestimmte aktuelle Reize ganz bestimmte Assoziationsprozesse auslösen. Diverse adaptive Vorgänge bei der Lokalisation legen nahe, daß sich der Zusammenhang zwischen Hörereignis- und Schallereigniseigenschaft mit Assoziationsprozessen erklären läßt [8]. Das gleiche gilt für verschiedene „Lokalisationsfehler“. Wenn Assoziationsprozesse bewirken, daß Hörereignisort und Schallquellenort normalerweise deckungsgleich auftreten, so müssen entsprechende Lokalisationsabweichungen gerade mit diesen Assoziationsprozessen erklärbar sein. Tatsächlich scheint das nicht nur für das Phänomen „Phantomschallquelle“ zuzutreffen [20]. Folgende weitere Beispiele für Divergenzen zwischen Hörereignisort und Schallquellenort entstehen unter speziellen, widernatürlichen Umständen:

1. Die Hörereignisrichtung in der Medianebene steht im direkten Zusammenhang mit dem Signalspek-

trum; sie verhält sich bei schmalbandigen Signalen von der Schalleinfallrichtung unabhängig. Die Abhängigkeit vom Spektrum läßt sich aus den linearen Verzerrungen ableiten, die vom Außenohr verursacht werden (Blauert [16]) und zur Prägung von richtungsbestimmenden Reizmustern führen.

- Die Hörereignisrichtung in der Medianebene steht im direkten Zusammenhang mit dem Signalspektrum und -spektrum [21]. Bei Lautsprecherdarbietung im kleinen Schallfeld verhält sie sich abhängig von der Hörerfahrung [9]: Entscheidend ist das Verhältnis der Lautheit und Klangfarbe zur assoziierten Lautstärke und „Tonlage“ der Schallquelle.
- Die Hörereignisrichtung entspricht nicht der Schalleinfallrichtung bei Fehlanpassung an die bisherige Hörerfahrung, beispielsweise nach plötzlicher Normalisierung eines einseitig krankhaft veränderten Gehörs (Operation, [22]).
- Der Hörereignisort läßt sich, in bestimmten Grenzen unabhängig vom Schallquellenort, durch Assoziationslenkung beeinflussen, beispielsweise durch begleitende akustische oder optische Reize (z. B. [14, 23 („räumliche Komplikation“), 24, 25, 26]).
- Die vielfältigen Ursachen der Im-Kopf-Ortung lassen sich zusammenfassen, wenn man davon ausgeht, daß die Lokalisation über einen Reizmustervergleich zwischen aktuellem Reiz und erlernten Reizmustern erfolgt [27, 28]. Im-Kopf-Ortung tritt auf, sobald der Reiz sich nicht einem ortsbestimmenden Reizmuster zuordnen läßt; sie kann durch Assoziationslenkung verhindert werden (z. B. optische Information, [29]).

Bevor das vollständige Lokalisationsmodell entwickelt wird, soll die Bedeutung des Assoziationsprinzips auf anderen Forschungs- und Anwendungsgebieten kurz umrissen werden. Es zeigt sich dabei, daß wesentliche Bereiche des Wahrnehmungsprozesses mit einer Art „assoziativen Mustererkennung“ erklärt und auch in stark vereinfachter Form beschrieben werden können, daß andererseits jedoch Assoziationsvorgänge – beispielsweise aus neurophysiologischer Sicht – in der nervösen Verarbeitung des zentralen Nervensystems nicht einheitlich als erwiesen gelten.

Aus wahrnehmungspsychologischer und informationstheoretischer Sicht geben Assoziations- und Mustererkennungsmechanismen aber die oft einzig mögliche Erklärung bestimmter Phänomene der auditiven (und visuellen) Wahrnehmung. Dazu gehört z. B. die Erscheinung der „simultanen Hörereignisse“ bei Darbietung verschiedener Signale über einen Lautsprecher. Das entspricht nachrichtentechnisch dem Empfang getrennter Signale nach Übertragung über einen einzigen Kanal mit der Bandbreite eines Signals. Diese grundlegende Fähigkeit unseres Gehörs ist leicht nachzuweisen und nachrichtentechnisch wohl nur mit der informationsreduzierenden Wirkung von Mustererkennungsmechanismen erklärbar.

Assoziative Mustererkennung ist ein Prozeß, der ein aktuelles Muster mit einem gespeicherten Muster verknüpft, selbst dann, wenn nur Teile des

gespeicherten Musters im aktuellen Muster enthalten sind.

Dieser bis heute noch hypothetische Mechanismus ist seit langer Zeit auf mehreren Forschungsgebieten Gegenstand des Interesses, so besonders in der Kybernetik, vergleiche Flechtner [30]. In der Neurophysiologie wird die Existenz eines „sensorischen Assoziationssystems“ für die Verarbeitung der sensorischen Erregungen im zentralen Nervensystem vermutet (z. B. [31]); es gilt indes als nahezu unerforscht. Die hier erkennbaren Vorstellungen sind stark beeinflusst von kybernetischen Modellen. Besonders in jüngerer Zeit versucht man intensiv, die speziellen Fähigkeiten des menschlichen Gehirns nachzubilden, um sie zu untersuchen oder technisch anzuwenden. Assoziative Informationsspeicherung und assoziativer Informationsruf stellen dabei offenbar ein Grundprinzip der Verarbeitung dar [32 bis 40].

Ein „Modell eines neuralen Netzwerkes“ von Wigstroem [41] beispielsweise:

„... das für die Beschreibung der Gehirnrinde gedacht ist, besteht aus einem Netzwerk von Zellen, die von erregendem und hemmendem Typ sind. Es hat die Fähigkeit für assoziatives Lernen. Es wird gezeigt, daß unter geeigneten Bedingungen das Ausgangsmuster von nur einer Hauptkomponente gebildet wird, auch wenn das erregende Eingangsmuster aus einem Gemisch von mehreren Mustern, die beim Lernprozeß Verwendung fanden, besteht. Diese Hauptkomponente ist ein Teil des speziellen Ausgangsmusters, das beim Einlernen mit dem Eingangsmuster assoziiert wurde. Das Verhalten wird durch einen dynamischen Prozeß erreicht, in dem die Musterseparationseigenschaften der Rückkopplungskreise eine wichtige Rolle spielen. Die Funktion des Modells kann als Mustererkennung betrachtet werden.“

Als weiteres Beispiel sei die Arbeit von Willwacher [42] angegeben, der die „Fähigkeiten eines assoziativen Speichersystems im Vergleich zu Gehirnfunktionen“ untersuchte. Das vorgestellte Netzwerk vermag in stark vereinfachter Weise Leistungen des menschlichen Gehirns nachzuahmen: parallele Assoziation (vollständiger Aufruf eines Musters durch Eingabe eines Teils des Musters), serielle Assoziationen (Aufruf einer zeitlichen Mustersequenz durch Eingabe eines Anfangsmusters), Zuordnung eines unbekanntes (nicht gespeicherten) Musters, Zuordnung von Mustern aus zwei Systembereichen, Assoziation einer wahrscheinlichen Musterfolge, Störung des Assoziationsvorgangs, „Eselsbrücke“, „Abstraktion des Gemeinsamen“, „Umlernen“, „produktiver Einfall“.

Diesen und anderen kybernetischen Ansätzen ist gemeinsam, daß die verwendeten Systembausteine ein den realen Neuronen analoges Verhalten zeigen. Da bis heute keine bis ins einzelne gehenden Befunde der Neurologie über die Struktur des Gehirns und die Art der Mustererkennung vorliegen, synthetisiert man mit Hilfe der bekannten Bausteine entsprechende Gehirnleistungen, um daraus Hypothesen zur Gehirnfunktion zu gewinnen. Die zentrale Hypothese lautet, daß Assoziationsvorgänge ein Grundprinzip der sensorischen Reizverarbeitung darstellen.

2.1.3. Ein Funktionsschema der Ortsassoziationsstufe

Die Ortsassoziationsstufe leistet die Lokalisationsreizselektion. Der Verarbeitungsprozeß dieser Stufe läßt sich beispielsweise mit der Wirkung eines orteabhängigen Filters beschreiben, dessen Parameter

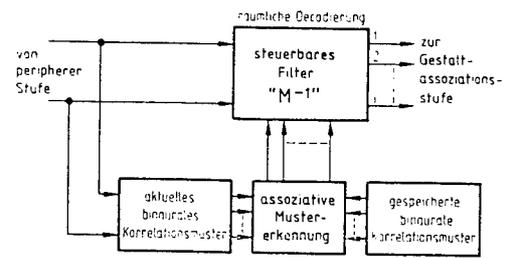


Bild 3

Funktionsschema der Ortsassoziationsstufe

Das adaptive Filter „M-1“ erzielt die räumliche Decodierung entsprechend Bild 1, sobald die assoziative Mustererkennung erfolgt

aufgrund der assoziativen Mustererkennung gesteuert werden. Als zweckdienliches Signalmuster für den Erkennungsprozeß wird ein in [8] definiertes „binaurales Korrelationsmuster“ gewählt. Das Funktionsschema gibt Bild 3 wieder.

Es sei angenommen, daß die Verknüpfung zwischen Ein- und Ausgangssignal das Verhalten dieser hypothetischen Verarbeitungsstufe beschreibe. Das Schema soll nicht die innere Struktur der sensorischen Verarbeitung nachbilden. Die Funktion der Stufe setzt voraus, daß der Erfahrungsprozeß stattgefunden hat, daß also das gespeicherte binaurale Korrelationsmuster vorliegt.

