

# Geräuschunterdrückung und Klassifizierung der akustischen Umgebung

Beamformer

Auditorische Szenenanalyse

# Das digitale Hörgerät

Technologische Möglichkeiten werden laufend erweitert.  
Wie können die Hörbehinderten davon profitieren?

- Psychoakustische Modelle als Grundlage komplexer Signalverarbeitung für Hörgeräte
- Fortschritte in der Störgeräuschreduktion mit Ein- und Mehrmikrofonsystemen
- Binaurale Algorithmen zur verbesserten Lokalisation und Geräuschunterdrückung
- Wahrnehmung und Klassifizierung auditorischer Objekte (automatische Programm-Selektion)

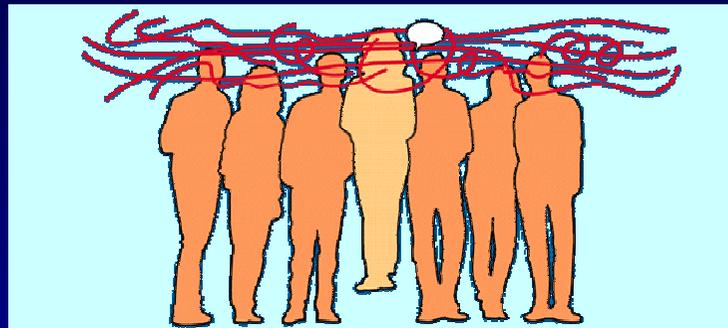
# Störschallunterdrückung

- Abschwächung konstanter Hintergrundgeräusche
- Aktivierung von Richtungsmikrofonen im Störschall

