

V. Psychophysik des normalen und gestörten Gehörs

V.1 Wahrnehmungs-Grundgrößen

Das menschliche Gehör ist in faszinierender Weise optimal an das Empfangen von akustischen Signalen aus der Umwelt angepaßt, insbesondere an das Verstehen von Sprache. Weil das akustische Sprachsignal sich zeitlich stark ändert und unterschiedliche Frequenzanteile aufweist, benötigt das Ohr für die Sprachwahrnehmung die Fähigkeit, zu jedem Zeitpunkt die Intensität wahrzunehmen, mit der jede Frequenz momentan vorliegt. Wesentliche Grundgrößen der Wahrnehmung sind daher die Umsetzung verschiedener Schallintensitäten in subjektiv empfundene **Lautheit**, die Umsetzung verschiedener Frequenzen in subjektiv empfundene **Tonhöhen**, die Umsetzung verschiedener Zeitdauern und Rhythmen in subjektiv empfundene **Zeitmuster**, die Umsetzung von akustischen Signalen in subjektiv empfundene **Klänge**, und das Trennen verschiedener Klänge (z. B. Nutzsignal vom störenden Hintergrundsignal). Diese Beziehungen zwischen dem akustischen Reiz und der subjektiven Wahrnehmung, die durch psychologische Meßmethoden erfaßt werden können, werden in der **Psychoakustik** untersucht. Im folgenden soll die Psychoakustik dieser Wahrnehmungs-Grundgrößen näher betrachtet werden.

V.1.1 Intensitätsabbildung (Lautheit) und Intensitätsauflösung

Als **Ruhehörschwelle** wird der Pegel eines soeben wahrnehmbaren Sinustons bei einer vorgegebenen Frequenz in Ruhe (d. h. ohne Vorliegen störender Signale) bezeichnet. Bei niedrigen und bei hohen Frequenzen muß ein wesentlich höherer Schalldruckpegel erzeugt werden als bei mittleren, damit der Ton von normalhörenden Menschen schwellenhaft wahrgenommen werden kann (vgl. Abbildung 5.1). Wenn ein Sinus-Ton im Pegel kontinuierlich erhöht wird, nimmt die Versuchsperson nach Überschreiten der Ruhehörschwelle einen immer lauterem Ton wahr. Der Grad, in dem diese Lautheitswahrnehmung mit zunehmenden Tonpegel ansteigt, hängt von der Frequenz ab (vgl. Abbildung 5.1): Ausgehend von einem Sinus-Ton bei 1000 Hz kann man bei jeder Frequenz denjenigen Tonpegel bestimmen, der zum gleichen Lautheitseindruck führt. Diese Kurve wird als **Isophone** oder Kurve gleicher Pegellautstärke bezeichnet. Während die 10 Phon-Isophone bei 1000 Hz genau 10 dB über der Ruhehörschwelle ist und zu niedrigen und hohen Frequenzen ähnlich steil ansteigt wie die Ruhe-Hörschwelle, flachen die zu höheren Pegeln gehörenden Isophonen zunehmend ab, so daß die 120 Phon-Isophone,

die ungefähr der Schmerzschwelle bei Normalhörenden für alle Frequenzen entspricht, etwa den gleichen Schalldruck für alle Frequenzen angibt.

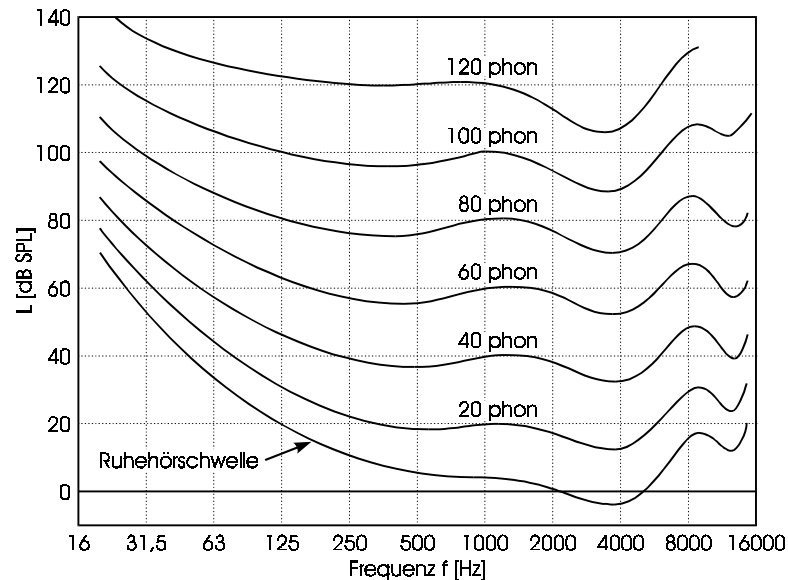


Abbildung 5.1.1: Ruhehörschwelle und Isophonen

Um den Lautheitsanstieg mit zunehmendem Pegel zu bestimmen, d. h. beispielsweise die Pegelerhöhung, die für eine Verdopplung der empfundenen Lautheit notwendig ist, bedient man sich verschiedener subjektiver Skalierungsmethoden. Bei der **freien Größenschätzung** nach Stevens und Mitarbeitern (Stevens, 1957) soll die Versuchsperson die jeweils empfundene Lautheit mit einer beliebigen Zahl kennzeichnen. Trägt man nun die von der Versuchsperson gegebenen Zahlenwerte über die angebotenen Pegel auf, erhält man die nach Stevens benannte **Potenzfunktion**

$$N[\text{sone}] = \left(\frac{I}{I_0} \right)^\alpha \quad (5.1)$$

Dabei bezeichnet N die Lautheit in der Einheit „sone“, I die Intensität des Testtons und I_0 die Intensität des Referenz-Sinustons bei 1 kHz mit einem Pegel von 40 dB SPL, der genau der Lautheit 1 sone entspricht. Der Exponent α nimmt im Mittel über viele Versuchspersonen und Experimente den Wert von etwa 0,3 an. Das bedeutet, daß bei einer Erhöhung der Schallintensität um 10 dB die Lautheit ungefähr um den Faktor 2 ansteigt (da 10 dB eine Verzehnfachung der Intensität bedeutet und $10^{0,3} \cong 2$ ist). Die Lautheit hängt damit in anderer Weise von der Schallintensität ab als der Schallpegel. Die Lautheit in sone darf auch

nicht mit der Lautstärke (in Phon) verwechselt werden, da der erste Begriff die Größe einer subjektiven Empfindung kennzeichnet und der zweite Begriff einen Pegel bezeichnet (d. h. eine physikalische Reizgröße), die zu einer vorgegebenen Lautheitsempfindung führt (nämlich zu der Lautheit, die für alle Lautstärken auf einer Isophonen gleich ist). Daher wird der zweite Begriff in der Normung auch als „Pegellautstärke“ bezeichnet.

Eine andere Möglichkeit zur Bestimmung der subjektiven Lautheitsempfindung besteht bei der **kategorialen Skalierung** darin, daß der Versuchsperson eine gewisse Zahl von Lautheitskategorien (z. B. „sehr leise“, „leise“, „mittel“, „laut“ und „sehr laut“) als Antwortmöglichkeiten vorgegeben wird. Bei dem Verfahren der „Würzburger Hörfeldskalierung“ nach Heller (1985) wird jede dieser groben Kategorien in einem zweiten Schritt in 10 feinere Kategorien unterteilt. Bei der „Oldenburger Hörflächenskalierung“ (Hohmann und Kollmeier, 1995), die sich insbesondere für Messungen mit schwerhörigen Patienten eignet, werden neben den fünf Hauptkategorien vier Zwischenkategorien und die Grenzfälle „nicht gehört“ und „zu laut“ vorgegeben, so daß die Gesamtzahl der Antwortkategorien nicht 50, sondern 11 beträgt. Trägt man die Antwort der Versuchsperson in Kategorialeinheiten über dem Pegel des Testschalls auf, erhält man im mittleren Pegelbereich eine Gerade, d. h. die empfundene Kategoriale-Lautheit ist dort proportional zum Logarithmus der Intensität (vgl. Abbildung 5.2). Abweichungen zu der im logarithmischen Maßstab aufgetragenen Lautheit in sone treten bei sehr niedrigen und bei sehr hohen Pegeln auf:

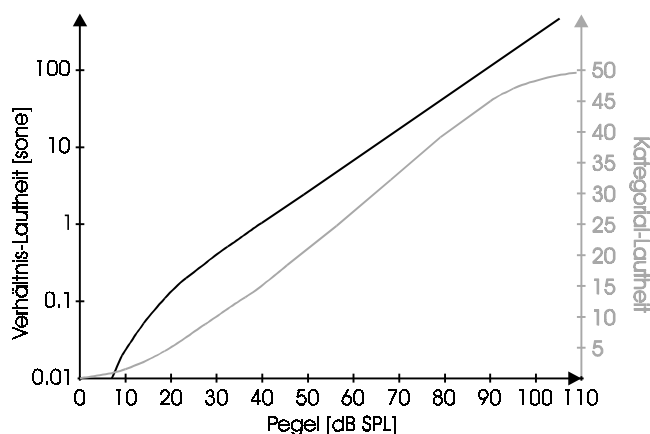


Abbildung 5.2: Sone-Skala und Kategoriale-Skala über den Pegel

Bei sehr niedrigen Pegeln stehen für die relativ grobe Kategoriale-Skalierung nicht genügend Antwortalternativen zur Verfügung, so daß die Kurve erst bei mittleren Pegeln ansteigt, während die Lautheit in sone gegen Null geht und ihr Logarithmus

dementsprechend steil abfällt. Bei sehr hohen Pegeln ist der Antwortbereich der Kategorial-Lautheit ebenfalls begrenzt, so daß hier ein Sättigungseffekt auftritt, während der Logarithmus der Lautheit in some weiter linear ansteigt. Dieser Effekt ist auf den „Logarithmic Response Bias“ (Poulton, 1989) zurückzuführen, d. h. die Versuchspersonen beurteilen bei der Erhöhung einer großen Zahl nicht mehr den absoluten Zuwachs dieser Zahl, sondern das Verhältnis der neuen Zahl zur alten Zahl. Damit tendieren die Versuchspersonen bei hohen Pegeln bei der absoluten Größenschätzung dazu, ihre Empfindungsgröße mit dem Logarithmus der angenommenen, großen Zahlen zu beschreiben. Dieser Effekt tritt bei der Kategorial-Skalierung mit einer wesentlich kleineren Zahl von Antwortalternativen nicht auf, so daß die ungefähre Proportionalität zwischen dem Logarithmus der Lautheit in some und dem absoluten Wert der Kategorial-Skala plausibel erscheint (paralleler Verlauf der Kurven in Abbildung 5.2 über einen relativ großen Pegelbereich).

Eine logarithmische Beziehung zwischen der Reizgröße (in unserem Fall die Intensität I) und der Empfindungsgröße E folgt auch aus dem **Weber-Fechner'schen Gesetz**, das eines der ersten und wichtigsten Gesetze der Psychophysik darstellt. Weber und Fechner konnten nachweisen, daß der Reizunterschied ΔI , der zu einer eben wahrnehmbaren Änderung der Empfindung ΔE führt, proportional zu der absoluten Größe I des Reizes ist, so daß der sogenannte **Weber-Bruch** $\frac{\Delta I}{I}$ für alle Werte von I annähernd konstant sein muß und proportional der kleinsten wahrnehmbaren Empfindungsänderung ΔE ist:

$$\Delta E = k \cdot \frac{\Delta I}{I} \quad (5.2)$$

Durch Aneinanderreihen von eben merklichen Unterschieden ΔE (was mathematisch der Aufintegration von Gleichung 5.2 entspricht) läßt sich nun eine Beziehung zwischen der Empfindungsgröße E und der Reizgröße I herleiten:

$$E = k' \cdot \log_{10} I + k'' \quad (5.3)$$

k' und k'' sind Konstanten, die aus den Randbedingungen festgelegt werden. Diese logarithmische Abhängigkeit der Wahrnehmungsgröße von der Reizgröße entspricht gerade dem Zusammenhang zwischen Kategorial-Lautheit und Tonpegel, d. h. dem Logarithmus der Intensität. Dieses Gesetz entspricht auch der sone-Skala unter der Annahme, daß die empfundene Lautheit durch den Logarithmus der von der Versuchsperson angegebenen Zahl, und nicht durch die Zahl selber

wiedergegeben wird.