Kennzeichnend ist die selektive Eigenschaft. Das von der peripheren Stufe ankommende Signal enthält eine Rauminformation und eine Gestaltinformation (einer Schallquelle). Die Rauminformation wird mit Hilfe des Mustererkennungsprozesses erkannt, die Gestaltinformation mit Hilfe des adaptiven Filters „M-1“ diskriminiert und der Gestaltassoziationsstufe zugeführt. Mit anderen Worten: Das ankommende Signal wird mit Hilfe des ortsangepaßten Filters „M-1“ von der Beeinflussung durch das Außenohr entsprechend dem Schallquellenort befreit, die Bewertung des Sendesignals durch die wirksame Übertragungsfunktion wird aufgehoben, das reine Sendesignal sowie die gewonnene Richtungs- und Entfernungsinformation werden **getrennt** weitergeleitet.

2.2. Die Wirkungsweise des Assoziationsmodells

Die vorangegangenen Überlegungen führen auf das in Bild 4 gezeigte Funktionsprinzip des Assoziationsmodells. Es umfaßt neben der peripheren Stufe, in der die Ohrsignale zunächst mit Hilfe von Filterbänken im Bereich etwa konstanter relativer Bandbreite spektral zerlegt werden (vergleiche [15, 43, 44, 45]), die beiden zentralen Verarbeitungsstufen Ortsassoziationsstufe und Gestaltassoziationsstufe. Jeder der beiden Verarbeitungsprozesse geschieht auf dem Wege einer assoziativ gesteuerten Musterselektion: Bestimmte Ohrsignale lösen nach der peripheren Verarbeitung in der ersten zentralen Stufe eine Ortsassoziation und in der höhergelagerten Stufe eine Gestaltassoziation aus.

Beide zentralen Auswertungsstufen enthalten die im Prinzip gleichen Verarbeitungsmechanismen. Ebenso wie der für die Ortsassoziationsstufe bereits beschriebene Mechanismus zur Lokalisationsreizselektion enthält die Gestaltassoziationsstufe einen Mechanismus zur Selektion der Gestalt.

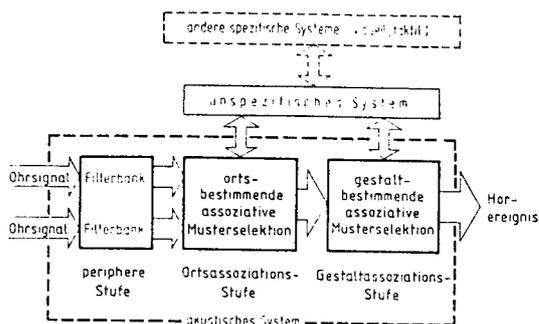


Bild 4

Das Funktionsprinzip des Assoziationsmodells

Geht man davon aus, daß beide Assoziationsprozesse nur aufgrund verschiedener gespeicherter Merkmalsmuster unterschiedlich wirken, so kann von den Fähigkeiten der Gestaltassoziationsstufe in bestimmten Grenzen [8] auf die Fähigkeiten der Ortsassoziationsstufe geschlossen werden: Die Fähigkeit der Gestaltassoziationsstufe beispielsweise, zwei (oder mehr) Schallereignisse auch bei identischer räumlicher Eigenschaft gleichzeitig diskriminieren zu können, entspricht der hypothetischen Fähigkeit der Ortsassoziationsstufe, zwei (oder mehr) Schallereignisse bei identischer Gestalteigenschaft gleichzeitig diskriminieren zu können.

Die Grenzen der Lokalisationsreizselektion sind aus informationstheoretischer Sicht erreicht, wenn die Informationsreduktion, die bei der assoziativen Signalverarbeitung erreicht wird, optimal genutzt wird. Sie sind in [8] in Hinblick auf die gleichzeitige Diskriminierbarkeit von zwei bestimmten Schallereignissen beliebiger räumlicher Konstellation näher untersucht worden.

Es wird angenommen, daß die Funktion des akustischen Assoziationsystems von einem sogenannten „unspezifischen System“ im Sinne einer Assoziationslenkung beeinflusst wird. (In der Neurophysiologie weiß man, daß dieses „Retikulärsystem“ mit allen Bereichen des Großhirns in Verbindung steht und man vermutet, daß es eine Koordination der Erregungsverarbeitung, beispielsweise der Verarbeitung akustischer und optischer Erregungen, sowie eine bewußtseins- und aufmerksamkeitsabhängige Erregungsauswahl bewirkt; vergleiche Caspers [31].) Beide Verarbeitungsstufen des akustischen Systems sind Bestandteil des Kurzzeitgedächtnisses und des Langzeitgedächtnisses; sie wirken im Sinne der für den Wahrnehmungsvorgang notwendigen Informationsreduktion.

Die Eigenschaft der Ortsassoziationsstufe ergibt sich aus der bereits beschriebenen Lokalisationsreizselektion, die sich als Wirkung eines adaptiven Filters auffassen läßt, das die räumliche Decodierung entsprechend den Gleichungen (5) und (6) vornimmt, sofern der entsprechende Mustererkennungsprozeß die für die Steuerung des Filters notwendige Information liefern kann.

Die Gestaltassoziationsstufe hat die Eigenschaft, einen empfangenen Reiz unabhängig von seiner räumlichen Information zu verarbeiten. Diese Stufe repräsentiert alle Mechanismen, die für eine inhalt-

liche Wahrnehmung des Reizes erforderlich sind und deren Teilbereiche für eine Vielzahl unterschiedlicher Hörphänomene verantwortlich sind. Hierzu gehören Mechanismen der Signalverschmelzung sowie der Erkennung und Bewertung von Musik und Sprache.

Von Bedeutung für das räumliche Hören sind im Assoziationsmodell die Gesetzmäßigkeiten der binauralen Signalerkennung (BMLD und BILD, siehe Blauert [15]), die der Gestaltassoziationsstufe zugeschrieben werden. Entsprechende Modelle, z. B. das „Akkumulationsmodell“ von Schenkel [46], das „EC-Modell“ von Durlach [47] oder das „Korrelationsmodell“ von Osman [48], sind also der Lokalisationsreizselektion nachgeschaltet.

Das Modell zur Tonhöhenwahrnehmung von Klängen von Terhardt [49] läßt sich besonders gut mit der assoziativen Verarbeitung der Gestaltassoziationsstufe vereinbaren. Es basiert ebenfalls auf Verarbeitungsprozessen, die sich aus der angelernten Kenntnis der Tonhöhenbeziehungen ergeben. Auch die Tonhöhen von Klängen werden hier als Hörereignis-Eigenschaft aufgefaßt, die – unter bestimmten Voraussetzungen – infolge der Hörerfahrung wahrgenommen wird. Im Sinne des Assoziationsmodells geschieht dies, nachdem durch die Wirkung der Ortsassoziationsstufe die spektralen Informationen, die aufgrund der Übertragungsfunktion des Außenohres auftreten, eliminiert werden.

Das Assoziationsmodell sagt voraus, daß die Tonhöhenwahrnehmung von Klängen ebenso wie die Wahrnehmung der Klangfarbe unabhängig ist vom Schallquellenort. Diese Erkenntnis entspricht der Erfahrung, wurde aber noch nicht experimentell belegt. Sie ergibt sich aus der Wirkung der Ortsassoziationsstufe: Wegen der Lokalisationsreizselektion gelangt allein das diskriminierte Sendesignal an die Gestaltassoziationsstufe; erst hier erfolgen die Prozesse, die zur Tonhöhen- und Klangfarbenwahrnehmung führen.

3. Diskussion und Konsequenzen des Modells

3.1. Summenlokalisierung

Ein wichtiger Spezialfall der Lokalisation im überlagerten Schallfeld liegt bei Stereophonie über Lautsprecher vor. Kennzeichnendes Merkmal dieser Technik ist offenbar die Tatsache, daß Hörereignisort und Schallquellenorte nicht übereinstimmen. Man spricht dann pauschal von einer „Phantomschallquelle“, weil ein Hörereignis dort wahrgenommen wird, wo sich keine Schallquelle befindet.

Im Sinne der vorangestellten Definition kann aber das Wort „Schallquelle“ nicht zur Beschreibung des Hörereignisraumes herangezogen werden. Die Phantomschallquelle wird deshalb in der Literatur auch im Sinne einer „gedachten Schallquelle“ verstanden, mit deren Hilfe man die physikalische Seite der Lokalisation zu beschreiben sucht. Man geht davon aus, daß die Phantomschallquelle grundsätzlich eine Ersatzschallquelle darstellt, in deren Schallfeld am Hörort die gleichen Ohrsignalmerkmale erzeugt werden wie im überlagerten Schallfeld.

Alle bekannten Phantomschallquellentheorien basieren gemeinsam auf der grundsätzlichen Annahme, daß „Summenlokalisierung“ [50] stattfindet. Darunter

wird die Vorstellung verstanden, daß aus den Schallfeldüberlagerungen an den Ohren Summensignale resultieren, deren Komponenten das Gehör nicht mehr trennen kann. Man vermutet deshalb bei der Lokalisation einer „Phantomschallquelle“ und bei der Lokalisation einer entsprechenden, am Ort der Phantomschallquelle befindlichen Realschallquelle eine Äquivalenz der Ohrsignalmerkmale am linken Ohr bzw. am rechten Ohr. In den Arbeiten über Summenlokalisation werden verschiedene Äquivalenztheorien beschrieben, deren Gültigkeitsbereiche sich jedoch auf das Richtungshören beschränken, oft sogar speziell auf das Richtungshören in der Horizontalebene [7, 51, 52, 53, 54].