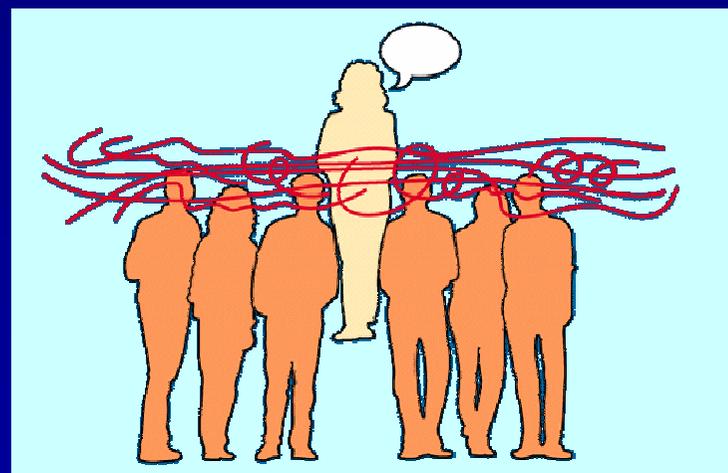


# Schallvereinfachung

## Verständlichkeitsschwelle

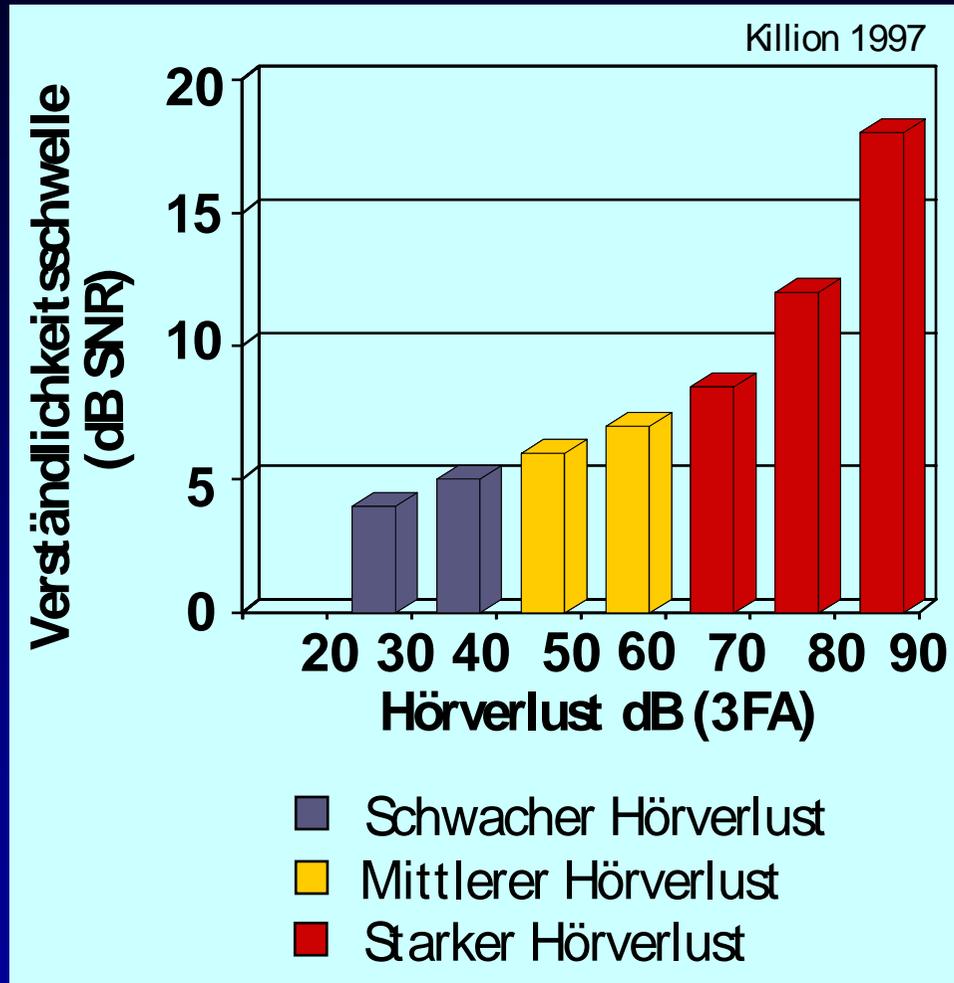


**Intaktes Gehör:**  
niedriger SNR

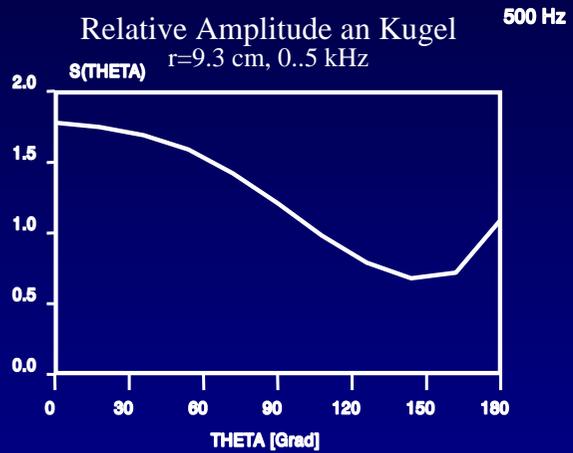


**Cochleäre  
Schwerhörigkeit:**  
erhöhter SNR

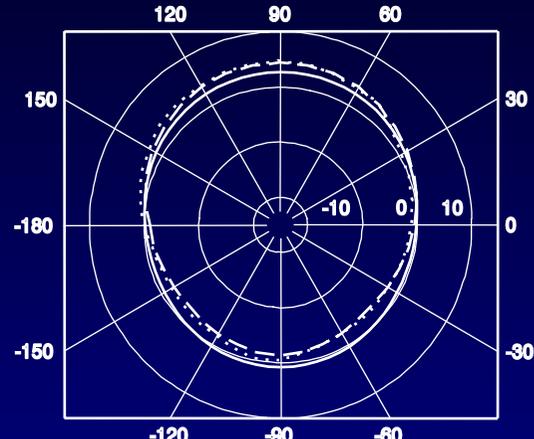
# Schallvereinfachung



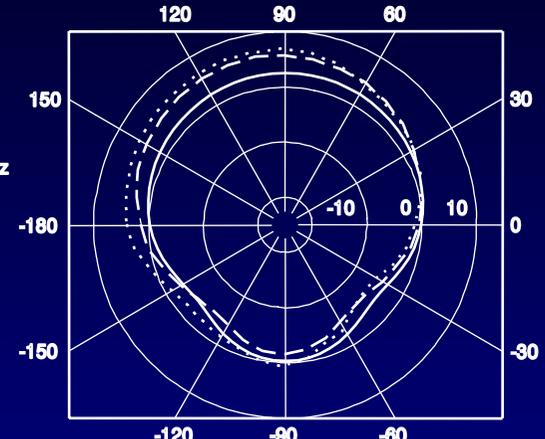
# Kopfschattenmodell (Kugel)