Neben der bisher behandelten absoluten Wahrnehmung der Schallintensität interessiert auch die **differentielle Wahrnehmung**, d. h. die kleinste wahrnehmbare Änderung einer Schallintensität. Sie wird auch als JND („just noticeable difference“) bezeichnet und tritt als ΔI im Weber-Fechner'schen Gesetz (Gleichung 5.2) auf. Für die meisten Schalle beträgt die JND etwa 1 dB, d. h.

$$\frac{\Delta I + I}{I} = 10^{\frac{1 \text{ dB}}{10}} \Rightarrow \frac{\Delta I}{I} = 0,259 \quad (5.4)$$

In der Tat mußte die erste logarithmische Pegelskala, deren Einheit das Bel ist (nach Alexander Graham Bell, der unabhängig von Philip Reis das Telefon erfand und zu einem kommerziellen Erfolg führte, $1 \text{ Bel} = \log_{10}(I/I_0)$) in eine 10-fach feinere Skala, das **Dezibel (dB)** unterteilt werden. Damit entspricht eine Einheit dieser Skala gerade einem kleinsten hörbaren Pegelunterschied. Während bei breitbandigen Signalen über einen weiten Pegelbereich Webers Gesetz gültig ist und der kleinste wahrnehmbare Pegelunterschied 1 dB beträgt, ist bei schmalbandigen Signalen (wie Sinustönen) dieser Unterschied bei kleinen Pegeln größer als 1 dB und nimmt mit zunehmendem Pegel stetig ab. Diese Abweichung wird als „near miss to Weber's law“ bezeichnet und hängt mit der speziellen Anregung der Basilarmembran bei Sinustönen zusammen.

V.1.2 Tonhöhe und Frequenzauflösung

Wird die Frequenz eines Sinustons erhöht, führt dies zu der Wahrnehmung einer ansteigenden Tonhöhe. Wenn diese Tonhöhenwahrnehmung dem Weber-Fechner'schen Gesetz entsprechend proportional zum Logarithmus der Frequenz wäre, entspräche dies genau der musikalischen Tonhöhenempfindung: Eine Oktave (Frequenzverhältnis 2:1) würde als doppelt so hoch empfunden werden, unabhängig von der jeweiligen Frequenz. Tatsächlich folgt die Tonhöhenwahrnehmung dieser Gesetzmäßigkeit nur in einem begrenzten Bereich mittlerer Frequenzen, der bei musikalisch ausgebildeten Versuchspersonen sich noch etwas erweitern läßt. Läßt man Versuchspersonen jedoch unabhängig von musikalischen Intervallen die empfundene Tonhöhe im Vergleich zum Referenzton bei 1 kHz angeben, so erhält man die Mel-Skala als subjektive Tonhöhenempfindungsskala. Sie entspricht ungefähr der Frequenz-Orts-Transformation auf der Basilarmembran, bei der die verschiedenen Frequenzen zu einer maximalen Auslenkung an unterschiedlichen Orten führen (gepunktete

Linie in Abb. 5.3). Die gleiche Beziehung zwischen subjektiver Tonhöhe und objektiver Frequenz liefert die in Abb. 5.3 dargestellte Bark-Skala (benannt nach dem deutschen Physiker und Akustiker Heinrich Barkhausen), die auf dem Konzept der Frequenzgruppe beruht (Zwicker und Mitarb. 1957).

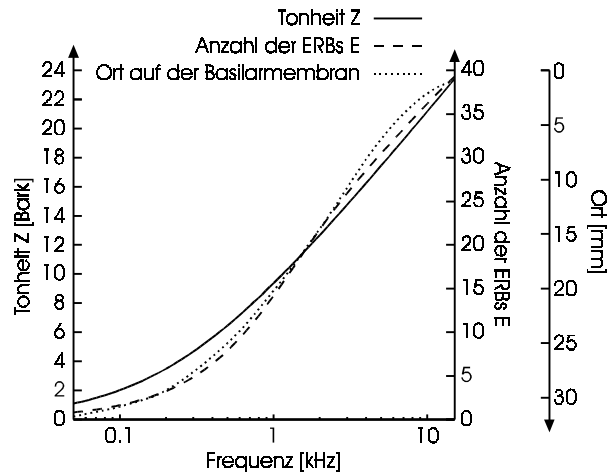


Abbildung 5.3: Mel-Skala, Bark-Skala, ERB-Skala

Sie basiert auf der Frequenzabhängigkeit der Lautheitswahrnehmung und der Maskierung: Bei der **Lautheitssummation** wird die empfundene Lautheit eines schmalbandigen Signals (z. B. eines Schmalbandrauschens) bei gleicher Leistung, aber verschiedener Bandbreite bestimmt (vgl. Abbildung 5.4). Wenn dieselbe Leistung auf einen größeren Frequenzbereich verteilt wird, steigt die wahrgenommene Lautheit auf etwa das Doppelte an (entspricht ungefähr 10 dB Lautstärkengewinn). Bleibt die Leistung jedoch in einem Frequenzbereich, der kleiner als die **Frequenzgruppenbreite** ist, hängt die wahrgenommene Lautheit nicht von der Bandbreite ab, so daß aus dieser Bandbreitenabhängigkeit auf die Größe der Frequenzgruppe geschlossen werden kann. Die Frequenzgruppenbreite beträgt etwa 100 Hz unterhalb von 500 Hz und etwa 1/5 der Frequenz oberhalb von 500 Hz (vgl. Abbildung 5.3).

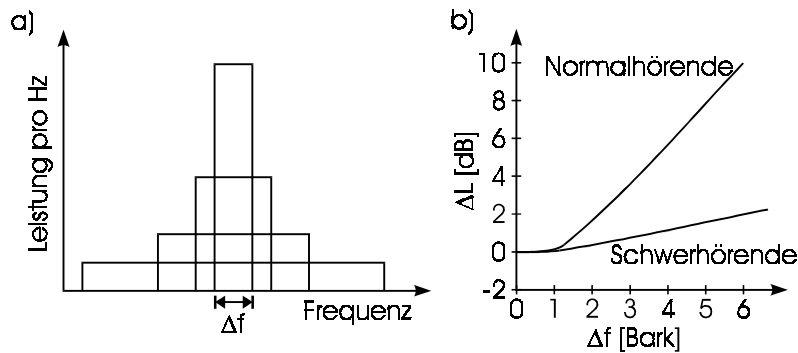


Abbildung 5.4: Bandbreitenabhängigkeit der Lautheit

Anschaulich versteht man unter der Frequenzgruppe diejenige Bandbreite im Gehör, innerhalb derer sämtliche Signale gemeinsam verarbeitet und zu einem „Erregungspegel“ zusammengefaßt werden, der der Lautheitswahrnehmung zugrunde liegt. Wenn das Signalspektrum in verschiedene Frequenzgruppen fällt, wird nicht mehr die Leistung aufsummiert, sondern die Teil-Lautheiten in den verschiedenen Frequenzgruppen.

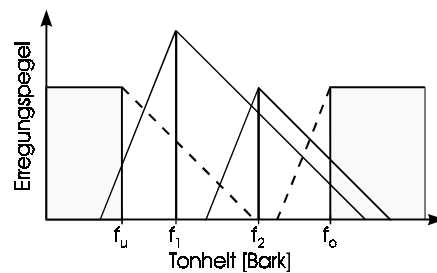


Abbildung 5.5: Erregungspegelmuster

Eine Vorstellung zum Zustandekommen der Frequenzgruppe und der Frequenzauflösung im auditorischen System ist die Erregungsverteilung auf der Basilarmembran (vgl. Abbildung 5.5). Bei einem Sinus-Ton mit f_1 wird die Basilarmembran nicht nur bei dieser Frequenz, sondern im schwächerem Maße auch bei den darüber- bzw. darunterliegenden Frequenzen angeregt. Wenn auf der x-Achse die Frequenz in Bark aufgetragen ist, kann diese Verbreiterung der Erregung als ein dreieckförmiges Muster dargestellt werden, dessen Flanken zu tiefen Frequenzen mit etwa 25 dB pro Bark ansteigen und zu hohen Frequenzen mit etwa 10 dB pro Bark abfallen. Bei hohen Pegeln werden diese Flanken flacher. Wird nun ein weiterer Sinus-Ton mit einer Frequenz f_2 mit einem Pegel angeboten, der unterhalb dieses Erregungspegels bei der Frequenz f_2 liegt, wird dieser zusätzliche Ton vom ersten Ton vollständig **maskiert**, d. h. der zweite Ton wird im Beisein des ersten Tons nicht mehr gehört. Durch derartige Verdeckungsexperimente kann die Form und Steilheit der Erregungspegel-Verteilung über der Frequenz ausgemessen werden.

Um aus derartigen Experimenten auf die Frequenz-Gruppe zu schließen, wird meist ein **Notched-Noise** verwendet, bei dem ein Rauschen nur unterhalb einer Grenzfrequenz f_u und oberhalb einer Frequenz f_o vorliegt. Je schmaler die zwischen f_u und f_o liegende spektrale Lücke („notch“) ist, desto höher liegt die **Mithörschwelle** eines in diesem Frequenzbereich liegenden Tons, d. h. der Pegel des Tons muß relativ hoch sein, damit er gehört wird. Aus der Abnahme dieser Mithörschwelle mit zunehmender Breite der spektralen Lücke kann auf die Frequenzgruppenbreite geschlossen werden, die in Einheiten einer „equivalent rectangular bandwidth“(ERB) nach Moore und Glasberg (1987) gemessen wird. Darauf beruht die ebenfalls in Abb. 5.3 angegebene ERB-Skala, die sehr ähnlich der Bark-Skala ist.

Die Verteilung der Erregungspegel über der Frequenz kann für die Berechnung der Lautheit verwendet werden, wobei die sich überlappenden Bereiche des Erregungspegel-Verlaufs für verschiedene Frequenzen nur einmal berücksichtigt werden. Dazu wird bei jeder Frequenzgruppe der Erregungspegel gemäß dem Potenzgesetz (Gleichung 5.1) umgeformt. Die Gesamt-Lautheit wird dann durch Integration (Aufsummation) der spezifischen Lautheiten über sämtliche Frequenzen gebildet. Diese Berechnung entspricht dem nach Iso 532 B genormten **Lautheitsmodell nach Zwicker**, das für stationäre Schallsignale die subjektiv empfundene Lautheit in sone sehr gut berechnen kann. Dabei werden sowohl spektrale Maskierungseffekte als auch Lautheitssummations-Effekte und Kombinationen dieser Effekte richtig vorhergesagt. Eine Anwendung auf zeitlich stark schwankende Schall (wie z. B. Sprache) ist jedoch problematisch, da das Modell für stationäre Signale entwickelt wurde.

Die Frequenzgruppe läßt sich auch mit einer Reihe weiterer Maskierungsexperimente bestimmen, bei denen sich der maskierende Schall (meistens ein Rauschen) stark im Frequenzbereich ändert, z. B. sinusförmige Änderung im Frequenzbereich („ripple-noise“), Tiefpaßrauschen oder Hochpaßrauschen. Für alle Fälle werden ähnliche Frequenzgruppenbreiten gemessen, wobei sich neben der Breite dieser Frequenzgruppen auch die effektive Form der Filter im Gehör angeben läßt, mit denen jeweils ein begrenzter Frequenzbereich analysiert und zu einer Frequenzgruppe zusammengefaßt wird. Die einfachste Form ist rechteckförmig, eine realistischere Filterform ist ein gerundetes Exponentialfilter dessen Flanken die Verhältnisse im Erregungspegelmuster widerspiegeln. Man darf sich eine Frequenzgruppe nicht als eine starre, bei einer festen Frequenz gelegene rechteckförmige Filterung des Schallsignals vorstellen, sondern eine kontinuierliche, bei jeder Mittenfrequenz auftretende Filterung mit einer abgerundeten, kontinuierlichen Filterform.