Gegen das hypothetische Prinzip der Summenlokalisation wurde erst in jüngster Zeit ein „spektraler Einwand“ formuliert [8, 20]. Er besagt, daß die spektralen Eigenschaften die Ohrsignale, die aus der Schallfeldüberlagerung resultieren, sich nicht mit den Eigenschaften des zugeordneten Hörereignisses vereinbaren lassen. Weder die Entfernung, noch die Elevation, noch die Klangfarbe des Hörereignisses „Phantomschallquelle“ lassen sich den resultierenden Summensignalen an den Ohren unter der Annahme zuordnen, daß diese Ohrsignale einen einzigen Lokalisationsreiz darstellen. Vielmehr sind diese Eigenschaften des Hörereignisses „Phantomschallquelle“ von den Lokalisationsreizen bestimmt, die die Lautsprecher signale getrennt hervorrufen. Einzelheiten dazu sind in [8] ausführlich beschrieben.

Die Phantomschallquelle läßt sich hinsichtlich der an den beiden Ohren wirksamen Spektren nicht als Ersatzschallquelle auffassen. Diese Erkenntnis steht im Widerspruch zu den Summenlokalisationstheorien, wird dagegen vom Assoziationsmodell vorhergesagt: Infolge der Lokalisationsreizselektion werden die Lautsprecher signale aufgrund der unterschiedlichen Lautsprecherorte zunächst diskriminiert (Wirkung der Ortsassoziationsstufe). Erst nach erfolgter räumlicher Decodierung verschmelzen die Reize, weil die Lautsprecher hinreichend ähnliche Signale abstrahlen (Wirkung der Gestaltassoziationsstufe). Detaillierte Angaben findet man in [5, 8, 20].

Besonders in der Phantomschallquellsituation wird deutlich, daß der Hörereignisort (Richtung, Entfernung und Elevation) nicht einem Verarbeitungsprozeß zugeschrieben werden kann, welcher der Lokalisation einer entsprechenden Ersatzschallquelle gleichkommt. Das Assoziationsmodell befindet sich mit diesem Sachverhalt in Übereinstimmung, weil nicht grundsätzlich der eigentliche ortsbestimmende Verarbeitungsprozeß den Hörereignisort festlegt. Vielmehr führen im Modell die Ortsassoziationen nur in Verbindung mit zugeordneten Gestaltassoziationen zu Hörereignissen; es gibt keinen Hörereignisort ohne zugeordnete Hörereignisgestalt.

3.2. Ersatzschallquelle oder Phantomschallquelle?

Unterschiedliche Verarbeitungsprozesse im Gehör können also im Prinzip identische Hörereignisorte hervorbringen. Dies geschieht abhängig von den Ohrsignalen immer dann, wenn unterschiedliche Orts- und Gestaltassoziationen in ihrer Gesamtwirkung zur gleichen Lokalisation führen. Zwei Lautsprecher in Stereostandardaufstellung beispielsweise können

mit bestimmten Signalen genau die Ohrsignale erzeugen, die eine Realschallquelle verursachen würde (z. B. Tradis-Verfahren [55]); der Hörereignisort kommt als Folge einer Ortsassoziation zustande, die fiktive Schallquelle ist eine **Ersatzschallquelle**. Zwei andere Lautsprecher signale können aber auch ein Schallfeld erzeugen, welches entsprechend der wirksamen Schallereignisorte zwei Ortsassoziationen auslöst und dennoch den gleichen Hörereignisort wie im ersten Fall verursacht; die fiktive Schallquelle ist in diesem Fall eine **Phantomschallquelle**.

Kommt die Lokalisation eines Hörereignisses im überlagerten Schallfeld als Folge einer einzigen Ortsassoziation zustande, ist die fiktive Schallquelle eine **Ersatzschallquelle**.

Kommt die Lokalisation eines Hörereignisses im überlagerten Schallfeld als Folge von wenigstens zwei Ortsassoziationen zustande, so ist das Hörereignis die Folge einer **Phantomschallquelle**.

Aus diesen Definitionen leiten sich zwei weitere ab, die für die Beurteilung stereofoner Übertragungsverfahren von Bedeutung sind:

Kopfbezogene Stereophonie liegt unabhängig von der Beschallungsart immer dann vor, wenn eine Ersatzschallquelle existiert.

Raumbezogene Stereophonie liegt unabhängig von der Beschallungsart immer dann vor, wenn eine Phantomschallquelle existiert.

Wie sich zeigen wird, ermöglichen die Definitionen eine zweckmäßige Systemübersicht für die unterschiedlichen Möglichkeiten stereofoner Übertragung. Die Definitionen sind zweckmäßig, weil sie sich unmittelbar aus der Funktion des Gehörs ableiten. Die Funktion des Gehörs bei der Lokalisation im überlagerten Schallfeld wird mit dem Assoziationsmodell, keinesfalls aber mit Summenlokalisationsmodellen ausreichend einheitlich und allgemeingültig beschrieben.

Diese Aussage des Assoziationsmodells entspricht nicht dem in der Literatur zu findenden Verständnis über die Funktion des Gehörs beim räumlichen Hören. Man geht bisher davon aus, daß der überaus komplexe Prozeß, der zur Lokalisation führt, eindeutig beschreibbar sein müsse, wenn nur die räumlichen Eigenschaften des Hörereignisses eindeutig festliegen. Die Vorstellung ist unrichtig, weil unterschiedliche Ohrsignale die gleichen räumlichen Hörereigniseigenschaften hervorrufen können, sobald mehr als eine Schallquelle das Hörereignis bestimmt. Man geht weiterhin davon aus, daß der Lokalisationsprozeß grundsätzlich auf die Wirkung einer ortsbestimmenden Verarbeitungsstufe im Gehör zurückführbar sein müsse. Diese Vorstellung ist nicht zwingend und erscheint unwahrscheinlich, weil dann die ortsbestimmende Verarbeitungsstufe bestimmte Ohrsignale unterschiedlich interpretieren würde.

Demgegenüber ergeben sich die Zusammenhänge beim räumlichen Hören nach dem Assoziationsmodell wie folgt:

1. Eine Einzelschallquelle ruft in der ortsbestimmenden Stufe (unter bestimmten Bedingungen) eine Ortsassoziation hervor, welche allein den Hörereignisort festlegt.

2. Zwei Schallquellen, die hinsichtlich Ort und Sendesignal verschieden und unabhängig sind, rufen (unter bestimmten Bedingungen) in der ortsbestimmenden Stufe zwei Ortsassoziationen und in der gestaltbestimmenden Stufe zwei Gestaltassoziationen hervor. Entsprechend der Schallereignisse sind auch die Hörereignisse hinsichtlich Ort und Gestalt „entkoppelt“, so daß die Lokalisation allein auf die Wirkung der Ortsassoziationsstufe zurückzuführen ist.
3. Zwei Schallquellen, die nur hinsichtlich des Ortes verschieden und hinsichtlich des Sendesignals verkoppelt sind, rufen (unter bestimmten Bedingungen) in der ortsbestimmenden Stufe nach wie vor zwei Ortsassoziationen hervor. In der Gestaltassoziationsstufe kann aber entsprechend der Verkopplung der Sendesignale keine Entkopplung stattfinden. Da ein Hörereignis grundsätzlich von beiden Verarbeitungsstufen bestimmt ist, treten infolge der gemeinsamen Gestalteeigenschaften entsprechend verkoppelte Hörereignisse auf. Identität der Sendesignale führt deshalb zur vollständigen Hörereignisverschmelzung, also auch zur Verschmelzung der Hörereignisorte. In diesem Fall muß die Lokalisation sowohl auf die Wirkung der Orts- als auch der Gestaltassoziationsstufe zurückgeführt werden.

Allgemein gilt, daß die Lokalisation nur dann mit der Funktion einer ortsbestimmenden Verarbeitungsstufe erklärt werden kann, wenn allein ein Lokalisationsreiz zur Bildung des Hörereignisses beiträgt. Ein Lokalisationsreiz liegt vor, wenn hinreichend breitbandige, binaurale Ohrsignale ausschließlich durch die Wirkung von Kopf und Außenohren im Schallfeld miteinander verkoppelt sein können.

3.3. Gesetz der ersten Wellenfront, Cocktailparty-Effekt

Die grundsätzliche Unterscheidung einer ortsbestimmenden und einer gestaltbestimmenden Verarbeitungsstufe im Gehör führt auf ein besonderes Verständnis für Hörphänomene, die im überlagerten Schallfeld auftreten. Speziell solche Erscheinungen, die von der Beziehung der Sendesignale abhängen (Phantomschallquelle, Gesetz der ersten Wellenfront, Cocktailparty-Effekt), lassen sich im Assoziationsmodell infolge der „zweidimensionalen“ Verarbeitung prinzipiell erklären. Die Ortsassoziationsstufe diskriminiert einzelne Sendesignale aufgrund der räumlichen Informationen; erst die Reizantworten dieser Stufe unterliegen in der nachfolgenden Stufe der gestaltbestimmenden Verarbeitung, erst hier kommt also die Beziehung der Sendesignale zur Wirkung. Sofern die beteiligten Sender selegierbare und simultane Lokalisationsreize hervorrufen, können bei festen Sendeorten durch Verändern der Sendesignalbeziehungen nur die Gesetzmäßigkeiten der Gestaltassoziationsstufe untersucht werden, – und zwar auch dann, wenn die Sendesignalbeziehungen den Hörereignisort beeinflussen.