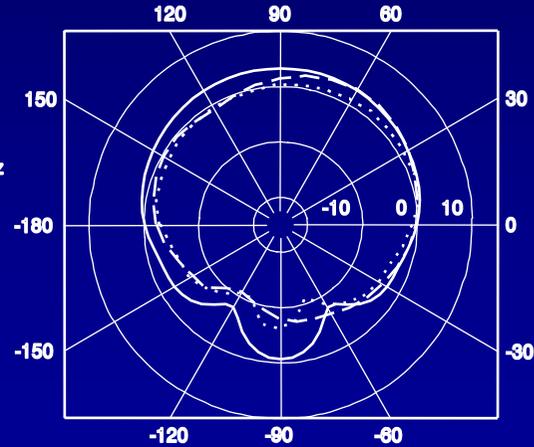
500 Hz



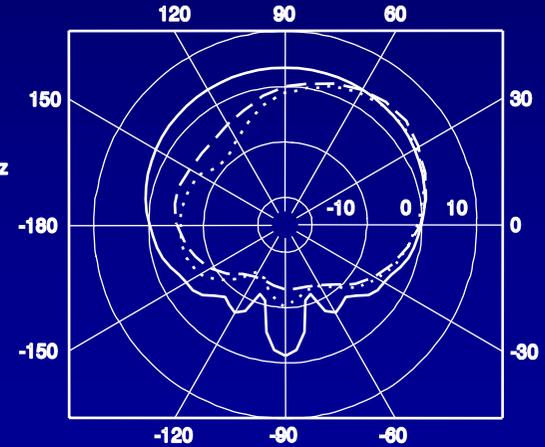
1000 Hz