Im Gegensatz zu der relativ geringen Frequenzauflösung durch die Frequenzgruppenfilter bei der Maskierung steht die sehr hohe Frequenzauflösung bei der **Diskrimination** verschiedener Frequenzen, die nicht gleichzeitig, sondern in Folge angeboten werden. Dies bei der Musikwahrnehmung (z. B. beim Stimmen von Instrumenten) wichtige Phänomen der **Tonhöhenunterscheidung** ermöglicht einen kleinsten hörbaren Frequenzunterschied von etwa 3 Hz für Frequenzen unterhalb von 500 Hz und etwa 0,6 % für Frequenzen über 1000 Hz. Dies entspricht ungefähr 1/30 Bark, d.h. die Tonhöhenunterscheidung ist wesentlich feiner als die Maskierung im Frequenzbereich. Dies liegt an der massiven Parallelverarbeitung im Gehör: Wenn die Frequenz eines einzelnen Sinustons um wenige Herz verschoben wird, verschiebt sich das ganze Erregungsmuster auf der Basilarmembran, so daß im Mittel über sämtliche beteiligten Nervenfasern selbst ein kleiner Unterschied in der Verschiebung des Erregungsmusters feststellbar ist, obwohl das von einem einzelnen Ton hervorgerufene Erregungsmuster selbst relativ breit sein kann. Daher mißt die Frequenzauflösung (Diskrimination) eine andere Leistung des Gehörs als die Maskierung im Frequenzbereich, die auch als **spektrale Integration** bezeichnet werden kann.

V.1.3 Zeitliche Verarbeitung im Hörsystem

Neben der spektralen Verschmierung (Integration) und Frequenzauflösung spielt die zeitliche Verarbeitung von akustischen Signalen im auditorischen System eine wichtige Rolle, die ebenfalls durch eine zeitliche Verschmierung (Zeitintegration) und eine Zeitauflösung gekennzeichnet werden kann. Die **zeitliche Integration** bezeichnet dabei die Fähigkeit, einen lang andauernden, stationären Klang bei gleichem Pegel als lauter wahrzunehmen als einen kurzen Klang. Diese Eigenschaft kann mit einem Maskierungsexperiment demonstriert werden (vgl. Abbildung 5.6), bei der die Mithörschwelle für einen Signalton bestimmt wird, d. h. der Pegel, bei dem der Ton im Rauschen soeben noch hörbar ist. Sie nimmt mit zunehmender Dauer T des Testtons ab und erreicht für Werte von etwa 200 ms einen stabilen Wert, der durch weitere Verlängerungen des Tons nicht mehr verändert wird. Diese Eigenschaft kann durch eine Energieintegration über einen Bereich von 200 ms erklärt werden. Bei Innenohrschwerhörigen ist diese Energie-Summation gestört, so daß die Schwelle bei zunehmender Dauer T nicht weiter abnimmt.

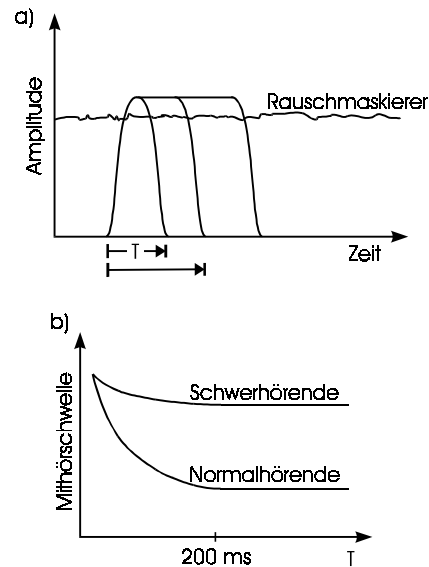


Abbildung 5.6: Zeitliche Integration (Schema, s. Text)

Die zeitliche Verschmierung der internen Repräsentation akustischer Signale läßt sich durch die psychoakustisch meßbaren Phänomene der **Nachverdeckung** und **Vorverdeckung** beschreiben. Bei der Nachverdeckung wird ein Testsignal (z. B. ein kurzer Testtonpuls) zeitlich nach dem Abschalten eines Maskierungssignals (z. B. Rauschen) angeboten und der Pegel des Tons wird solange variiert, bis er soeben nicht mehr hörbar ist (Mithörschwelle). Diese Mithörschwelle nimmt mit zunehmendem Abstand des Testtons vom Maskierende ab und erreicht bei etwa 200 ms die Ruhehörschwelle. Dabei ist der Verlauf dieser Nachverdeckungskurve abhängig von der Maskiererdauer, d. h. vom Adaptationszustand des auditorischen Systems (vgl. Abbildung 5.7). Bei der Vorverdeckung wird der Testton dagegen nicht nach dem Abschalten, sondern zeitlich vor dem Anschalten eines Maskierers angeboten, wobei sich eine ähnliche Verschmierung der Mithörschwelle als Funktion der Verzögerungszeit ergibt wie bei der Nachverdeckung. Allerdings ist der Zeitbereich, über den sich die Vorverdeckung erstreckt, wesentlich kürzer (etwa 10 ms) und es ergibt sich keine vergleichbar starke Abhängigkeit von der Dauer des Maskierers.

Einen kombinierten Effekt von Vor- und Nachverdeckung beobachtet man bei der **Lückendetektion** (Englisch: Gapdetection), bei der die kleinste in einem Rauschen wahrnehmbare Pausendauer gemessen wird (vgl. Abb. 5.8). Da aufgrund der zeitlichen Verschmierung im auditorischen System kleine Lücken im Rauschen nicht wahrgenommen werden können, liegt bei Normalhörenden bei der Verwendung von breitbandigen Rauschen die minimal detektierbare Lückendauer bei etwa 8 ms. Da sie ein einfach und schnell zu bestimmendes Maß für die zeitliche

Verarbeitung ist, wird sie oft zur Charakterisierung der gestörten zeitlichen Verarbeitung bei Schwerhörigen eingesetzt, wobei aufgrund starker interindividueller Schwankungen Werte zwischen 8 und 100 ms auftreten können.

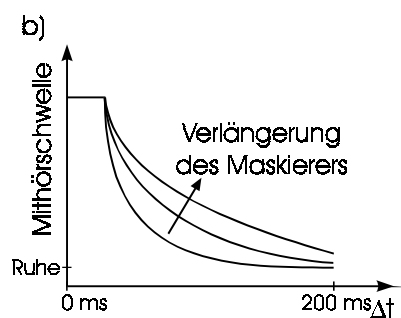
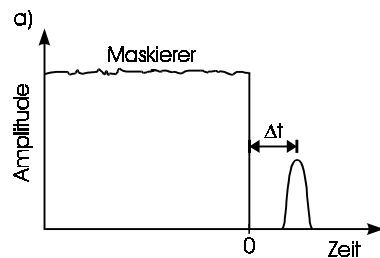


Abbildung 5.7: Nachverdeckung

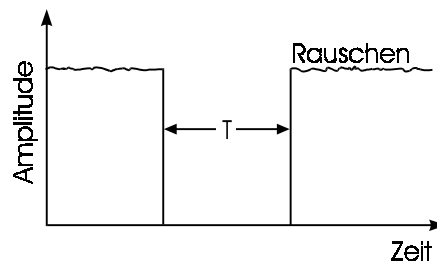


Abbildung 5.8: Gap-detection

Eine weitere Eigenschaft der zeitlichen Verarbeitung akustischer Signale ist die **Modulationswahrnehmung**, d. h. die Wahrnehmung von Schallen mit einer aufgeprägten (z. B. sinusförmig variierenden) Einhüllenden (vgl. Abb. 5.9). Als Trägersignale können Sinussignale oder Rauschen verwendet werden, während als Modulationssignale (d. h. Verformung der Einhüllenden) zumeist sinusförmige Signale mit einer bestimmten Modulationsfrequenz f_{mod} verwendet werden. Eine derartige Signalform ist sehr ähnlich der Form von Sprache, weil bei der Sprachartikulation ein Trägersignal (z. B. die Schwingung der Stimmritze) durch den Vokaltrakt in der Intensität zeitlich stark verändert wird. Im Modulationsfrequenzbereich zwischen 0 und etwa 10 Hz werden Modulationen als Schwankungen der Lautstärke wahrgenommen, die

Empfindung wird als **Schwankungsstärke** bezeichnet. Bei Frequenzen zwischen etwa 10 und 20 Hz entsteht ein „knarrender“ Klangeindruck, der auch als R-Rauhigkeit bezeichnet wird, während Modulationen zwischen etwa 20 Hz und 80 Hz als **Rauhigkeit** des Klanges wahrgenommen werden können. Bei noch höheren Modulationsfrequenzen treten im Frequenzspektrum des akustischen Signals Seitenlinien auf, die aufgrund der spektralen Verfärbung als Klangänderungen wahrgenommen werden können.

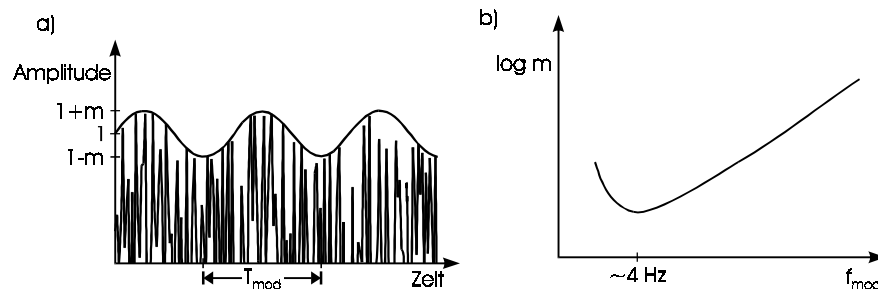


Abbildung 5.9: Moduliertes Signal, PMTF für breitbandige Träger, d. h. der kleinste detektierbare Modulationsgrad als Funktion der Modulationsfrequenz

Als Maß für die Zeitauflösung kann die **psychoakustische Modulationstransfer-Funktion** bestimmt werden, d. h. für jede Modulationsfrequenz derjenige Modulationsgrad m , der soeben noch wahrgenommen werden kann. Sie zeigt für Modulationen von etwa 4 Hz ein Maximum (dies entspricht in etwa der Sprachsilbenfrequenz und ist daher gut an die zu empfangene Sprache angepaßt) und für zunehmende Modulationsfrequenzen wird die Empfindlichkeit geringer. Diese Modulations-Tiefpaßcharakteristik des auditorischen Systems korrespondiert zu der zeitlichen Verschmierung, die in den Vor- und Nachverdeckungsexperimenten aufgezeigt wurde.