3.3.1. Das Gesetz der ersten Wellenfront

Die Horizontalrichtung einer Phantomschallquelle hängt von der eingefügten Laufzeit- und/oder Pegeldifferenz der beiden breitbandigen Lautsprechersignale ab. Diese Erscheinung (Laufzeitstereofonie

ebenso wie Intensitätsstereofonie) erklärt das Assoziationsmodell mit der Lokalisationsreizelektion einerseits und mit den Gesetzmäßigkeiten der Gestaltassoziationsstufe andererseits: Der erste Lokalisationsreiz (bzw. derjenige größerer Intensität) an der Gestaltassoziationsstufe hat ein größeres Gewicht. Im Zusammenhang mit Kompatibilitätsfragen bei Kunstkopfstereofonie wird hierüber ausführlich in [5] berichtet.

Die jeweils zuerst an der Gestaltassoziationsstufe eintreffenden Reizantworten dominieren im Verarbeitungsprozeß dieser Stufe, die späteren werden gedämpft; dies um so mehr, je größer die Zeitdifferenz der eintreffenden Lokalisationsreize wird. Erst nach Überschreitung einer Zeitdifferenz von etwa 10 bis 30 ms bekommt der nacheilende Lokalisationsreiz wieder zunehmendes Gewicht, er wird oberhalb der Echoschwelle (Definition z. B. in [15]) als getrenntes Hörereignis wahrnehmbar. Darüber hinaus vermittelt das Modell, daß im gesamten Zeitdifferenzbereich zwischen 0 und 10 bis 30 ms kein prinzipieller Unterschied in der Reizauswirkung besteht. Vielmehr wird angenommen, daß die Gesetzmäßigkeiten sowohl der Laufzeitstereofonie als auch der ersten Wellenfront auf eine zeitabhängige Bewertung der nacheinander eintreffenden Lokalisationsreize zurückführbar sind.

Die Ortsassoziationsstufe wirkt im überlagerten Schallfeld als Filter zur Diskrimination der Sendesignale, so daß nur in der nachgeschalteten Gestaltassoziationsstufe die Sendesignalbeziehung bewertet wird. Zwei Sendesignale mit Laufzeitunterschieden führen dazu, daß zwei Lokalisationsreize nicht zugleich eintreffen. Alle resultierenden Gesetzmäßigkeiten für die zugeordneten Hörereignisorte lassen sich als „Gesetz des ersten Lokalisationsreizes“ verstehen.

Darüber hinaus sagt das Modell aus, daß Laufzeitstereofonie auch für stationäre breitbandige Signale möglich ist, dagegen das Gesetz der ersten Wellenfront nur an Signalen mit ausgeprägten Hüllkurven nachzuweisen ist. Und es sagt weiter aus, daß die durch Signalverzögerungen verursachten Effekte im Prinzip auch bei Kopfhörerdarbietung auftreten müssen, sofern eine ausreichende Lokalisierbarkeit der Einzelschallquellen vorliegt. Nähere Einzelheiten dazu in [8].

3.3.2. Der Cocktailparty-Effekt

Die Selektionseigenschaften sowohl der Ortsassoziationsstufe als auch der Gestaltassoziationsstufe ermöglichen es, gleichzeitig unterschiedliche Muster zu diskriminieren. Abhängig von den Orten der Sender und von den Eigenschaften der Sendesignale treten die zugeordneten Ortsassoziationen und Gestaltassoziationen auf. Wenn davon ausgegangen wird, daß jede Selektionsstufe gleichzeitig wenigstens zwei Muster diskriminieren kann, ergibt sich ein prinzipieller Ansatz, um auch den Cocktailparty-Effekt mit Hilfe des Assoziationsmodells erklären zu können.

Die Phantomschallquelle verschwindet, sobald die beiden Lautsprechersignale hinreichend unähnlich werden, das heißt sobald sie unterschiedliche Hörereignisgestalten hervorrufen. Dies ist der allgemeine, natürliche Fall eines überlagerten Schallfelds.

des. Zwei Schallquellen rufen normalerweise nicht nur zwei verschiedene Ortsassoziationen, sondern zusätzlich zwei verschiedene Gestaltassoziationen hervor. Die resultierenden beiden Hörereignisse treten dann also nach zweistufiger Selektion auf, woraus sich die größtmögliche Auflösung ableitet. Trägt nur die Gestaltassoziationsstufe zur Hörereignisauflösung bei, weil beide Lautsprecher sich am gleichen Ort befinden, so geht die Hörereignisauflösung auf einen Grenzwert zurück (Grenzauflösung der Gestaltassoziationsstufe, [8]). Man erkennt auch hier die Bedeutung der Gestaltassoziationsstufe für das räumliche Hören: Simultane Hörereignisse können nur auftreten, wenn die auslösenden Reizmuster zusätzlich verschiedene Hörereignisgestalten hervorufen.

Umgekehrt ergibt sich der Cocktailparty-Effekt als Wirkung der Ortsassoziationsstufe; simultane Hörereignisgestalten können unter bestimmten Bedingungen erst auftreten, wenn die auslösenden Reizmuster zusätzlich als Lokalisationsreize wirken. Beispielsweise kann aus einem ausreichend vielstimmigen Stimmengewirr nicht ein bestimmter Sprecher „herausgehört“ werden, wenn ein Ohr verschlossen wird, weil damit die Ortsassoziationsstufe keinen Selektionsbeitrag mehr liefern kann. Die Lokalisationsreizselektion ist ein sinnvoller Vorselektionsprozeß, der der höhergelagerten Mustererkennung vorangeht. Jede der beiden Verarbeitungsstufen selektiert die Muster hinsichtlich unterschiedlicher, voneinander unabhängiger Merkmale. Die resultierende Auflösung der unterschiedlichen Muster gewährleistet dann, daß das Gehör einzelne Hörereignisse unterscheiden kann. Dabei könnte die Konzentration auf ein Hörereignis dazu führen, daß infolge einer entsprechenden Assoziationslenkung die Auflösung eines bestimmten Musters weiter erhöht wird (Assoziationslenkung durch das Bewußtsein).

Die Lokalisation im überlagerten Schallfeld ist nach dieser Darstellung ein zweistufiger Prozeß, der aufgrund der Hörerfahrung die Hörereigniszuordnung leistet. Eine grundlegende Hörerfahrung lautet: Die empfangenen Ohrsignale sind zurückzuführen auf einen bestimmten Sendeort und auf ein bestimmtes Sendesignal. Beide Sendereigenschaften verhalten sich voneinander unabhängig; sie treten nur paarweise auf. Das Assoziationsmodell ist dieser Hörerfahrung vollständig angepaßt. Dies ist ein Grund für die weitreichende Aussagekraft.

3.4. Lokalisation oder Lateralisation?

Binaurale Ohrsignale oder Ohrsignalanteile, die keinen Lokalisationsreiz darstellen, sondern beliebig miteinander verkoppelt sind, unterliegen nicht den Gesetzmäßigkeiten einer ortsbestimmenden Verarbeitungsstufe. Dieser Sachverhalt ist besonders bei Lateralisationsexperimenten gegeben. Er läßt sich noch deutlicher darstellen, wenn die Wirkung der beiden Verarbeitungsstufen im Assoziationsmodell betrachtet wird:

Die Ortsassoziationsstufe ist der Gestaltassoziationsstufe vorgeschaltet. Damit erhält das Gehör die Eigenschaft, aus einem aktuellen Reizangebot vorzugsweise Lokalisationsreize zu selektieren. Die Lokalisationsreizselektion wirkt im überlagerten Schall-

feld als Filter zur Diskrimination einzelner Sendesignale, so daß prinzipiell erst in der Gestaltassoziationsstufe eine etwaige Verkopplung der Sendesignale zum Tragen kommt. In der Phantomschallquellsituation beispielsweise verhindert die Ortsassoziationsstufe Summenlokalisation; die Lokalisationsreizselektion bewirkt, daß trotz der Schallfeldüberlagerung in der Gestaltassoziationsstufe die gleichen Signale verarbeitet werden wie bei Kopfhörerbeschallung.

Lateralisationsexperimente geben lediglich Auskunft über die Funktion der Gestaltassoziationsstufe, weil die beiden Sendesignale unabhängig von der Senderentfernung einzeln diskriminiert und der Gestaltassoziationsstufe zugeführt werden. Lateralisationsexperimente lassen deshalb grundsätzlich keine Rückschlüsse zu auf die Funktion der Ortsassoziationsstufe; sie führen nur auf Gesetzmäßigkeiten der Phantomschallquelle („Phantomschallquelle im Kopf“). Eine „Ersatzschallquelle im Kopf“ ist nicht vorhanden (bei adaptiertem Gehör, vgl. Plenge [28]).