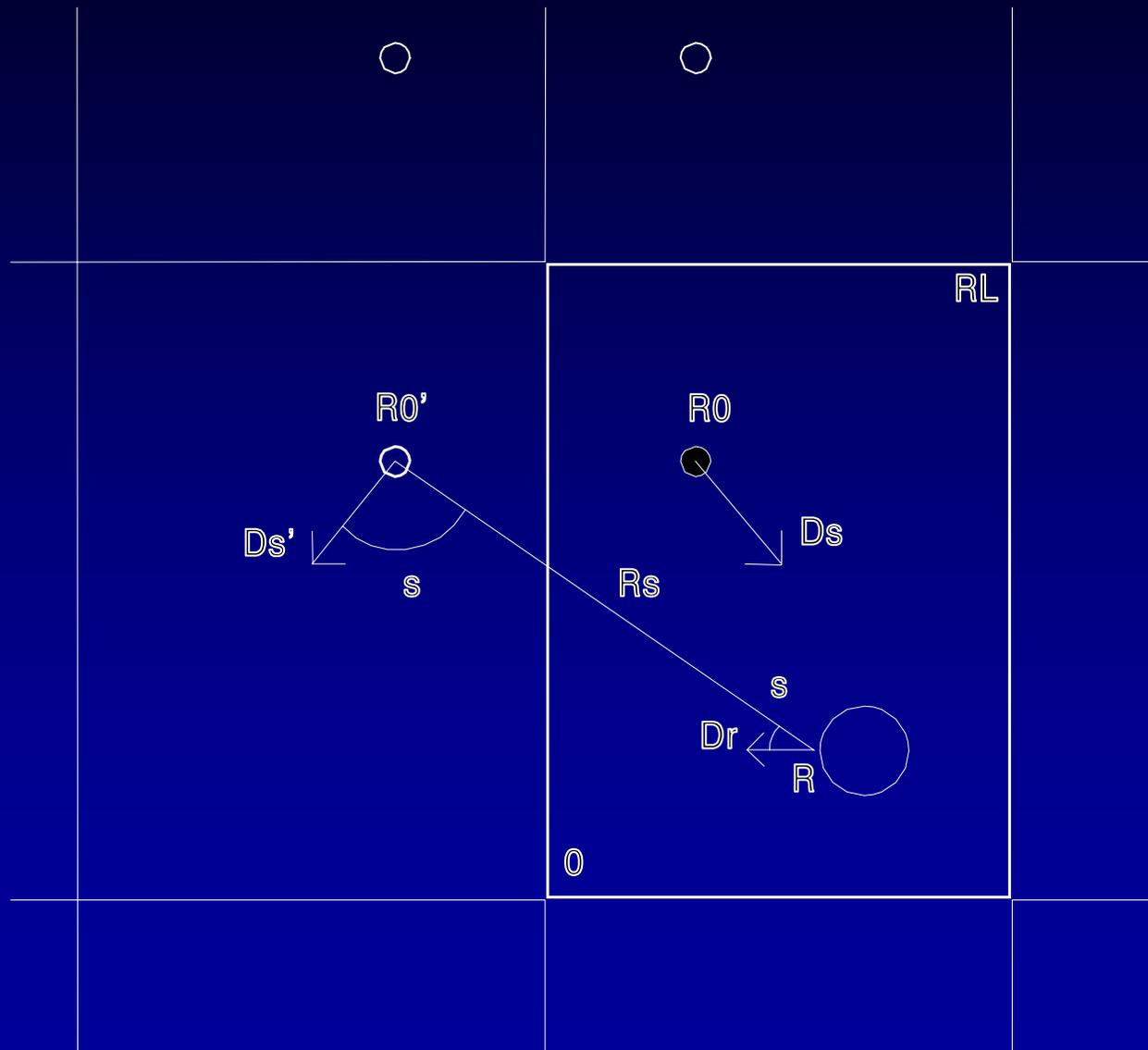
2000 Hz



4000 Hz

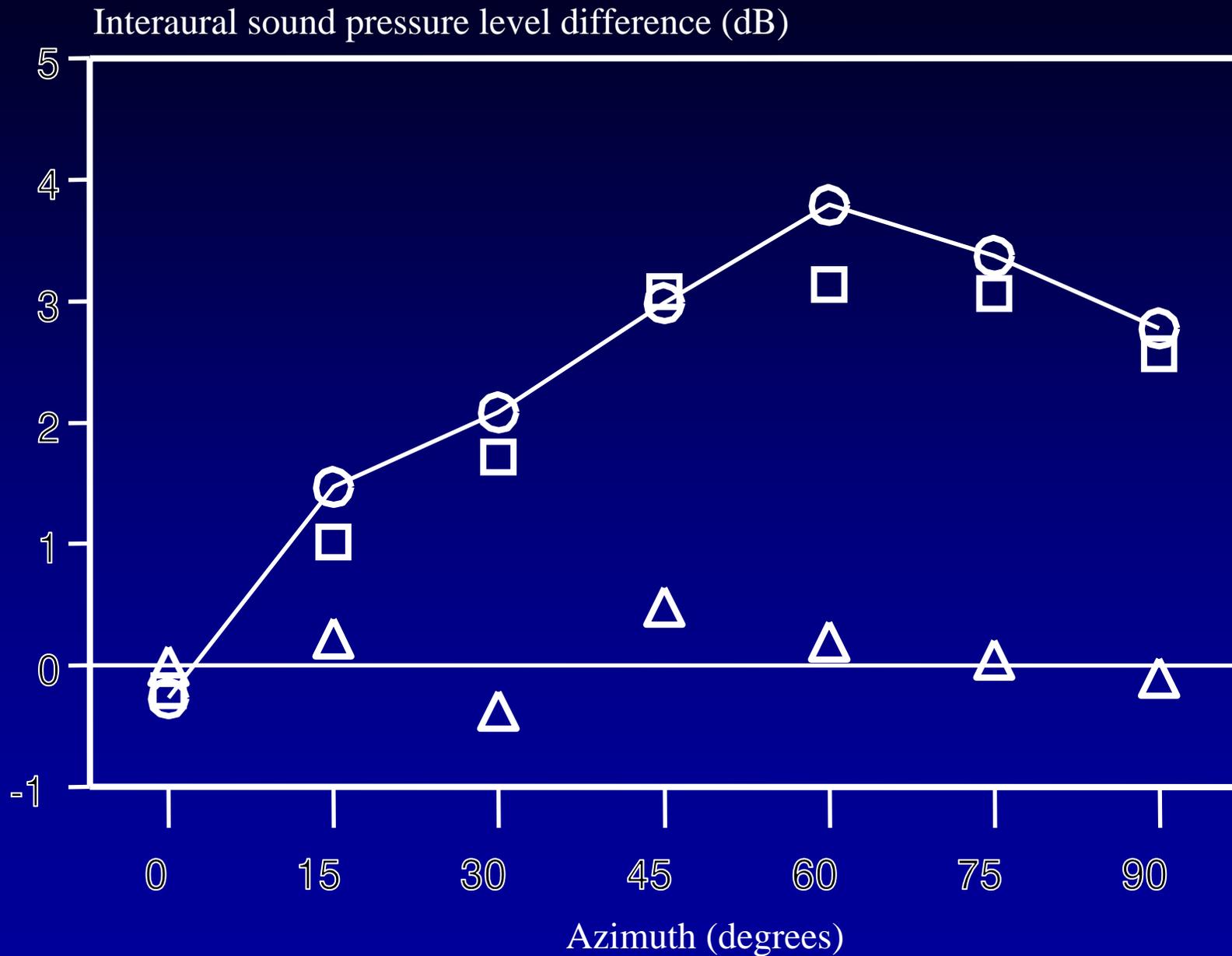


# Beispiel einer Reflexionsberechnung