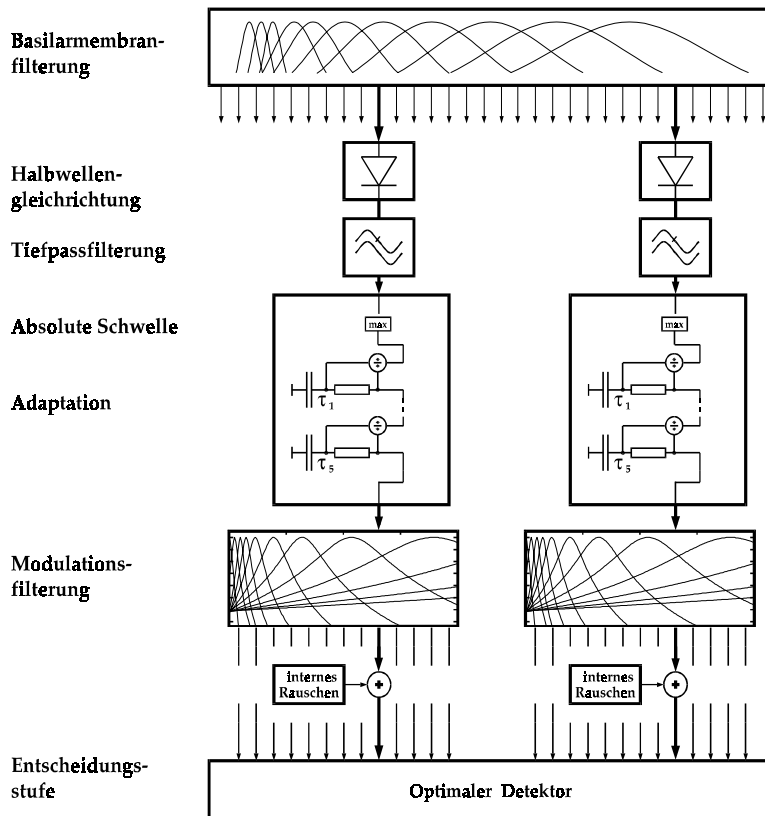


Abbildung 5.10: Verarbeitungsmodell nach Dau et al.

Um die verschiedenen zeitlichen Eigenschaften des Hörsystems (zeitliche Integration, Vor- und Nachverdeckung, Modulationswahrnehmung) mit möglichst wenig Annahmen beschreiben zu können, wurden Hörmodelle zur zeitlichen Verarbeitung aufgestellt. Bei dem in Abbildung 5.10 schematisch dargestellten Verarbeitungsmodell nach Dau et al. (1997) wird die Verarbeitung unterschiedlicher Frequenzen zunächst durch eine Filterbank beschrieben, die das Eingangssignal in verschiedene Frequenzgruppenbreite Bandpaß-Signale aufteilt. In jedem Bandpaßsignal wird eine Halbwellengleichrichtung mit Tiefpaßfilterung durchgeführt, das ungefähr der Aktion der Haarzellen entspricht und eine Extraktion der Einhüllenden bewirkt. Die zeitlichen Verarbeitungseigenschaften werden durch die nachgeschalteten Adaptations-Schleifen nachgebildet, bei denen das Eingangssignal durch das tiefpaßgefilterte Ausgangssignal geteilt wird. Dadurch wird eine gewisse Adaptation an den Mittelwert des Eingangssignals ermöglicht, während schnelle Änderungen (Fluktuationen) ohne weitere Beeinflussung durchgelassen werden. Durch Hintereinanderschalten mehrerer dieser **Nachregelschleifen** mit unterschiedlichen Zeitkonstanten zwischen 1ms und 500 ms wird approximativ eine logarithmische Kompression für langsame Änderungen im Eingangssignal erreicht und eine unveränderte Durchlässigkeit für schnelle Änderungen im Eingangssignal. Mit dieser Modellstruktur lassen sich Vor- und Nachverdeckung sowie die Lückendetektion (Gap detection)

bereits sehr gut beschreiben. Die zeitliche Integration von Testtönen und die Modulationswahrnehmung kann durch Einfügen einer Modulationsfilterbank erreicht werden, die für jede Mittenfrequenz die Einhüllenden-Fluktuationen in verschiedene Modulationsfrequenzen aufspaltet. Am Ausgang dieser Modulationsfilterbank steht ein bestimmtes, zeitlich variables Muster zur Verfügung, das die **interne Repräsentation** von akustischen Signalen darstellen soll. Mit dieser internen Repräsentation lassen sich verschiedene Experimente einfach dadurch beschreiben, daß eine bestimmte, minimale Änderung in dieser Repräsentation zur Wahrnehmung eines Unterschieds im Klang führt, der gerade zur Detektion des „gesuchten“ akustischen Stimulus führt.

V.1.4 Binaurale Interaktion

Neben den bisher beschriebenen Leistungen des Gehörs bei der Intensitäts-, Frequenz-, und Zeitverarbeitung im auditorischen System ist die binaurale Interaktion, d. h. die im Gehirn stattfindende vergleichende Verarbeitung der an beiden Ohren anliegenden Signale von besonderer Wichtigkeit für das Hören in natürlicher Umgebung. Es trägt signifikant zur **Hallunterdrückung**, zur Ortung (**Lokalisation**) von Schallquellen im Raum und zur **Unterdrückung** von „unerwünschten“ Störgeräuschquellen in realen akustischen Situationen bei.

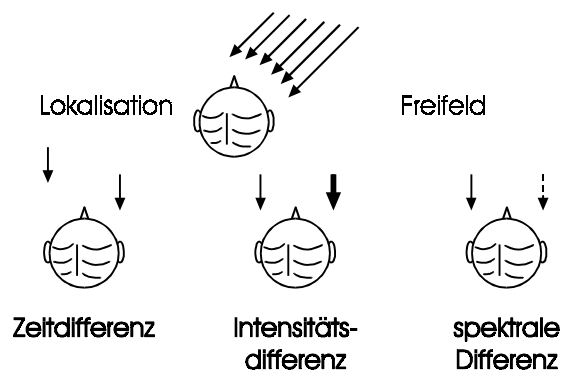


Abbildung 5.11: Physikalische Merkmale für die Lokalisation

Wenn eine Schallquelle so im Raum angeordnet ist, daß der von ihr ausgesendete Schall mit einem bestimmten Einfallswinkel auf den Kopf des Hörers fällt, bewirkt dieser (winkelabhängige) Schalleinfall eine **interaurale Zeitverzögerung** (d. h. der Schall erreicht das der Schallquelle zugewandte Ohr eher als das abgewandte Ohr), einen **interauralen Intensitätsunterschied** (d. h. das der Schallquelle zugewandte Ohr empfängt eine höhere Schallintensität), sowie eine von der Einfallrichtung abhängige **spektrale Verfärbung** (d. h. der Frequenzgehalt des empfangenen Schalls wird je nach Einfallrichtung

unterschiedlich verändert vgl. Abb. 5.11.). Aufgrund dieser akustischen Merkmale ist der normalhörende Mensch bei breitbandigen Signalen in der Lage, den Ort einer Schallquelle aus der Vorne-Richtung mit einer Ungenauigkeit von etwa 1 Grad bzw. bei seitlichem Einfall oder Schalleinfall von oben mit einer Ungenauigkeit von etwa 5 Grad aufzulösen. Diese erstaunlich hohe Ortungsleistung wird allerdings bei schmalbandigen Signalen und bei Vorliegen von Nachhall bzw. Echo eingeschränkt.

Dieser Effekt ist darauf zurückzuführen, daß die bei einer bestimmten Frequenz anliegenden interauralen Intensitäts- und Zeitunterschiede nicht eindeutig zu einer bestimmten Einfallsrichtung gehören, sondern daß mehrere Schalleinfallrichtungen bei derselben Frequenz zu denselben interauralen Intensitäts- und Zeitunterschieden führen können. Beispielsweise führen sämtliche Einfallsrichtungen in der Medianebene zu einer interauralen Zeit- und Laufzeitdifferenz von ungefähr 0, so daß bei schmalbandigen Signalen nicht sicher zwischen der Vorne-, Oben-, und Hinten-Einfallsrichtung unterschieden werden kann und es zu **Richtungsverwechslungen** kommt. Ein Schalleinfall von vorne links kann in ähnlicher Weise mit einem Schalleinfall von hinten links und Einfallsrichtungen verwechselt werden, die auf den sogenannten „**Cone of Confusion**“ liegen. Bei schmalbandigen akustischen Signalen werden zudem bestimmte Frequenzen zu bestimmten Schalleinfallrichtungen zugeordnet („**richtungsbestimmende Bänder**“ nach Blauert, 1974), so daß hier die spektrale Information teilweise die Information aus den interauralen Zeit- und Intensitätsunterschieden überschreibt. Erst bei breitbandiger Schalldarbietung kann durch den Vergleich über mehrere Frequenzen hinweg eine eindeutige, sämtliche Mehrdeutigkeiten bei schmalbandiger Signaldarbeitung vermeidende Lokalisation durchgeführt werden.

Die Rolle der interauralen Zeit- und Intensitätsdifferenzen für die Lokalisation und die binaurale Verarbeitung kann mit Kopfhörer-Experimenten erforscht werden, bei denen jeder dieser physikalischen Parameter einzeln variiert werden kann. So stellt sich heraus, daß die interauralen Zeitdifferenzen insbesondere bei niedrigen Frequenzen (unter 1,5 kHz) eine dominierende Rolle für die Lokalisation spielen und der kleinste wahrnehmbare interaurale Zeitunterschied etwa 20 μ s beträgt, während er oberhalb von 1,5 kHz etwa 50 μ s beträgt. Die interauralen Intensitäts-Unterschieds-Schwellen betragen dagegen unterhalb von 1,5 kHz etwa 3 dB und oberhalb von 1,5 kHz etwa 1 dB, so daß sie besonders bei hohen Frequenzen für die Lokalisation dominant sind. Diese bereits von Lord Rayleigh (1877) formulierte **Duplex-Theorie** des binauralen Hörens hat ihre physikalische Begründung darin, daß bei niedrigen Frequenzen aufgrund der (relativ zum Kopfdurchmesser) großen Schallwellenlänge nur geringe Intensitätsunterschiede zwischen

den beiden Ohren auftreten, so daß Intensitätsunterschiede erst bei hohen Frequenzen verlässlich ausgewertet werden können. Bei niedrigen Frequenzen können die Neuronen im auditorischen System der Schallwellenform noch sehr gut folgen und daher interaurale Zeit-Unterschiede gut detektieren, während bei hohen Frequenzen interaurale Zeit-Differenzen nur im Einhüllenden-Verlauf der Signale, nicht jedoch in der (von den Neuronen nicht mehr aufgelösten) Feinstruktur der übertragenden Signale erfaßt werden können.

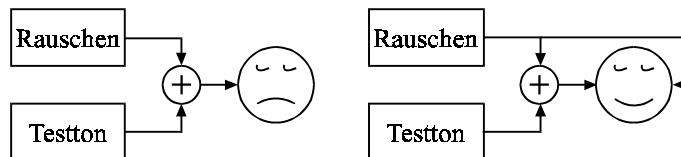


Abbildung 5.12: Binaurale Störgeräuschbefreiung

Zur Erfassung der **Störgeräuschbefreiung** im binauralen System kann ein psychoakustisches Experiment verwendet werden, bei dem die Detektion von Testtönen im Rauschen zwischen der binauralen und der monauralen Situation verglichen werden (vgl. Abb. 5.12). Zunächst wird der Versuchsperson ein Ton im Rauschen angeboten, dessen Pegel solange verringert wird, bis die Versuchsperson ihn nicht mehr hört (Mithörschwelle). Wenn anschließend dasselbe Rauschen am anderen Ohr gleichzeitig angeboten wird, wird die wahrgenommene Lautheit des Rauschens zwar höher, der Ton wird jedoch wieder hörbar, so daß eine Verringerung der Mithörschwelle gemessen werden kann, die als binaurale **Masking Level Difference (MLD)** bezeichnet wird. Ihre Größe hängt von der interauralen Phasenlage des Testtons und der zwischen den beiden Ohren bestehenden Beziehung zwischen dem Rauschen ab. Beispielsweise wird die MLD maximal (ca. 14-18 dB), wenn dasselbe Rauschen und ein um 180 Grad (Phase π) verschobener Sinuston bei etwa 500 Hz verwendet wird. Keine MLD tritt dagegen auf, wenn für das Rauschen und das Signal dieselbe Beziehung an beiden Ohren vorliegt (z. B. bei identischem Rauschen und identischem Ton an beiden Ohren oder bei Phaseninvertierung zwischen beiden Ohren).