Die Relevanz von Lateralisationsexperimenten für das räumliche Hören ist bisher unbekannt. Man nimmt an, daß der Lokalisationsprozeß in Teilaspekte zerlegbar ist, welche sich mit Hilfe von Kopfhörern isoliert untersuchen lassen. Stellvertretend für das in der Literatur zu findende Verständnis hierzu sei Blauert [15] zitiert:

„Die Entstehung von Hörereignissen in seitlichen Richtungen setzt unterschiedliche Ohrsignale voraus. Ohrsignalunterschiede spezifischer Art entstehen infolge Beugung, Abschattung und Resonanzen an Kopf und Außenohren. Durch Lateralisationsversuche mit Kopfhörern kann man diese sehr differenzierten Signalunterschiede auf ihre „wirksamen“ Bestandteile untersuchen.“

Speziell im Zusammenhang zur Summenlokalisation heißt es:

„Es liegt nahe, den Einfluß der einzelnen Signalanteile auf die Hörereignisrichtung näher zu untersuchen. Dies kann z. B. so geschehen, daß man synthetisch erzeugte Impulsgruppen über Kopfhörer darbietet. Die einzelnen Signalanteile sind dann beliebig variierbar.“

Man geht davon aus, daß die „Auswertung unterschiedlicher Ohrsignale“ (Kapitel 2.4 in [15]), die das Gehör bei der Lokalisation einer Schallquelle vornimmt, sich mit zwei hinreichend nahe an den Ohren befindlichen Schallquellen untersuchen läßt. Man beachtet zwar die resultierende Im- oder Am-Kopf-Ortung formal, indem das Experiment als Lateralisationsexperiment, also als „Experiment mit mangelhafter Entfernungswahrnehmung“ ausgewiesen wird. Es gibt jedoch keinen Nachweis für die in der Literatur bisher gültige Annahme, daß die Ergebnisse aus Lateralisationsexperimenten – und die daraus abgeleiteten Hypothesen bezüglich der Vorgänge bei der Bildung von Hörereignissen in seitlichen Richtungen – sich „auch auf das räumliche Hören im freien Schallfeld verallgemeinern lassen“ [15].

3.5. Die Im-Kopf-Ortung

Für die stereofone Wiedergabe über Kopfhörer hat das Phänomen der Im-Kopf-Ortung eine überragende Bedeutung. Bezüglich des Hörereignisortes hat bereits Plenge [9] ein Modell beschrieben, das vermutlich eine umfassende Erklärung der Im-Kopf-Ortung zuläßt. Es enthält einen Kurzzeit- und Langzeitpeicher und einen Reizverarbeitungsmechanis-

mus, welcher gespeicherte Reizmuster mit aktuellen Reizmustern vergleicht. Die Reizverarbeitung führt dann nicht zur Lokalisation (außerhalb des Kopfes), wenn entweder der Kurzzeitspeicher keine oder (noch) falsche Information über Schallquelle und Darbietungsort enthält, oder wenn die Reize so geartet sind, daß „sie keinem im Langzeitspeicher enthaltenen Reizmuster zugeordnet werden können“. Bemerkenswert daran ist, daß die Lokalisation auf einen Lernprozeß zurückgeführt wird und dabei sowohl dem Kurzzeitgedächtnis als auch dem Langzeitgedächtnis eine spezielle Bedeutung zukommt.

Der Ansatz „Reizverarbeitung durch Vergleich mit gespeicherten Reizmustern“ [9] ist im Assoziationsmodell enthalten, die beschriebene „assoziativ gesteuerte Mustererkennung“ der Ortsassoziationsstufe stellt dessen Weiterentwicklung dar. Die Vorhersagen bezüglich des Hörereignisortes stimmen deshalb überein. Bezüglich der Hörereignisklangfarbe dagegen führt das Assoziationsmodell zu weiteren Konsequenzen, weil zusätzlich die ortsbestimmende Verarbeitungsstufe die diskriminierten Sendesignale an eine gestaltbestimmende Stufe weiterleitet. Hieraus erklärt sich vor allem die Unabhängigkeit der Hörereignisklangfarbe von der Richtung **und der Entfernung** der Lautsprecher.

Diese Aussage gilt auch für die Entfernung Null, das heißt also für Kopfhörerbeschallung. Dazu sind folgende Bemerkungen zu machen:

- Die Schall- oder Hörereignisentfernung Null liegt am Kopf. Im Gegensatz zu Laws [21] u. a., die die Entfernung vom Kopfmittelpunkt (Koordinatenursprung) aus messen, kann im Sinne des Assoziationsmodells die Entfernung Null nur an den Begrenzungsflächen des Empfängers vorliegen. Ein Schallereignis im Kopf gibt es nicht. Ein Hörereignis im Kopf kann deshalb auch nicht aufgrund eines Lokalisationsprozesses auftreten; es widerspricht der Hörerfahrung. Ein „Hörereignis ohne Entfernung“ kennzeichnet die Im-Kopf-Ortung.
- Die Hörereignisklangfarbe verhält sich natürlich nur insoweit ortsunabhängig, wie dies vom Leistungsdichtespektrum des Schalldrucks am Hörort gewährleistet ist. Insbesondere sind für den Übergang der Lautsprecherentfernung gegen Null die Schalldruckverhältnisse im Nahfeld zu berücksichtigen („Nahfeldverzerrung“). Allerdings ist bisher nicht untersucht worden, ob und unter welchen Bedingungen dieser Einfluß vom Gehör ebenfalls erlernt und demzufolge decodiert wird.
- Die Decodierung der Entfernung der Lautsprecher (bzw. Kopfhörerkapseln) geschieht nur unter zwei Voraussetzungen:
 1. Die Lokalisationsbedingung „breitbandiges Spektrum“ muß eingehalten werden (vgl. 3.6.).
 2. Eine Ersatzschallquelle darf nicht existieren (keine kopfbezogenen Signale bzw. keine Kopfhörer mit „lautsprecher-simulierender Entfernung“ [21, 57, 58]).

Die „mangelhafte Entfernungswahrnehmung“ bei Kopfhörerwiedergabe erklärt das Assoziationsmodell mit der „Lautsprecherwiedergabe Null“; auch die

Schallquellenorte „an den Ohren“ verursachen im Prinzip zwei Lokalisationsreize, die entsprechende Ortsassoziationen auslösen. Hieraus ergeben sich Folgerungen für die Kopfhörerentzerrung (vgl. 4.). Weiterhin ergeben sich Folgerungen für Untersuchungen der Gehörfunktionen bei der Lokalisation: Die Auswertung unterschiedlicher Ohrsignale, die das Gehör bei der Lokalisation einer Schallquelle vornimmt, kann grundsätzlich **nicht** im Lateralisationsexperiment, sondern nur unter Lokalisationsbedingungen untersucht werden [8].

3.6. Schmalbandige Signale

Eng verbunden mit dem Effekt „mangelnde Entfernungswahrnehmung“ ergibt sich eine weitere wichtige Konsequenz: Die geforderte Lokalisationsbedingung wird nicht eingehalten, wenn das Schallereignis ein schmalbandiges Spektrum aufweist. Denn der Wahrnehmungsprozeß, der zur Lokalisation führt, ist nur möglich, wenn spektrale Merkmale die Zuordnung der Hörereignisentfernung zulassen.

Eine Unterschreitung der minimalen Signalbandbreite hat beispielsweise in der Phantomschallquellsituation zur Folge, daß keine Lokalisationsreizelektion stattfindet. In diesem Fall trifft die Bezeichnung „Summenlokalisierung“ den Sachverhalt. Die interauralen Summensignale sind hinsichtlich ihrer „Summanden“ nicht mehr zerlegbar (vgl. „binaurales Korrelationsmuster“ in [8]). Die resultierenden interauralen Phasen- und Intensitätsbeziehungen verursachen wie bei der Lateralisation lediglich seitliche Auslenkungen. Diese allerdings können durch entsprechende Voreingenommenheit von der Versuchsperson leicht als Richtung des Hörereignisses interpretiert werden, z. B. infolge optischer Eindrücke [9, 29].

Ein weiteres Beispiel stellt der Franssen-Effekt dar [59]. Der Lautsprecher A wird mit einem Ton von einigen Sekunden Dauer angesteuert, der mit exponentieller Hüllkurve an- und abklingt [Signal $u_A(t)$]. Lautsprecher B erhält ein Signal $u_B(t)$, das so beschaffen ist, daß die Summe $u_A(t) + u_B(t)$ einen Ton mit rechteckiger Hüllkurve ergeben würde. Lautsprecher B gibt also den (breitbandigen) Einschaltvorgang wieder, dagegen Lautsprecher A den von Schaltvorgängen befreiten (schmalbandigen) Ton. Es zeigt sich, daß der Hörereignisort ausschließlich vom Einschaltvorgang bestimmt ist. Das stimmt mit dem Assoziationsmodell überein, wenn man annimmt, daß der Ton keinen Lokalisationsreiz darstellt, sondern wegen fehlender Entfernungsinformation dem vorhandenen Lokalisationsreiz zugeordnet wird. Entsprechend hat Wendt [60] festgestellt, daß hartgeschaltete Töne sich bei Stereophonie ebenso verhalten wie breitbandige Signale. Sie stehen mit den Hörereignisorten, die er für Dauertöne und Tonimpulse ermittelt hat, in keinem erkennbaren Zusammenhang.