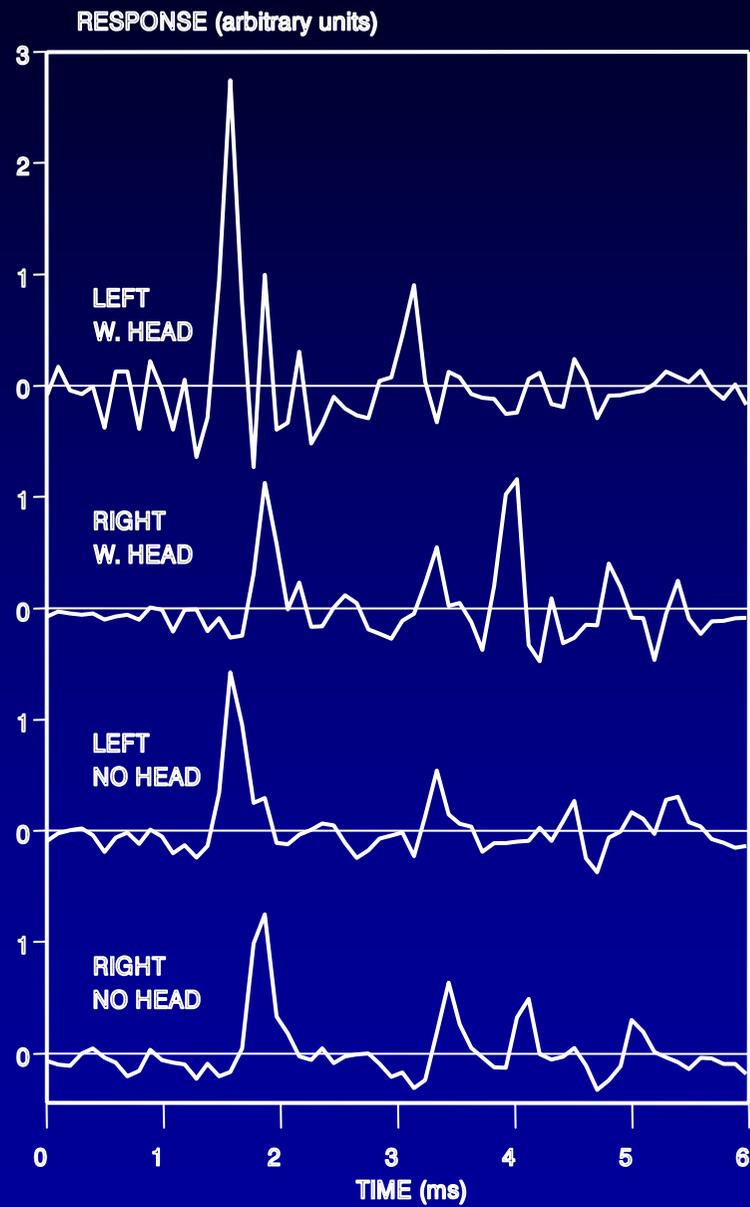
Störgeräuschunterdrückung (Dillier: Vorlesung Medizinische Akustik)