In einer guten Näherung kann diese Störgeräuschunterdrückung des binauralen Systems durch eine (mit Restfehlern behaftete) **Subtraktion** der beiden Ohrsignale aufgefaßt werden, die im Gehirn nach einer Anpassung der vom rechten und linken Ohr stammenden Signale (z. B. durch entsprechende Zeitverzögerung und Amplituden-Anpassung) vorgenommen wird. Mit dieser als „Equalization and Cancellation Theory“ (EC-Theorie) nach Durlach (1972) bezeichneten Modellvorstellung kann bereits eine Vielzahl der binauralen Maskierungsexperimente quantitativ gedeutet werden. Ein eher auf physiologischen Vorstellungen basierendes Modell der binauralen Signalverarbeitung wurde dagegen von Jeffress (1948) aufgestellt. Es sieht eine Art neuronales Kreuzkorrelations-Netzwerk vor, bei dem die an den beiden Ohren ankommenden

Signale in einer Laufzeitkette sukzessiv zeitverzögert werden und durch Koinzidenz-Dektoren die Laufzeitunterschiede zwischen den beiden Ohren jeweils abgegriffen werden. Auf der Basis dieses Modells wurden in der neueren Literatur eine Reihe weiterer Modelle entwickelt, die die binaurale Signalverarbeitung einschließlich einer Reihe von Effekten richtig vorhersagen (z. B. die Lateralisation bzw. Lokalisation von Schallereignissen mit bestimmter interauraler Zeit- und Pegeldifferenz, das Gesetz der ersten Wellenfront, die MLD für verschiedene Signale und Maskierer, sowie zeitliche und spektrale Eigenschaften dieser Effekte). Da auf diesem Gebiet noch weiter intensiv geforscht wird, sei auf die weiterführende Literatur von Blauert (1983) und Colburn (1996) verwiesen.

V.2 Signaldetektion und Wahrnehmung komplexer Signale

V.2.1 Signaldetektionstheorie

In der Psychoakustik untersucht man nicht nur die bereits beschriebene Wahrnehmung von Schall-Grundgrößen (und ihre Änderungen, die Diskrimination), sondern man interessiert sich auch für die **Detektion** von Signalen, d. h. das Erkennen von bestimmten Signalen in einem Hintergrundsignal. Für die Praxis ist dies beispielsweise wichtig bei der Erkennung von Warnsignalen oder der Wahrnehmung von Sprache in einer lärmbehafteten Umgebung. Um die Signaldetektion eines menschlichen Beobachters formal möglichst exakt beschreiben zu können wurde in den 50er Jahren von Green und Swets die **Signal-Detection-Theory** in die Psychophysik eingeführt, die ursprünglich in der Radartechnik im 2. Weltkrieg entwickelt wurde (dort war es z. B. notwendig, die Anwesenheit eines feindlichen Flugzeugs im rauschbehafteten Radarsignal mit einer ausreichenden Sicherheit und einer möglichst geringen Störanfälligkeit zu erkennen). Übertragen auf die Psychoakustik ergeben sich die folgenden möglichen experimentellen Paradigmen:

Single Interval Yes/No Aufgabe (Abb. 5.13): Bei ihr soll die Versuchsperson entscheiden, ob in einem Intervall, das mit Rauschen angefüllt ist, zusätzlich ein Ton enthalten ist. Dabei können die beiden Reiz-Fälle „N“ (für Noise: Nur Rauschen liegt vor) und „S + N“ (Signal und Rauschen liegt vor) unterschieden werden, sowie die beiden Antwortalternativen „No“ (kein Signal lag vor) und „Yes“ (das Signal lag vor). Die vier Möglichkeiten der Kombination aus Reiz-Darbietung und Antwort der Versuchsperson sind in Tabelle 5.1 dargestellt. Weiterhin ist die Wahrscheinlichkeit aufgetragen, daß die Versuchsperson bei vorgegebenem Reiz eine der beiden Antwortalternativen wählt.

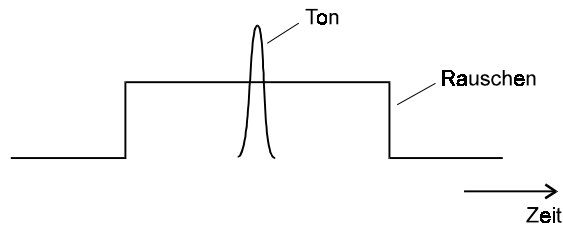


Abbildung 5.13: Schematische Anordnung des Single-Intervall Yes/No-Verfahrens: In einem Rausch-Intervall ist ein Ton vorhanden (oder nicht).

Tabelle 5.1 Mögliche Reizkonfigurationen und mögliche Antworten (samt zugehöriger Wahrscheinlichkeit) im Single-Intervall-Yes/No-Task:

Antwort	Reiz	
	N (nur Rauschen)	S+N (Signal+Rauschen)
No	$P(\text{No} N)$ „Correct rejection“	$P(\text{No} S+N)$ „Miss“
Yes	$P(\text{Yes} N)$ „False Alarm“	$P(\text{Yes} S+N)$ „Hit“
	$\Sigma=1$	$\Sigma=1$

Die Spaltensumme der Wahrscheinlichkeiten muß dabei jeweils = 1 sein, so daß in jeder Spalte die Angabe einer Wahrscheinlichkeit für die vollständige Charakterisierung der in der Matrix angegebenen Größen ausreicht. So wird beispielsweise meistens die Falsch-Alarmrate ($P(\text{Yes}|N)$) und die Trefferrate ($P(\text{Yes}|S+N)$) angegeben.

2-Alternative-Forced Choice (2 AFC): Bei diesem Versuchsparadigma muß sich die Versuchsperson für eine von 2 Antwortalternativen entscheiden, so daß sie zu einer Aussage „gezwungen“ wird. Beispielsweise können (wie in Abb. 5.14 angegeben) zwei Rauschintervalle angeboten werden, von denen nur eines (das zufällig ausgewählt wurde) den Testton enthält. Die Aufgabe der Versuchsperson ist es nun, zu sagen, ob das erste oder das zweite Rauschen den Testton enthalten hat. Diese Kopplung der Alternativen an zwei verschiedene Zeit-Intervalle wird auch als 2-Intervall-Alternative-Forced Choice (2I-AFC) bezeichnet.

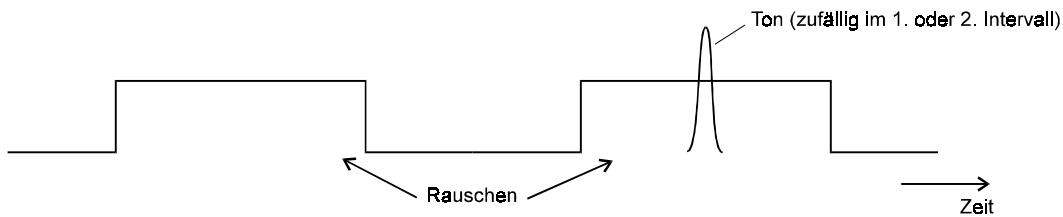


Abbildung 5.14: Schematische Darstellung des 2-Intervall-AFC-Verfahrens: Der Ton befindet sich nur in einem (zufällig ausgesuchten) der beiden angebotenen Intervalle mit einem Hintergrund-Rauschen.

N-Alternative-Forced Choice (n AFC): Bei dieser Verallgemeinerung des 2 AFC-Verfahrens stehen n-Antwort-Alternativen zur Verfügung, wobei in der Psychoakustik dies oft gleichbedeutend mit n Intervallen ist. Am verbreitetsten ist beispielsweise das 3-AFC-Verfahren, bei dem in einem von drei aufeinanderfolgenden Rauschintervallen der Testton enthalten sein kann.

Das Entscheidungsverhalten der Versuchsperson in diesen verschiedenen experimentellen Paradigmen (die in ähnlicher Form in der Radar-Technik vorkommen) lässt sich nun durch ein mathematisches Modell formal beschreiben. Dies geht davon aus, daß in jedem Intervall der „innere Eindruck“ X bei der Versuchsperson entsteht, der gewissermaßen die Ähnlichkeit des empfangenen Signals mit dem zu erwartenden Stimulus (Nutzsignal im Störgeräusch) beschreibt (bei der Theorie des „optimalen Detektors“ wird diese Ähnlichkeit sogar durch Kreuzkorrelation zwischen der aktuellen „internen Repräsentation“ des Eingangsstimulus mit der „internen Repräsentation“ des gesuchten Eingangsstimulus ausgedrückt (vgl. Dau et al., 1996)). Im allgemeinen ist dieser innere Eindruck eine Zufallsvariable, die entweder der Häufigkeitsverteilung $P(X | N)$ entstammt (für den Fall, daß das Rauschen ohne das Signal angeboten wurde) oder der Verteilung $P(X | S+N)$ für den Fall, daß X der Verteilung von Nutzsignal + Störsignal entstammt. Im weiteren nehmen wir an, daß die Anwesenheit des Signals die Verteilung $P(X | S+N)$ gegenüber der Häufigkeitsverteilung für Rauschen ohne Signal verschiebt, d. h. mit zunehmendem Pegel (zunehmender Detektierbarkeit des zu detektierenden Signals) wird die Überlappung zwischen den beiden Verteilungen immer geringer. Wir nehmen nun an, daß sowohl $P(X | N)$ und $P(X | S+N)$ **normalverteilt** sind und jeweils die gleiche Standardabweichung σ aufweisen. Damit kann der Abstand zwischen diesen beiden Verteilungen mit der Standardabweichung normiert werden, so daß das **normierte Detektierbarkeitsmaß** d' resultiert:

$$d' = \frac{\text{Abstand}}{\sigma} \quad (5.5)$$

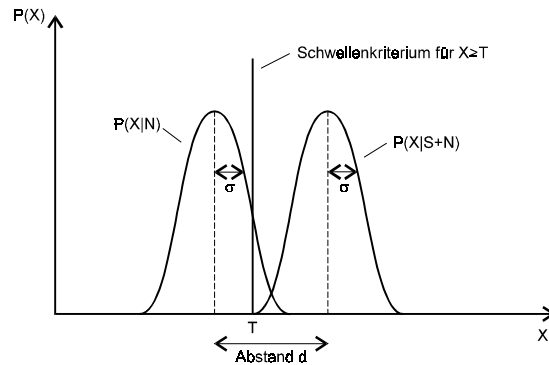


Abbildung 5.15: Wahrscheinlichkeitsverteilungen für den internen Eindruck X bei Vorliegen eines Rauschens bzw. eines Rauschens mit Signal, wobei X in beliebigen Einheiten angegeben ist. Zusätzlich ist das Schwellenkriterium T angezeigt, daß entweder lax (T bei sehr kleinen Werten von X) oder konservativ (T bei sehr großen Werten von X) gelegt werden kann.

Die Annahme ist nun, daß die Versuchsperson ein (zunächst frei wählbares) Kriterium T benutzt, bei dessen Überschreiten sie antwortet, daß das Signal vorgelegen hat und bei dessen Unterschreiten sie antwortet, das Signal sei nicht im Rauschen vorhanden gewesen, also:

„Yes“, falls $X \geq T$ und „No“, falls $X < T$

Die Trefferrate $P(Y | S+N)$ ist daher ebenso wie die „Falsch-Alarm-Rate“ $P(Y | N)$ von diesem Kriterium abhängig: Bei laxem Kriterium (d. h. bei einer Lage von T bei sehr kleinem X) sind beide Werte recht groß, während bei sehr konservativem T (Kriterium sehr weit nach rechts verschoben in Abb. 5.15) beide Werte sehr klein sind. Dabei ist die Trefferrate gerade das Integral über die Fläche rechts von T der Verteilung $P(X | S+N)$, während die „Falsch-Alarm-Rate“ $P(Y | N)$ gerade das Integral der Fläche von $P(X | N)$ für $X > T$ ist. Bei systematischer Variation von T kann das Integral dieser beiden Wahrscheinlichkeitsfunktionen gerade ausgewertet werden und gegeneinander aufgetragen werden, so daß sich die in Abb. 5.16 dargestellte **Receiver Operating Characteristic (ROC)** ergibt.