Äquivalent dazu sind auch neuere Untersuchungsergebnisse von Blauert/Cobben [45] zu interpretieren, die die Hörereignisrichtung für schmalbandgefilterte und breitbandige Impulse in Abhängigkeit der Laufzeitdifferenz entsprechend **Bild 5** gemessen haben. Die Lokalisationskurven der Schmalbandsignale decken sich weder im Bereich der Laufzeit-

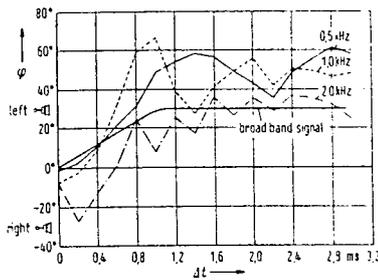


Bild 5

Höreignisrichtung φ als Funktion der Verzögerung Δt des rechten Lautsprechersignals (entnommen aus [45])

stereofonie noch bei größeren Laufzeitdifferenzen mit der Kurve, die für das breitbandige Signal gemessen wurde. Man erkennt an der Periodizität dieser Kurven, daß im ganzen Verzögerungsbereich Summenlokalisation stattgefunden hat, daß also die resultierenden interauralen Signaldifferenzen die Höreignisrichtung bestimmt haben. Auch das Gesetz der ersten Wellenfront läßt sich in Übereinstimmung mit dem Assoziationsmodell mit Schmalbandsignalen nicht untersuchen.

Hörversuche, bei denen die Lokalisationsbedingungen nicht eingehalten werden, lassen grundsätzlich keine Rückschlüsse zu auf die Funktion des Gehörs bei der Lokalisation einer Einzelschallquelle.

4. Konsequenzen für die elektroakustische Übertragung

Die Wirkung der Ortsassoziationsstufe im Modell kann mit einem adaptiven Filter beschrieben werden, dessen Übertragungsfunktion M^{-1} sich invers verhält zur ortsabhängigen Übertragungsfunktion M des Außenohres (vgl. Bild 1). Dieser Sachverhalt ermöglicht die Darstellung der gesamten „psychoakustischen Funktionskette“ für den Fall einer elektroakustischen Übertragung. Kennzeichnend ist, daß die Funktion des Gehörs bei der Lokalisation im Sinne der vorangegangenen Darlegungen einen Bestandteil der Übertragungsstrecke darstellt. Anhand dieser Darstellungsform können dann psychoakustische Zusammenhänge unter den unterschiedlichen Übertragungsverfahren leicht erkannt und beschrieben werden.

Bild 6 gibt die Funktionskette wieder. Eingetragen sind die Bezeichnungen der Übertragungsfunktionen sowie der dazugehörigen Ein- und Ausgangssignale. Es bedeuten

- q = Schallereigniseigenschaft,
- q' = Höreigniseigenschaft,
- a = Signal am Ausgang des Aufnahmemikrofons,
- b = zu übertragendes Signal,
- c = vom Lautsprecher bzw. Kopfhörer abgestrahltes Signal,
- e = nach der räumlichen Decodierung diskriminiertes Signal;
- A = ortsabhängige Übertragungsfunktion des Mikrofons,

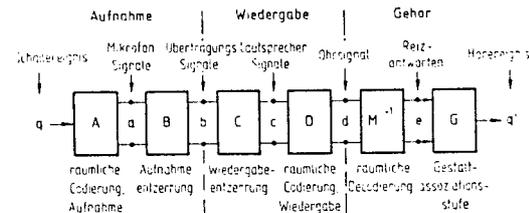


Bild 6

Schema der „psychoakustischen Funktionskette“

- B = Übertragungsfunktion der Mikrofonentzerrung,
- C = Übertragungsfunktion der Lautsprecher- bzw. Kopfhörerentzerrung,
- D = ortsabhängige Übertragungsfunktion des Außenohres,
- M^{-1} = Übertragungsfunktion des adaptiven Filters zur räumlichen Decodierung,
- G = Wirkung der Gestaltassoziationsstufe.

Die allgemeine Übertragungsgleichung lautet

$$q' = A \cdot B \cdot C \cdot D \cdot M^{-1} \cdot G \cdot q$$

Für das natürliche Hören gilt beispielsweise:

- Die elektroakustischen Übertragungsglieder fallen fort: $A = B = C = 1$
- Selektierbare Lokalisationsreize erzeugt D (natürliches Hören!): $D \cdot M^{-1} = 1$
- Ergebnis: $q' = G \cdot q$

Die Beziehung $q' = G \cdot q$ soll ausdrücken, daß keine Fehler beim räumlichen Hören auftreten. Schallereignisort und Höreignisort sind identisch, jedoch ergibt sich die Höreignisgestalt über die in diesem Lokalisationsmodell nicht definierte Abbildungseigenschaft G der Gestaltassoziationsstufe.

Die Übertragungsgleichung $q' = G \cdot q$ ergibt sich erwartungsgemäß ebenfalls bei **kopfbezogener Stereofonie**:

- Die systemgerechte Entzerrung ist vorhanden, wenn für Aufnahme- und Wiedergabeseite eine identische Definition eingehalten wird: $B \cdot C = 1$
- Für Kopfhörerwiedergabe ist die ortsabhängige Übertragungsfunktion des Außenohres eliminiert: $D = 1$
- Aus $B \cdot C \cdot D = 1$ folgt für die Ohrsignale $d = A_K \cdot q = a$. A_K erzeugt einen selektierbaren Lokalisationsreiz (Kunstkopf!): $A_K \cdot M^{-1} = 1$
- Ergebnis: $q' = G \cdot q$

Benutzt man anstelle des Kopfhörers Lautsprecher, die über das Tradis-Filter [55] zur Simulation der Kopfhörerwiedergabe gespeist werden, so gilt theoretisch $C \cdot D = 1$. Um ebenfalls $B \cdot C \cdot D = 1$ zu erreichen, muß $B = 1$ werden, d. h. das Tradis-Filter C muß genau abgestimmt sein auf die Entzerrung des Kunstkopfsignals.

Diese Betrachtungen lassen sich fortsetzen. Doch hier sollen mit ihrer Hilfe vor allem die Kompatibilitätsaspekte aufgezeigt werden.

Zunächst gilt für **raumbezogene Stereophonie** über Lautsprecher:

- Die Entzerrungen von Stereomikrofon und Lautsprecher sind in erster Näherung gleich (ebener Frequenzgang für Bezugsrichtung 0°): $B \cdot C = 1$
- Die selektierbaren Lokalisationsreize erzeugt D (Phantomschallquellensituation): $D \cdot M^{-1} = 1$
- Ergebnis (Index r : übliches Stereomikrofon): $q' = A_r \cdot G \cdot q$

Die gesamte Übertragungsstrecke läßt sich also beschreiben mit der Wirkung der Signale eines Stereomikrofons auf die Gestaltassoziationsstufe. Diese Wirkung ist in 3.3. beschrieben worden (Hörereignis-auslenkung zwischen den Lautsprechern infolge der Intensitäts- oder Laufzeitunterschiede); sie verursacht den erwünschten Stereophonie-Effekt.

Hinsichtlich der Kompatibilität raum/kopfbezogener Signale findet man folgende Aussagen:

1. Tritt an die Stelle des Stereomikrofons der Kunstkopf, so verhindert D die kopfbezogene Stereophonie: es bleibt $D \cdot M^{-1} = 1$; $A_K \cdot M^{-1} \neq 1$! Nur für den Fall $C \cdot D = 1$ (Einsatz des Tradis-Filters für C) wird $A_K \cdot M^{-1} = 1$.
2. Die Übertragungsgleichung $q' = A \cdot G \cdot q$ bleibt erhalten. Jedoch wirkt A_K anstelle von A_r auf G. In [5] wurde die Verknüpfung $q' = A_K \cdot G \cdot q$ ausführlich in Hinblick auf den Hörereignisort untersucht und gezeigt, daß dieser sich aus der Summe der Wirkung von Laufzeit- und Pegeldifferenz ergibt. Wie in [5] gezeigt, ist die „gemischte Stereophonie“ kompatibel zur Intensitätsstereophonie.

4.1. Die optimale Kunstkopffentzerrung

Weiterhin wurde in [6] die Übertragungsgleichung $q' = A_K \cdot G \cdot q$ in Hinblick auf die Klangfarbe untersucht. Es wird davon ausgegangen, daß durch den Fusionsprozeß in der Gestaltassoziationsstufe die Klangfarbe unbeeinflusst bleibt von den gegebenen Zeitdifferenzen der Kunstkopfsignale. Vielmehr wird angenommen, daß das klangfarbenbestimmende Gesamtspektrum sich aus den jeweils dominierenden Anteilen zusammensetzt. Damit gestaltet sich das Klangfarbenproblem sehr übersichtlich: Die Gleichung $q' = A_K \cdot G \cdot q$ besagt, daß die ortsabhängige Übertragungsfunktion A_K die Klangfarbenbestimmung durch G direkt beeinflusst. Komplexe, klangfarbenbeeinflussende „Verrechnungen“ dieser Spektren für die Lokalisation gibt es nicht.

Für die Bestimmung der optimalen Entzerrung werden die Funktionen B und C eingesetzt: Man kann jetzt nicht davon ausgehen, daß $B \cdot C = 1$ ist. Es gilt:

$$q' = A_K \cdot B_X \cdot C \cdot G \cdot q$$

In dieser Gleichung ist B_X die gesuchte Entzerrung des Kunstkopfes, C die Entzerrung der Lautsprecher. Geht man davon aus, daß die Lautsprecherentzerrung gegeben ist mit $C = C_{FF}$ (Freifeldentzerrung), dann ist ein B_X zu suchen, das sich invers verhält zu A_K . Die Forderung lautet:

$$A_K \cdot B_X = 1$$

A_K ist die **ortsabhängige** Übertragungsfunktion des Kunstkopfes, B_X dagegen die **ortsunabhängige** Übertragungsfunktion des Entzerrers.