# Kopfschatten: Messung und Modell

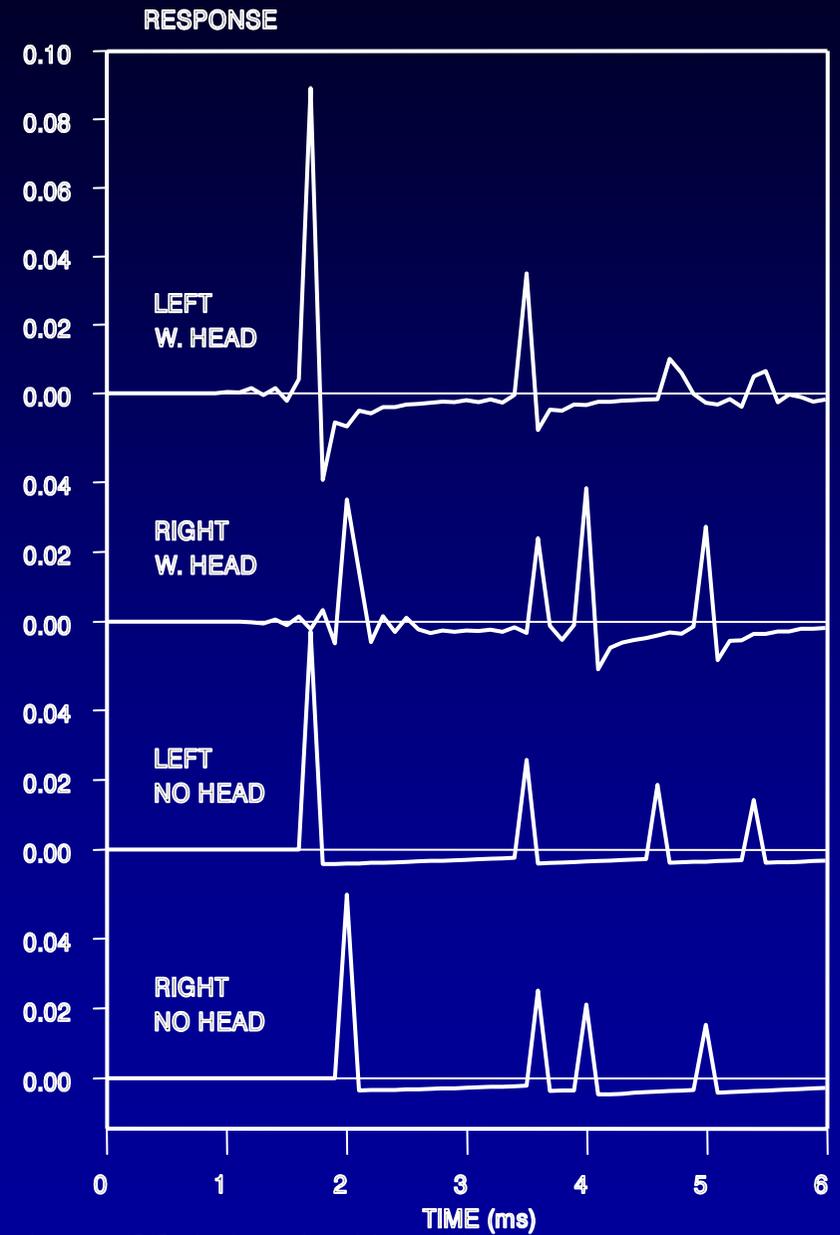


Störgeräuschunterdrückung (Dillier: Vorlesung Medizinische Akustik)

# Gemessene Transferfunktionen



# Modellierte Transferfunktionen



Störgeräuschunterdrückung (Dillier: Vorlesung Medizinische Akustik)

# Schallvereinfachung

## Ziele

Hervorhebung von Sprachschallen

Absenkung von Nebengeräuschen

## Strategien

Erhöhung des SNR

Reduktion des Störschalls

## Mittel

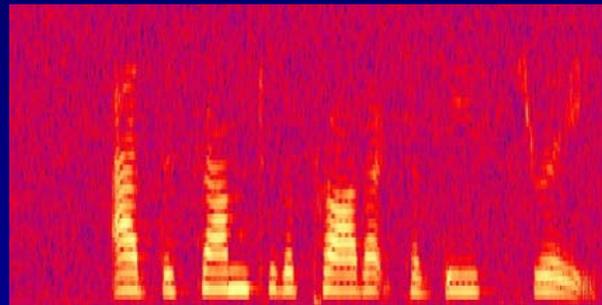
- Unterdrückung quasi-stationärer Geräusche
- Richtmikrofon-Technologie
- FM-Technologie
- Schall-Enthallung
- Automatische Aktivierung der Störschallunterdrückungsfunktionen

# Schallvereinfachung

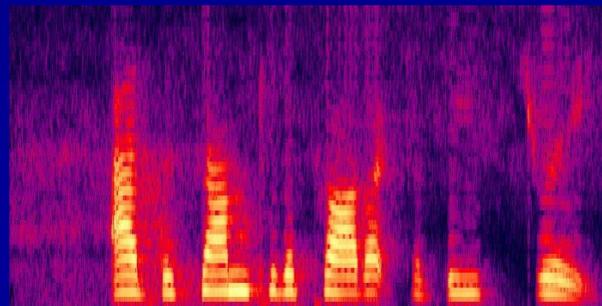
Unterdrückung quasi-stationärer Geräusche

- Verändert die Verständlichkeitsschwelle kaum, reduziert aber die Höranstrengung.