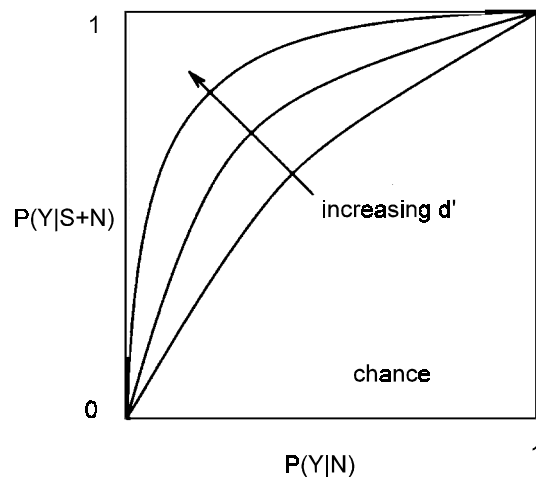


Abbildung 5.16: Receiver Operating Characteristic, d. h. Trefferrate gegen Falsch-Alarm-Rate für verschiedene Werte des Detektierbarkeits-Maßes d' . (nach Luce, R. D.: Sound and Hearing. L. Erlbaum Verlag, New Jersey 1993)

Bei einem vorgegebenen Wert von d' (z. B. $d' = 1$) ergibt sich eine links der Winkelhalbierenden liegenden Kurve, die vom Punkt (0,0) für ein extrem konservatives Kriterium bis zum Punkt (1,1) für ein extrem laxes Kriterium geht. Falls $d'=0$ ist (d. h. die Situationen N und S+N unterscheiden sich überhaupt nicht, z. B. weil der Signalpegel unendlich klein ist), wird die Winkelhalbierende als ROC-Kurve angenommen (Zufallsauswahl), während mit zunehmender Detektierbarkeit des Signales (wachsende d') die ROC-Kurven sich immer mehr in die linke obere Ecke „drängen“. Mit Hilfe dieser Schar von ROC-Kurven ist es nun möglich, daß **Kriterium** aus der Detektions-Aufgabe zu **eliminieren**, in dem man bei einem vorgegebenen Experiment die Falsch-Alarm-Rate und die Trefferrate in die in Abb. 5.16 dargestellte Kurvenschar einträgt und sich diejenige ROC-Kurve heraussucht, die am nächsten bei dem eingezeichneten Datenpunkt liegt. Der zu dieser Kurve gehörende Detektierbarkeit Index d' dient dann als **kriterienfreies Maß** für die Detektierbarkeit des Signals. Derselbe Wert von d' könnte beispielsweise gemessen werden, wenn die Versuchsperson gezwungen wird, ein wesentlich laxeres Kriterium anzunehmen (z. B. wenn alle „Treffer“ bei den Antworten der Versuchsperson belohnt werden, ohne daß die „falschen Alarme“ geahndet werden) oder ein wesentlich konservativeres Kriterium annimmt (z. B. indem alle falschen Alarme bestraft werden). Daher kann man durch den entsprechenden Eintrag sowohl den gesuchten Wert für d' ermitteln als auch die Lage des individuellen Kriteriums.

Bei psychoakustischen Detektionsexperimenten (z. B. Maskierung eines Sinustons in einem Rauschen) hängt der Detektierbarkeits-Index d' oft

anhand eines einfachen, empirisch gefundenen Zusammenhanges mit dem Signal-Rausch-Verhältnis zusammen:

$$d' = m \cdot \left(\frac{E}{N_0} \right)^k \quad (5.6)$$

E bezeichnet dabei die Energie des Testtons (Leistung des Testtons mal Zeit) und N_0 die spektrale Leistungsdichte des Maskierers (Leistung des Rauschens geteilt durch Bandbreite), während m und k Parameter sind, die an die jeweilige Daten angepaßt werden müssen.

Beachte, daß der Signal-Rauschabstand nicht als das Verhältnis zweier Energien, sondern als das Verhältnis einer Energie und einer spektralen Rauschleistungsdichte (d. h. Energie pro Zeit pro Frequenz) dargestellt ist, was dimensionsmäßig dieselbe Größen darstellt. Sofern der Testton eine Dauer von ca. 200 ms nicht überschreitet, ist dieser Wert E/N_0 relativ unabhängig von der Testondauer und der spektralen Form des Maskierers und beträgt etwa 10 dB an der Detektionsschwelle für $d'=1$.

Als Alternative zu dem für die Entwicklung der ROC-Kurven zugrundegelegten Single-Interval Verfahrens kommt das o. g. 2 IFC-Verfahren als „kriterienfrei“ in Frage. Im Sinne der oben beschriebenen Modellbildung (vgl. Abb. 5.15) kann dieses Experiment als die Entnahme von 2 Beobachtungsgrößen X_1 (interner Eindruck während des Intervalls 1) und X_2 (interner Eindruck im Intervall 2) beschrieben werden, die miteinander verglichen werden. Man geht nun davon aus, daß die Versuchsperson einfach die Differenz zwischen den beiden Größen bildet und je nach Vorzeichen das erste oder zweite Intervall als dasjenige Intervall auswählt, das mit höherer Wahrscheinlichkeit den Testton enthalten hat, d. h.

$$\text{Entscheidung: } X_1 - X_2 \begin{cases} \geq 0: V_p \text{ sagt: Signal im Intervall 1} \\ < 0: V_p \text{ sagt: Signal im Intervall 2} \end{cases}$$

Durch diese Differenzbildung fällt das subjektive Schwellenkriterium T aus der Auswertung heraus, so daß das 2AFC-Verfahren oft als „**kriterienfreies Verfahren**“ bezeichnet wird. Man kann es in dem Bild von Abb. 5.15 auch so beschreiben, daß die Verteilung der Differenz zwischen den beiden dort abgebildeten Normalverteilungen wieder normalverteilt ist, wobei als Entscheidungskriterium der Wert „Null“ angenommen wird.

Bei der Generalisierung auf ein n-AFC-Verfahren läßt sich entsprechend das

Experiment als ein Vergleich von N Stichproben beschreiben, wobei dasjenige Intervall von der Versuchsperson „ausgewählt“ wird, aus dem die größte Stichprobe X stammt. Aufbauend auf diesen Annahmen lässt sich die psychometrische Funktion für die verschiedenen Verfahren aus dem Detektierbarkeits-Index d' ausrechnen, d.h. die Richtigantwort-Wahrscheinlichkeit als Funktion des Signalpegels (Abb. 5.17).

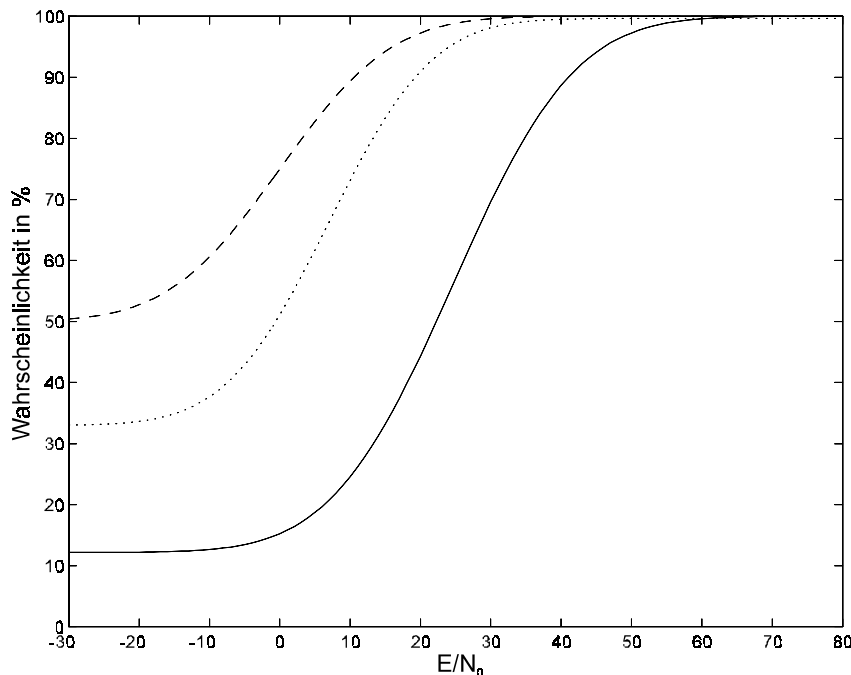


Abbildung 5.17: Psychometrische Funktion (d. h. Wahrscheinlichkeit einer richtigen Antwort als Funktion des Signalpegels) für ein Single-Intervall-Yes/No-Verfahren (durchgezogen), ein 2AFC-Verfahren (gestrichelt) und ein 3AFC-Verfahren (gepunktet).

In der Praxis interessiert man sich meist nicht für die genaue Form und Lage der gesamten psychometrischen Funktion, sondern man möchte mit möglichst geringem Zeitaufwand die **Detektionsschwelle** bestimmen, d. h. den Pegelwert, der zu einer vorgegebenen Richtig-Antwort-Wahrscheinlichkeit (bei einem vorgegebenen psychophysikalischen Paradigma) führt. Diese Schwellen-Wahrscheinlichkeit wird oft als Mittelwert zwischen zufälliger Entscheidung und vollständig richtiger Entscheidung angenommen (z. B. $P=0,75$ für das 2AFC-Verfahren). Dazu verwendet man **adaptive Pegelsteuerungs-Algorithmen**, die den Testsignal-Pegel in Abhängigkeit von den gegebenen Antworten der Versuchsperson möglichst in Richtung auf den gesuchten Schwellenpegel hin bewegen. Die in der Praxis am häufigsten eingesetzten Verfahren sind die von Levitt (1971) beschriebenen **transformed-up-down-Verfahren**, bei denen nach einer bestimmten Sequenz von Richtig/Falsch-Antworten der Testtonpegel erhöht wird und nach anderen, genau festgelegten

Sequenzen von Antworten der Pegel erniedrigt wird. Eine Aufstellung von derartigen adaptiven Testverfahren und die von ihnen angesteuerte „Ziel-Wahrscheinlichkeit ist in Tabelle 5.2 dargestellt:

Tabelle 5.2: Beispiel verschiedener transformierter Up-down-Algorithmen zur adaptiven Pegelsteuerung (nach Levitt, H.: Transformed Up-Down Methods in Psychoacoustics. J. Acoust. Soc. Am. 49, 1971), sowie die von ihnen angestrebte Zielwahrscheinlichkeit.

Entry	Response sequences		Response groupings	
	UP group increase level after:	DOWN group decrease level after:	Probability of a sequence from DOWN group= $P[\text{DOWN}]$	Probability of positive response at convergence
1	–	+	$P(X)$	$P(X) = 0.5$
2	+ – or –	++	$[P(X)]^2$	$P(X) = 0.707$
3	--	– + or +	$[1-P(X)]P(X)+P(X)$	$P(X) = 0.293$
4	++ – or + – or –	+++	$[P(X)]^3$	$P(X) = 0.794$
5	+++ – or ++ – or + – or –	++++	$[P(X)]^4$	$P(X) = 0.841$
6	----	---- + or -- + or – + or +	$1-[1-P(X)]^4$	$P(X) = 0.159$
7	Any group of 4 responses with 1 or more negative responses	++++	$[P(X)]^4$	$P(X) = 0.841$
8	-- – + – + --	++ + – + – + +	$[P(X)]^2[3-2P(X)]$	$P(X) = 0.5$

Das in der Praxis am häufigsten verwendete Verfahren ist das 1 up–2 down-Verfahren (Eintrag 2 in Tab. 5), bei dem der Testton-Pegel nach

zwei aufeinander folgenden richtigen Antworten erniedrigt wird und nach einer falschen Antwort erhöht wird. Es konvergiert auf eine Richtig-Antwort-Wahrscheinlichkeit von 0,707. Aufgrund der höheren Robustheit bzw. Effizienz bietet sich für die Praxis die Kombination eines 3AFC-Verfahrens mit dem 2-Schritt-Algorithmus oder die Kombination eines 3-Schritt-Algorithmus (Pegel-Erniedrigung erst nach 3 aufeinander folgenden richtigen Antworten) mit einem 2-AFC-Verfahren an (Kollmeier et al., 1988, Kollmeier, 1986).