Realisiert man eine Freifeldentzerrung der Kunstkopfsignale nach [61] (Bezugsrichtung 0° vorne), so gilt

$$A_{K \text{ vorne}} \cdot B_X = 1$$

Diese Forderung ist zwar exakt erfüllbar, doch eine Kunstkopffentzerrung der Form $A_{K \text{ vorne}}^{-1}$ berücksichtigt nicht die Ortsabhängigkeit von A_K . Gerade der Bezug „vorne“ ergibt in Hinblick auf die Gesamtheit aller Übertragungsfunktionen eine sehr ungünstige Entzerrung (8-kHz-Senke, vgl. [6]).

Man erkennt, daß für die Bestimmung von B_X nicht eine einzige Bezugsrichtung zugrunde gelegt werden kann, sondern vielmehr für die Gesamtheit der Schalleinfallrichtungen ein Optimum ermittelt werden muß. Geeignet ist beispielsweise die rechnerische Methode der Integralbildung über einen Katalog von Außenohrübertragungsfunktionen [62]. Noch günstiger ist der meßtechnische Weg dieser Integralbildung, nämlich die Bestimmung der Entzerrung, die einen ebenen Verlauf des Diffusfeld-Übertragungsmaßes gewährleistet. Die praktischen Aspekte für die Wahl und Realisierung dieser Diffusfeldentzerrung werden in [6] behandelt.

4.2. Die optimale Kopfhörerentzerrung

Wie schon gezeigt, ist die systemgerechte Entzerrung bei kopfbezogener Stereophonie vorhanden, wenn gilt: $B \cdot C = 1$. Nur dann gilt die zugehörige Übertragungsgleichung $q' = G \cdot q$. Da sich aus der Kompatibilitätsforderung ein diffusfeldentzerrter Kunstkopf ergibt (Übertragungsfunktion B_{DF}), muß für die Wiedergabe kopfbezogener Signale ein Kopfhörer mit Diffusfeldentzerrung C_{DF} benutzt werden.

Welche Kopfhörerentzerrung ist optimal für die Wiedergabe raumbezogener Signale? Die Lösung findet sich bei Betrachtung der Übertragungskette sehr einfach, indem die beiden Fälle „natürliches Hören“ und „Kopfhörerwiedergabe raumbezogener Signale“ gegenübergestellt werden (es sei $A = B = C = 1$):

$$\text{Natürliches Hören: } q'_N = D_N \cdot M^{-1} \cdot G \cdot q$$

$$\text{Kopfhörerwiedergabe: } q'_K = C_X \cdot D_K \cdot M^{-1} \cdot G \cdot q$$

Beim natürlichen Hören ist infolge der Lokalisationsreizselektion $D_N \cdot M^{-1} = 1$ und somit $q'_N = G \cdot q$.

Auch bei Kopfhörerwiedergabe von **nicht** kopfbezogenen Signalen adaptiert M^{-1} , und zwar auf die Verhältnisse für „Lautsprecherentfernung Null“ (vgl. 3.5.). Dies führt zur Im-Kopf-Ortung in Verbindung mit G. Für die Klangfarbenbestimmung ist wegen $D = 1$ deshalb auch M^{-1} ohne Einfluß. Es gilt daher $D_K \cdot M^{-1} = 1$ und somit $q'_K = C_X \cdot G \cdot q$.

In beiden Darbietungsfällen ist also die Klangfarbe konstant, auch beim natürlichen Hören unabhängig von der Schalleinfallrichtung. Es ist infolge der adaptiven Funktion von M^{-1} nicht möglich, durch Gleichsetzen der beiden Gleichungen C_X zu bestimmen. Doch es gibt die Möglichkeit, die Wirkung von M^{-1} zu unterdrücken: Die Lokalisationsreize sind zerstört, wenn ein Ohr geschlossen ist; eine Klangverfärbung durch die Wirkung der Außenohrübertragungsfunktion ist wahrnehmbar. In diesem Fall gilt damit:

$$q'_N = D_N \cdot G \cdot q$$

Aus der geforderten Klangfarbenäquivalenz beim einohrigen Hören folgt also mit $q'_N = q'_K$ die Forderung

$$C_X = D_N$$

Analog zur optimalen Kunstkopffentzerrung ergibt sich die optimale Kopfhörerentzerrung durch Integration von D_N über alle Schalleinfallrichtungen.

Ebenso wie für die Wiedergabe kopfbezogener Signale über Lautsprecher ist für die Wiedergabe raumbezogener Signale über Kopfhörer also eine Entzerrung optimal, die sich auf das Diffusfeld bezieht. Der Bezug auf eine bestimmte Schalleinfallrichtung im Freifeld ist ungünstig. Insbesondere der Bezug auf die Schalleinfallrichtung „vorne“ führt dazu, daß die „8-kHz-Senke“ die Klangfarbe beeinträchtigt [6].

Dieses Ergebnis bestätigt die Erfahrungen, die man mit der zur Zeit üblichen Freifeldentzerrung nach DIN 45 619 gemacht hat [6].

Allgemein gilt, daß für die Entzerrung der Schnittstelle zwischen Aufnahme- und Wiedergabeseite die Gesamtheit aller Übertragungsfunktionen des Außenohres durch eine entsprechende mittlere Übertragungsfunktion zu ersetzen ist, sobald kopfbezogene Signale über Lautsprecher oder raumbezogene Signale über Kopfhörer abgehört werden sollen. In beiden Übertragungsfällen kann das Gehör die vorhandenen spektralen Ohrsignalmerkmale, die sich aus der aufnahme- oder wiedergabeseitigen räumlichen Codierung ableiten, nicht decodieren. Infolgedessen gelangen die spektralen Merkmale „unverrechnet“ an die Gestaltassoziationsstufe, sie beeinflussen die Klangfarbe.

5. Ergebnis

Das Assoziationsmodell erweitert das bisherige Verständnis über die Funktion des Gehörs durch zwei bedeutsame Eigenschaften:

1. Die Hörereigniseigenschaften resultieren aus der Wirkung von zwei unterschiedlichen Verarbeitungsprozessen, einer für die Selektion des Ortes und ein zweiter, höher gelagerter Prozeß für die Selektion der Gestalt.
2. Die ortsbestimmende Verarbeitungsstufe hat die Eigenschaft, die räumliche Information des ankommenden Signals mit Hilfe eines ortsangepaßten adaptiven Filters zu decodieren, darüber hinaus aber die Gestaltinformation von der Ortsinformation zu trennen und der gestaltbestimmenden Stufe zuzuführen.

Die adaptive und diskriminierende Funktion der ortsbestimmenden Verarbeitungsstufe erklärt die Phänomene des räumlichen Hörens einheitlich. Der Ansatz hat Konsequenzen sowohl für zukünftige Untersuchungen als auch auf dem Gebiet der elektroakustischen Übertragung.

Insbesondere sagt das Modell den Zusammenhang zwischen Hörereignisort und Klangfarbe voraus. Daraus läßt sich eine Definition für die optimale Entzerrung von Kunstköpfen und Kopfhörern ableiten. Als „Universalschnittstelle“ zwischen Aufnahme- und Wiedergabeseite wird eine Entzerrung vorgeschlagen, die sich grundsätzlich auf das Diffusfeld bezieht.

Der Verfasser dankt den Prof. Dres. J. Blauert, G. Boerger, L. Cremer, G. Plenge sowie Dr. P. Schöne für wertvolle Anregungen und Diskussionen.