V.3 Psychoakustik des pathologischen Gehörs

-Intensitätsabbildung

Bei sämtlichen Formen der Schwerhörigkeit tritt ein **Hörverlust** auf, d. h. eine Verringerung der Sensitivität des Hörorgans für Schall. Der Pegel eines akustischen Signals muß daher gegenüber Normalhörenden angehoben werden, um eine Empfindung hervorzurufen, d. h. die Hörschwelle ist angehoben. Die Frequenzabhängigkeit der Hörschwelle wird im **Tonschwellenaudiogramm** erfaßt (vgl. 5.4). Bei weiterer Erhöhung des Pegels ergibt sich ein Anstieg der Lautheit, der bei Innenohrschwerhörigen in der Regel steiler ist als bei Normalhörenden (**Recruitment**, vgl. 3.2). Dieser Effekt führt zu dem bei vielen Schwerhörenden beobachteten Phänomen, daß sie bei niedrigen Sprachlautstärken nichts verstehen. Wenn der Gesprächspartner aufgefordert wird, doch etwas lauter zu reden, beklagen sich die Schwerhörigen darüber, daß es schnell zu laut ist und die Sprecher nicht gleich schreien sollen. Dieses Phänomen kann mit **überschwelligen Tests** erfaßt werden. In der Routine-Audiometrie wird bei seitendifferentem Hörverlust dazu der Fowler-Test verwendet (vgl. 5.4).

Den Lautheitsanstieg mit zunehmenden Pegel kann man für jedes Ohr separat mit der Hörflächenskalierung ermitteln, bei der sich bei Patienten mit Recruitment ein steilerer Anstieg der Kategoriallautheit ergibt. Die Steigung dieser Lautheitsfunktion ist jedoch individuell sehr unterschiedlich und kann möglicherweise von dem Verhältnis der Schädigung innerer und äußerer Haarzellen abhängen (vgl. 3.2). Daher ist der **Kompressionsverlust** des auditorischen Systems, der sich in der Steigung der Hörflächenskalierung ausdrückt, der als eine weitere, vom Hörverlust weitgehend unabhängige Komponente zu erfassen. Der Kompressionsverlust nimmt mit zunehmenden Hörverlust zwar tendenziell zu, läßt sich aufgrund der großen individuellen Schwankungen aber nicht daraus vorhersagen (vgl. Kießling, 1995, Launer et al., 1996). Bei Schalleitungs-Schwerhörigkeit ist die Pegel-Lautheitsfunktion der Hörflächenskalierung zu höheren Pegeln hin verschoben. Aufgrund des Adaptationseffektes ist diese Verschiebung jedoch nicht so groß wie der

an der Schwelle meßbare Hörverlust. Außerdem kann auch die Steigung leicht variieren, so daß die Hörfächenskalierung keine sichere Unterscheidungsmöglichkeit zwischen Schalleitungs- und Schallempfindungs-Schwerhörigkeit bietet.

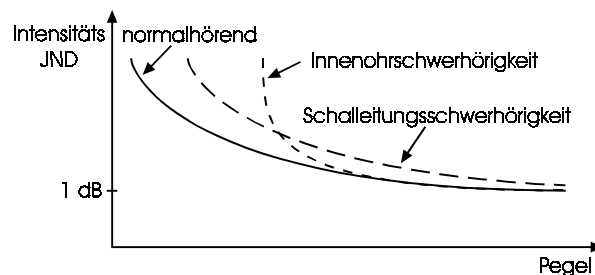


Abbildung 5.18: Intensitäts-JND für verschiedene Schwerhörigkeitsformen

Für die Schalleitungs-Schwerhörigkeit verschiebt sich die Pegelabhängigkeit der **Intensitätsauflösung** (d. h. der kleinste hörbare Intensitätsunterschied, Intensitäts-Just Noticeable Difference, JND) für Sinustöne zu höheren Pegeln (Abb. 5.18). Bei Innenohrschwerhörigen mit Recruitment geschieht die Abnahme der Intensitäts-JND dagegen in einem wesentlich kleineren Pegelbereich. Sie haben bei derselben wahrgenommenen Lautheit ungefähr die gleiche bzw. eine geringgradig erhöhte Intensitäts-JND wie Normalhörende. Die früher verbreitete Annahme, daß Innenohr-Schwerhörige eine kleinere Intensitäts-JND besitzen, weil sie einen steileren Lautheitsanstieg haben, hat sich als falsch herausgestellt (Hohmann, 1993). Wenn man die Intensitäts-Diskrimination allerdings knapp oberhalb der Hörschwelle mißt (z. B. 20 dB wie beim SISI-Test in der überschwelligen Audiometrie, vgl. 5.4), können Innenohrschwerhörige etwas niedrigere Diskriminations-Schwellen erreichen, weil sie dort eine größere Lautheit wahrnehmen als Normalhörende. Allerdings ist diese Diskriminationsschwelle sehr variabel, so daß der SISI-Test ein sehr indirektes Maß für ein Recruitment darstellt, was sich als sehr unzuverlässig herausgestellt hat. Bei breitbandigen Signalen (z. B. Rauschen oder Sprache) hängt die Intensitäts-Diskrimination sehr wenig vom Pegel ab. Innenohrschwerhörige zeigen daher in ihrem Dynamikbereich eine leicht erhöhte Intensitäts-JND gegenüber Normalhörenden, wobei jedoch die gleiche (geringere) Pegelabhängigkeit auftritt. Bei hochgradig Schwerhörigen mit einer sensorineuralen Schwerhörigkeit bzw. bei Patienten mit einer neuralen Schwerhörigkeit kann die Intensitäts-JND relativ stark erhöht sein, so daß es diesen Patienten schwerfällt, akustische Signale (wie z. B. Sprache) in ihrem auditorischen System richtig abzubilden und zu verstehen.

Bei Schalleitungsschwerhörigkeit treten keine nennenswerten Abweichungen in den überschwelligen Funktionen gegenüber Normalhörenden auf, sofern der Darbietungspegel auf die individuelle Hörschwelle bezogen wird. Im folgenden sollen daher nur die bei sensorineuralen Schwerhörigkeiten auftretenden Veränderungen der Frequenz- und Zeitabbildung betrachtet werden:

Aufgrund der bei Schädigung der äußeren Haarzellen verringerten Frequenzselektivität des Gehörs nimmt auch die psychoakustisch meßbare Frequenzauflösung des Gehörs ab. Dies macht sich in einer **Verbreiterung der Frequenzgruppe** bemerkbar, d. h. bei Innenohrschwerhörenden wird ein größerer spektraler Bereich im Gehör zusammengefaßt, um die Lautheit und die Maskierung zu bestimmen (vgl. 5.19). Die Verbreiterung der Frequenzgruppe führt zu einem **Anstieg der Mithörschwelle** im Rauschen, d. h. ein in einem Rauschen versteckter Testton muß stärker im Pegel erhöht werden als bei Normalhörenden, um detektiert werden zu können. Dies liegt daran, daß mehr Energie von benachbarten Frequenzen zur Maskierung des Ziel-Testtons beiträgt. Allerdings hängt diese Verbreiterung vom Signalpegel ab: Auch bei Normalhörenden tritt eine Verbreiterung der Frequenzgruppe mit hohem Pegel auf. Daher sollte man die Frequenzgruppe von Schwerhörenden nicht bei gleichem Pegel über der individuellen Hörschwelle, sondern bei gleichem absoluten Pegel vergleichen. Dabei ist die Frequenzgruppe von Schwerhörenden nur geringgradig gegenüber der von Normalhörenden verbreitert. Bei einer entsprechenden Anhebung des Pegels über die Hörschwelle des Innenohrschwerhörigen (z. B. durch ein Hörgerät) wird daher der Effekt der verbreiterten Frequenzgruppe geringer.

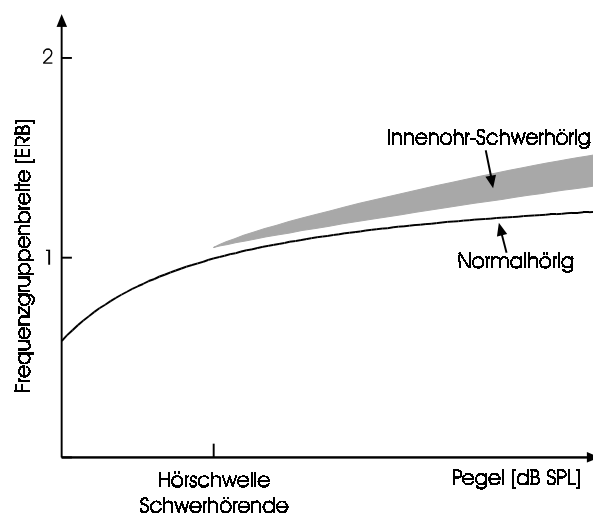


Abbildung 5.19: Frequenzgruppenbreite als Funktion des Pegels für Normal- und Innenohrschwerhörende

Die **Zeitauflösung** bei Innenohrschwerhörigen ist ebenfalls verringert, d. h. man mißt eine verlängerte Vor- und Nachverdeckung und eine Verlängerung der kleinsten detektierbaren Pausendauer im Rauschen (vgl. 5.1.3). Allerdings ist hier genau wie bei der Frequenzauflösung die Pegelabhängigkeit zu berücksichtigen, da bei den Innenohrschwerhörenden ein wesentlich kleinerer Pegelbereich zur Verfügung steht als bei Normalhörenden. Wenn man die Experimente bei dem gleichen Pegel oberhalb der Ruhehörschwelle (bzw. bei der gleichen wahrgenommenen Lautheit) durchführt und auch den Effekt des Recruitments berücksichtigt, ist die Zeitauflösung nicht wesentlich gegenüber der von Normalhörenden gestört. Dies ist auch der Grund, wieso die psychoakustisch meßbare Modulations-Transferfunktion (vgl. 5.1.3) ungefähr den gleichen Tiefpaßverlauf wie bei Normalhörenden aufweist. Allerdings ist die Schwelle insgesamt erhöht, so daß Schwerhörige nicht so geringe Modulationstiefen detektieren können wie Normalhörende.

Im Gegensatz zu den relativ wenig geänderten Zeitkonstanten bei der Zeitauflösung steht das geänderte Verhalten bei der **zeitlichen Integration** bei Innenohrschwerhörigen (vgl. 5.1.3): Während bei Normalhörenden die Wahrnehmbarkeit und die Lautstärke eines Tones mit zunehmender Tondauer bis zu einer Maximaldauer von etwa 200 ms zunimmt, tritt dieser Effekt bei Innenohrschwerhörigen nicht in gleicher Weise auf. Dies ist ein Grund dafür, daß Signale, die im Rauschen versteckt sind (maskierte Signale), von Innenohrschwerhörenden erst bei wesentlich höheren Signal-Rauschverhältnissen detektiert werden können als von Normalhörenden.

-Weitere auditorische Funktionen

Generell zeigt sich bei Innenohrschwerhörenden eine allgemeine Verschlechterung sämtlicher psychoakustisch meßbarer auditorischer Funktionen im Vergleich zu Normalhörenden. Zu diesen eingeschränkten Funktionen gehören neben der oben erwähnten Intensitäts-, Frequenz- und Zeitauflösung auch andersartig gemessene Detektions- und Diskriminationsleistungen im Zeit- und Frequenzbereich (z. B. Tonhöhenunterscheidung), komplexere Erkennungs- und Unterscheidungsleistungen, sowie die binaurale Interaktion (d. h. die Extraktion von Unterschieden zwischen den Signalen, die an den beiden Ohren anliegen), die zur Ortung von Schallquellen und zur Trennung zwischen Nutzsignalen und Störsignalen beiträgt. Dabei stellt man fest, daß sämtliche dieser auditorischen Funktionen primär vom Hörverlust beeinflusst werden. Im Mittel über alle Patienten verschlechtern sich alle Funktionen mit zunehmendem Hörverlust. Abgesehen von diesem

generellen Effekt gibt es jedoch nur sehr geringe gegenseitige Abhängigkeiten zwischen den Schädigungen in der Hörfunktion. Beispielsweise kann die Frequenzauflösung stark gestört sein und das binaurale Hören relativ intakt (oder umgekehrt). Dies führt zu der Vorstellung, daß sich Innenohr-Hörstörungen primär durch die **zwei Komponenten Hörverlust und Kompressionsverlust** charakterisieren lassen. Die darüber hinausgehenden Einschränkungen unterliegen einem diffusen, an unterschiedlichen Orten in verschiedener, nicht vorhersehbarer Weise wirkenden Schädigungsprozeß.

V.4 Ton- und überschwellige Audiometrie

Die am einfachsten durchführbaren audiometrischen Tests benutzen eine **Stimmgabel** mit einer Frequenz von 440 Hz oder 512 Hz, mit der sich eine Schalleitungs- von einer Schallempfindungs-Schwerhörigkeit unterscheiden läßt. Bei dem **Rinne-Versuch** wird die Stimmgabel solange mit der Fußplatte auf das Mastoid (Warzenfortsatz hinter dem Ohr) aufgesetzt bis der Proband angibt, nichts mehr zu hören. Dann wird sie dicht vor das zu testende Ohr gehalten. Wenn der Patient den Ton nicht wieder hört („Rinne-negativ“) ist die Knochenleitung besser als die Luftleitung, was für eine Schalleitungs-Schwerhörigkeit spricht und im umgekehrten Fall („Rinne-positiv“) für eine Schallempfindungs-Schwerhörigkeit. Beim **Weber-Versuch** wird dem Probanden der Stimmgabelfuß auf die Mitte des Kopfes aufgesetzt und der Proband wird gefragt, auf welcher Seite er den Ton wahrnimmt (lateralisiert). Während bei Normalhörigkeit oder symmetrischer Schwerhörigkeit die Lateralisation in der Mitte erfolgt, wird bei asymmetrischer Schalleitungs-Schwerhörigkeit der Schall auf der schlechter hörenden Seite lateralisiert und bei asymmetrischer Schallempfindungs-Schwerhörigkeit auf der besseren Seite.

Eine genauere Aussage erlaubt das **Tonaudiogramm**, das mit einem Audiometer über Kopfhörer (Luftleitung) oder über Knochenhörer (Knochenleitung) gemessen wird, der am Mastoid oder der Stirn aufgesetzt wird. Für Spezialmessungen (wie z. B. Sprachaudiometrie mit oder ohne Hörgeräte) wird auch eine Freifeldmessung durchgeführt, bei der Schall über Lautsprecher abgestrahlt wird. Beim Tonaudiogramm wird bei den audiometrischen Frequenzen (125, 250, 500 Hz 1,2,4,6 und 8 kHz) derjenige Pegel gemessen, bei dem der Proband den Ton soeben wahrnimmt (**Ruhehörschwelle**). Dazu wird meist ein Eingabelungsverfahren verwendet, d. h. der Testtonpegel wird solange erhöht, bis der Proband ihn sicher wahrnimmt, dann wieder erniedrigt, bis er ihn nicht mehr wahrnimmt und dann in kleineren Schritten soweit erhöht und erniedrigt, bis die Schwelle möglichst sicher und reproduzierbar

eingegrenzt ist. Die resultierenden Werte werden in dem in Abb. 5.20 dargestellten Audiogramm-Formular als Hörverlust nach unten aufgetragen, wobei verschiedene Symbole für Luftleitung und Knochenleitung für die linke Seite (kann durch blaue Farbe gekennzeichnet werden) und die rechte Seite (rot) verwendet werden. Die Differenz zwischen Luft- und Knochenleitung („Air-Bone-Gap“) gibt die Schalleitungskomponente des Hörverlusts an.

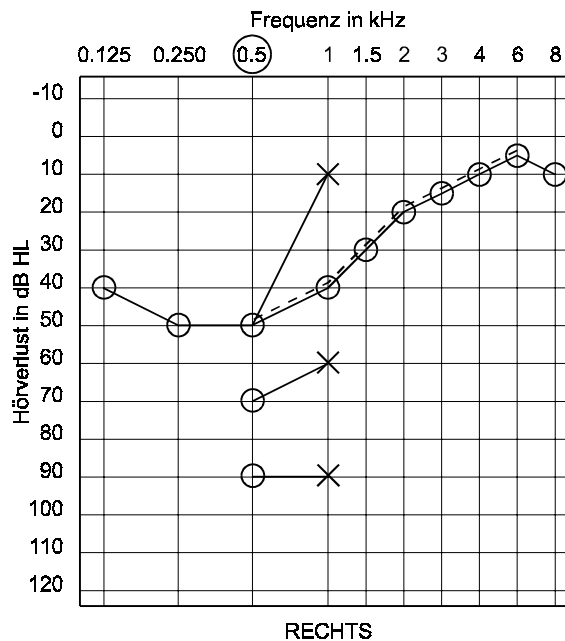


Abbildung 5.20: Audiogrammformular mit Luftleitung (durchgezogene Linie) und Knochenleitungsschwelle (gestrichelt) sowie eingetragendem Fowler-Test bei 500 Hz (aus Lehnhardt, E.: Praxis der Audiometrie. Thieme Verlag, Stuttgart 1996)

Bei der Tonaudiometrie gibt es eine Reihe von Fehlerquellen, so daß die Prüfperson eine gewisse Erfahrung benötigt. Beispielsweise führt ein schlechter Sitz des Kopfhörers zum Abdrücken (Verschluß) des äußeren Gehörgangs mit vorgetäuschter Schalleitungsschwerhörigkeit, oder eine unzureichende Instruktion des Patienten führt zu falschen Schwellen. Bei asymmetrischem Hörverlust ist auf das **Überhören** zu achten, das bei Luftleitung ab etwa 40-50 dB Seitenunterschied auftritt und bei Knochenleitung bereits zwischen 0 und 15 dB. Um es zu vermeiden, muß das in der Knochenleitung bessere Ohr beispielsweise mit einem Schmalbandrauschen vertäubt werden. Bei Knochenleitungs-Messungen wird bei hohen Pegeln die **Fühlschwelle** überschritten, d. h. der Proband registriert die Vibrationen des Vibrators mit der Haut, ohne eine akustische Wahrnehmung zu erhalten. Weitere Fehlerquellen sind die vom Patienten vorgetäuschte Hörstörung (Simulation) bzw. Vergrößerung des Hörverlustes (Aggravation) bzw. Verkleinerung des Hörverlustes (Dissimulation). Zudem kann bei neuralen Schwerhörigkeiten ein Schwellenschwund auftreten, so daß die zu Anfang gemessenen Hörschwellen bei

niedrigeren Pegeln liegen als die im späteren Verlauf der Untersuchung gemessenen Werte. Zu näheren Einzelheiten sei auf die Lehrbücher der Audiometrie verwiesen (z. B.: Böhme, G., Welzl-Müller, K.: Audiometrie. Hans Huber Verlag, Bern 1993 oder Lehnhardt, E.: Praxis der Audiometrie. Thieme Verlag, Stuttgart 1996).

Mit den **überschwelligen tonaudiometrischen Tests** werden Eigenschaften des krankhaften Innenohres oder Hörnerven erfaßt. Insbesondere wird das Vorliegen eines Recruitments als Zeichen für eine Innenohrschädigung ausgemessen. Bei seitendifferentem Hörverlust (Knochenleitungs-Differenz von > 20 dB) kann der **Fowler-Test** durchgeführt werden. Dazu wird der Pegel am „besseren“ Ohr in 20 dB-Schritten erhöht und der Pegel am „schlechteren“ Ohr wird soweit erhöht, bis der Proband subjektiv die gleiche Lautheit an beiden Ohren wahrnimmt. Die zu gleichem Lautheitseindruck führenden Pegel werden miteinander verbunden (vgl. Abb. 5.20). Bei hohen Pegeln stellt sich bei Vorliegen eines Recruitments ein Lautheitsausgleich dar, d. h. derselbe (hohe) Pegel ist zum Hervorrufen des gleichen Lautheitseindrucks auf beiden Ohren notwendig. Ein direkteres überschwelliges Verfahren zur Bestimmung des Recruitments ist die Hörflächenskalierung (vgl. 5.3), die auch bei symmetrischem Hörverlust durchgeführt werden kann. Ebenfalls für den Nachweis eines Recruitment geeignet ist das „Metz-Recruitment“ d. h. der verringerte Abstand zwischen der Hörschwelle und der Stapedius-Reflex-Schwelle aus der Impedanzaudiometrie. Wenn dieser Abstand kleiner als 60 dB ist, spricht dieses für das Vorliegen eines Recruitments.

Der ebenfalls überschwellig tonaudiometrisch angewandte **SISI-Test** (Short Increment Sensitivity Index) wird 20 dB oberhalb der Hörschwelle durchgeführt und beinhaltet 20 sprungartige Pegelanstiege von 1 dB (bzw. größere Pegelanstiege in der Übungsphase). Das Ergebnis wird als der Prozentsatz vom Probanden richtig wahrgenommener Pegelsprünge angegeben, der bei Normalhörenden sehr niedrig und bei Patienten mit Recruitment sehr hoch (80 bis 100 Prozent) sein sollte. In der Praxis ist dieser Test jedoch nicht sehr sensitiv und daher umstritten (vgl. 5.3).

Zur Aufdeckung einer etwaigen **neuralem Komponente** des Hörverlusts wird der tone-decay-Test durchgeführt, bei dem der Testton in der Nähe der Hörschwelle solange angeboten wird, bis der Proband angibt, ihn nicht mehr zu hören. Darauf wird der Pegel erhöht. Eine **pathologische Hörermüdung** wird dadurch nachgewiesen, daß die Schwellenabwanderung zu groß wird. Einen ähnlichen Hinweis auf eine pathologische Hörermüdung (bei zugrundeliegender neuraler Schwerhörigkeit) bietet der **Reflex-decay**, bei dem die Antwort des

Stapediusreflexes auf längere Beschallung mit einem um 10 dB über der Reflexschwelle liegenden Testton registriert wird. Wenn die Änderung der Mittelohrimpedanz zu hoch ist (Abfall des Reflexes) spricht dies ebenfalls für eine neurale Schwerhörigkeit.

Neben den rein tonaudiometrischen Tests stehen **sprachaudiometrische Testverfahren** (vgl. 6.2) und „zentrale“ Tests zur Bestimmung integrativer Leistungen des auditiven Systems zur Verfügung (z. B. Richtungshörtest, binauraler Summationstest, binaural alternierende Sprache, Feldmann-Test mit dichotischer Sprache). Außerdem stehen die in Kap. 2. bereits erwähnten Verfahren der Impedanz-Audiometrie (Messung der akustischen Impedanz im äußeren Gehörgang bzw. des Einfluß des akustischen Reflex darauf), und der Registrierung von Hirnstammpotentialen und otoakustischen Emissionen zur Verfügung. Zur Einordnung ihrer diagnostischen Wertigkeit sei auf Lehrbücher der Audiometrie verwiesen (z. B. Böhme und Welzl-Müller, 1993, Lehnhardt et al., 1996).