SCHRIFTTUM

- [1] Schlemm, W.: Kunstkopf-Stereophonie — eine Alternative zur Quadrofonie? *Funkschau* 46 (1974), S. 771 bis 774.
- [2] Steinke, G.: Möglichkeiten und Grenzen der kopfbezogenen Stereophonie. Teil 1 und 2. *radio fernsehen elektronik* 28 (1979), H. 6 und 7, S. 388 bis 392 und S. 456 bis 466.
- [3] Theile, G.: „Kunstkopf“ Stereophony: Results of recent investigations and development of a new Kunstkopf system. IBC 80, Brighton. Conf. Publ. No. 191, S. 350 bis 353.
- [4] Hudde, H.; Schröter, J.: Verbesserungen am Neumann-Kunstkopfsystem. *Rundfunktech. Mitt.* 25 (1981), S. 1 bis 6.
- [5] Theile, G.: Zur Kompatibilität von Kunstkopfsignalen mit intensitätsstereofonen Signalen bei Lautsprecherwiedergabe: Die Richtungsabbildung. *Rundfunktech. Mitt.* 25 (1981), S. 67 bis 73.
- [6] Theile, G.: Zur Kompatibilität von Kunstkopfsignalen mit intensitätsstereofonen Signalen bei Lautsprecherwiedergabe: Die Klangfarbe. *Rundfunktech. Mitt.* 25 (1981), in diesem Heft.
- [7] Wendt, K.: Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie. Dissertation. Technische Hochschule Aachen, 1963.
- [8] Theile, G.: Über die Lokalisation im überlagerten Schallfeld. Dissertation. Technische Universität Berlin, 1980.
- [9] Plenge, G.: Über das Problem der intracranialen Ortung von Schallquellen bei der akustischen Wahrnehmung des Menschen. Habilitationsschrift. Technische Universität Berlin, 1973.
- [10] Cremer, L.: Die wissenschaftlichen Grundlagen der Raumakustik. Band I. S. Hirzel Verlag, Stuttgart 1948.
- [11] Cherry, E. C.: Some experiments on the recognition of speech, with one and with two ears. *J. of Acoust. Soc. Am.* 25 (1953), S. 975 bis 979.
- [12] Cherry, E. C.; Taylor, W. K.: Some further experiments upon recognition of speech, with one and with two ears. *J. of the Acoust. Soc. Am.* 26 (1954), S. 554 bis 559.
- [13] Platte, H.-J.: Zur Bedeutung der Außenohrübertragungseigenschaften für den Nachrichtenempfänger „menschliches Gehör“. Dissertation. Technische Hochschule Aachen, 1979.
- [14] Plenge, G.; Tilse, U.: Cocktail-Party-Effekt with and without conflicting visual cues. 50th AES Convention, London 1975.
- [15] Blauert, J.: Räumliches Hören. S. Hirzel Verlag, Stuttgart 1974.
- [16] Blauert, J.: Untersuchungen zum Richtungshören in der Medianebene bei fixiertem Kopf. Dissertation. Technische Hochschule Aachen, 1969.
- [17] Marko, H.: Ein Funktionsmodell für die Aufnahme, Speicherung und Erzeugung von Informationen im Nervensystem. *Röntgen-Blätter* 24 (Nov. 1971).
- [18] Keidel, W. D.: Informationsverarbeitung. In: *Kurzgefaßtes Lehrbuch der Physiologie*. Hrsg. v. W. D. Keidel. Georg Thieme Verlag, Stuttgart 1973, S. 385 bis 401.
- [19] Engel, W.: Eine neue Methode zur Entkopplung mehrfach geregelter Systeme. Dissertation. Technische Hochschule München, 1966.
- [20] Theile, G.: Weshalb ist der Kammfiltereffekt bei Summenlokalisation nicht hörbar? Vortrag auf der 11. Tonmeistertagung in Berlin, Nov. 1978.
- [21] Laws, P.: Zum Problem des Entfernungshörens und der Im-Kopf-Lokalisiertheit von Hörereignissen. Dissertation. Technische Hochschule Aachen, 1972.
- [22] Roesser, D.: Schallrichtungsbestimmung bei krankhaft verändertem Gehör. Dissertation. Technische Hochschule Aachen, 1965.
- [23] Klemm, O.: Lokalisation von Sinneseindrücken bei disparaten Nebenreizen. *Psychol. Stud.* 5 (1909), S. 73 bis 162.
- [24] Blauert, J.: Ein Versuch zum Richtungshören bei gleichzeitiger optischer Stimulation. *Acustica* 23 (1970), S. 118 bis 119.
- [25] Massaro, D. W.; Warner, D. S.: Dividing attention between auditory and visual perception. *Perception & Psychophysics* 21 (1977), S. 569 bis 573.

- [26] Lehninger, F. J.: Über die Beeinflussung der Hörereignisrichtung durch das Schereignis bei der Personendarstellung in Telekonferenzen. *Acustica* 42 (1979), S. 240 bis 248.
- [27] Plenge, G.: Über das Problem der Im-Kopf-Lokalisation. *Acustica* 26 (1972), S. 241 bis 252.
- [28] Plenge, G.: On the differences between localization and lateralization. *J. of the Acoust. Soc. Am.* 56 (1974), S. 944 bis 951.
- [29] Jeffress, L. A.; Taylor, R. W.: Lateralization versus localization. *J. of the Acoust. Soc. Am.* 33 (1961), S. 482 bis 483.
- [30] Flechtner, H. J.: Grundbegriffe der Kybernetik. S. Hirzel Verlag, Stuttgart 1972.
- [31] Caspers, H.: Zentralnervensystem. In: Kurzgefaßtes Lehrbuch der Physiologie. Hrsg. v. W. D. Keidel. Georg Thieme Verlag, Stuttgart 1973, S. 467 bis 524.
- [32] Fukushima, K.: A model of associative memory in the brain. *Kybernetik* 12 (1973), S. 58 bis 63.
- [33] Poggio, T.: On optimal nonlinear associative recall. *Biol. Cybernetics* 19 (1975), S. 201 bis 209.
- [34] Kohonen, T.; Reukala, E.; Maekisara, K.; Vainio, L.: Associative recall of images. *Biol. Cybernetics* 21 (1976), S. 159 bis 168.
- [35] Kohonen, T.; Oja, E.: Fast adaptive formation of orthogonalizing filters and associative memory in recurrent network of neuron-like elements. *Biol. Cybernetics* 21 (1976), S. 85 bis 95.
- [36] Kohonen, T.: Associative memory. Springer-Verlag, Berlin-Heidelberg-New York 1977.
- [37] Wess, O.; Röder, U.: A holographic model for associative memory chains. *Biol. Cybernetics* 27 (1977), S. 89 bis 98.
- [38] Bohn, G.: A structure for associative information processing. *Biol. Cybernetics* 29 (1978), S. 193 bis 200.
- [39] Fukushima, K.; Miyake, S.: A self-organizing neural network with a function of associative memory: Feed-back-type cognitron. *Biol. Cybernetics* 28 (1978), S. 201 bis 208.
- [40] Murakami, K.; Akaishi, S.; Aibara, T.: On optimal associative recall by an incomplete key. *Biol. Cybernetics* 30 (1978), S. 95 bis 97.
- [41] Wigström, H.: A model of a neural-network with recurrent inhibition. *Kybernetik* 16 (1974), S. 103 bis 112.
- [42] Willwacher, G.: Fähigkeiten eines assoziativen Speichersystems im Vergleich zu Gehirnfunktionen. *Biol. Cybernetics* 24 (1976), S. 181 bis 198.
- [43] Zwicker, E.; Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. S. Hirzel Verlag, Stuttgart 1967.
- [44] Duifhuis, H.: Perceptual analysis of sound. Dissertation. Technische Hochschule Eindhoven, 1972.
- [45] Blauert, J.; Cobben, W.: Some considerations of binaural cross correlation analysis. *Acustica* 39 (1978), S. 96 bis 104.
- [46] Schenkel, K. D.: Accumulation theory of binaural-masked thresholds. *J. of the Acoust. Soc. Am.* 41 (1967), S. 20 bis 31.
- [47] Durlach, N. I.: Equalization and cancellation theory of binaural masking-level differences. *J. of the Acoust. Soc. Am.* 35 (1963), S. 1206 bis 1218.
- [48] Osman, E.: A correlation model of binaural masking level differences. *J. of the Acoust. Soc. Am.* 50 (1971), S. 1494 bis 1511.
- [49] Terhardt, E.: Zur Tonhöhenwahrnehmung von Klängen. Teil I und II. *Acustica* 26 (1972), S. 173 bis 186 und S. 187 bis 199.
- [50] Warncke, H.: Die Grundlagen der raumbezüglichen stereophonischen Übertragung im Tonfilm. *Akust. Z.* 6 (1941), S. 174 bis 188.
- [51] Leakey, D. M.: Some measurements on the effects of interchannel intensity and time differences in two channel sound systems. *J. of the Acoust. Soc. Am.* 31 (1959), S. 977 bis 986.
- [52] Franssen, N. V.: Stereophonie. Philips Tech. Bibl., Eindhoven 1963.
- [53] Makita, Y.: On the directional localisation of sound in the stereophonic sound field. *E.B.U. Rev. Part A Tech. No. 73*, June 1962, S. 102 bis 108.
- [54] Mertens, H.: Directional hearing in stereophony — Theory and experimental verification. *E.B.U. Rev. Part A Tech. No. 92*, August 1965, S. 146 bis 158.
- [55] Damaske, P.; Mellert, V.: Ein Verfahren zur richtungstreuen Schallabbildung des oberen Halbraumes über zwei Lautsprecher. *Acustica* 22 (1969/70), S. 153 bis 162.
- [56] Gardner, M. B.: Image fusion, broadening and displacement in sound location. *J. of the Acoust. Soc. Am.* 46 (1969), S. 339 bis 349.
- [57] Blauert, J.; Laws, P.: Verfahren zur orts- und klanggetreuen Simulation von Lautsprecherbeschallungen mit Hilfe von Kopfhörern. *Acustica* 29 (1973), S. 273 bis 277.
- [58] Wischgoll, K. J.: Kunstköpfe hören und sprechen. *radio mentor electronic* 40 (1974), S. 276 bis 278.
- [59] Franssen, N. V.: Some considerations of the mechanism of direction hearing. Dissertation. Technische Hochschule Delft, 1960.
- [60] Wendt, K.: Das Richtungshören bei Zweikanal-Stereophonie. *Rundfunktech. Mitt.* 8 (1964), S. 171 bis 179.
- [61] Platte, H.-J.; Laws, P.: Technische Probleme beim Einsatz kopfbezogener stereofoner Übertragungsverfahren. *Rundfunktech. Mitt.* 22 (1978), S. 22 bis 27.
- [62] Schröter, J.: Katalog von Außenrohrübertragungsfunktionen. Ruhruniversität Bochum, 1980.