Beispiel: Sprache in gleichmässigem Geräuschhintergrund



Original

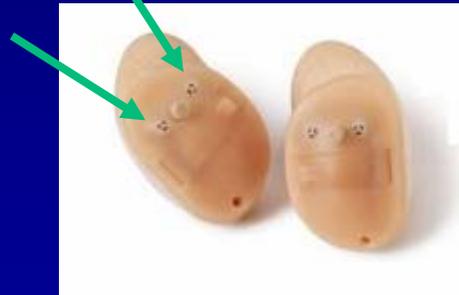


Mit Geräuschunterdrückung

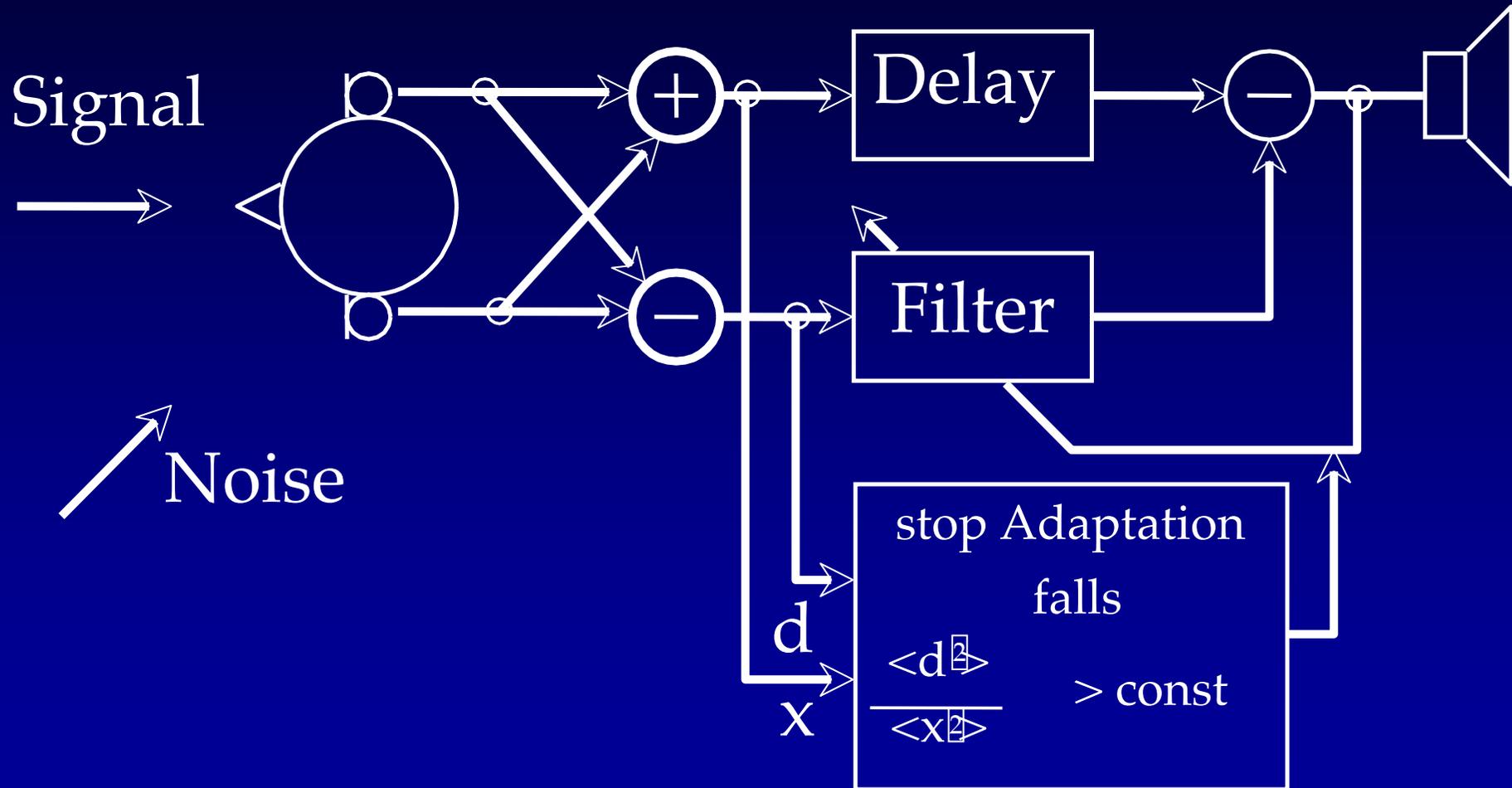
# Schallvereinfachung

## Richtmikrofon-Technologie

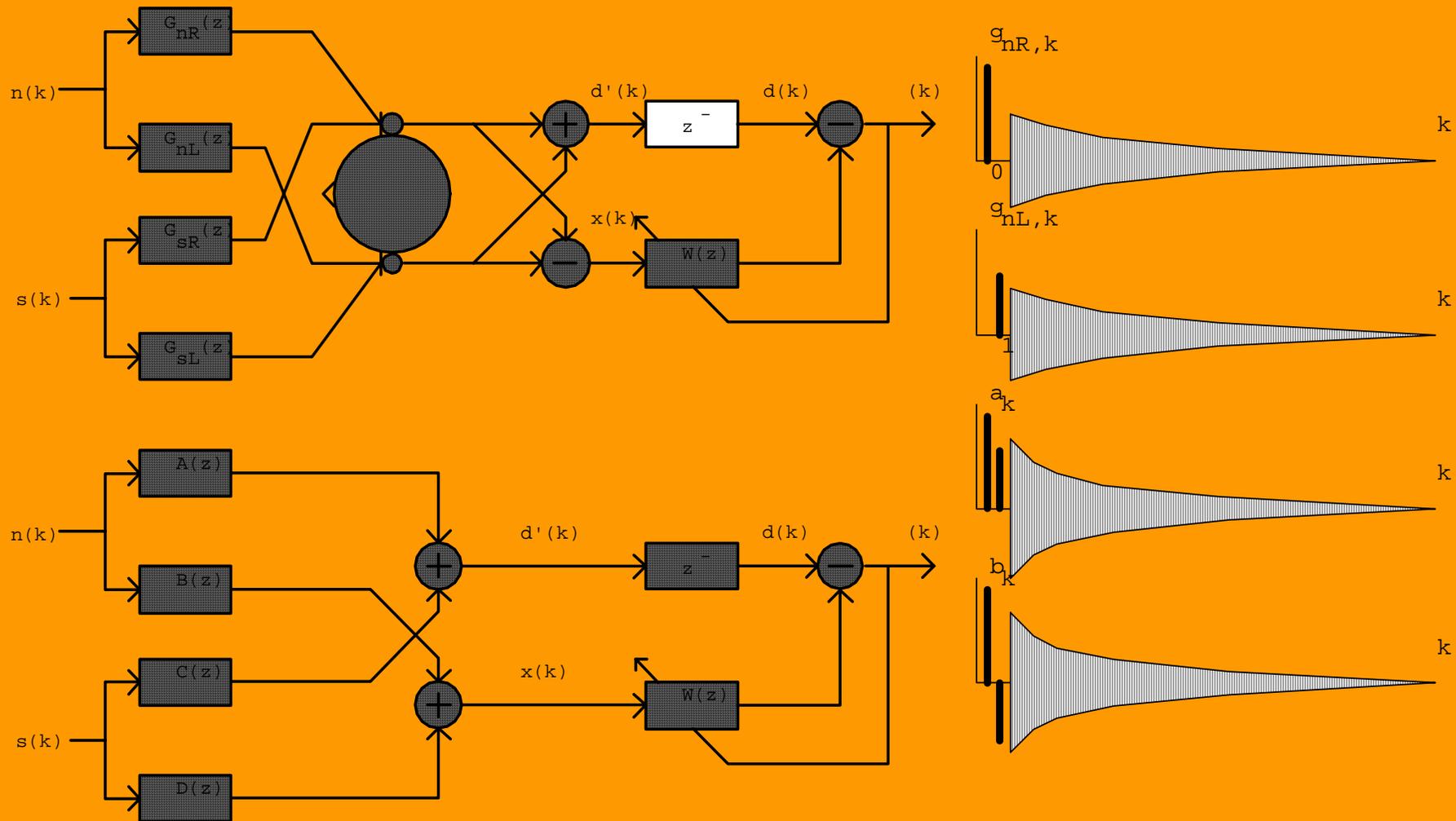
- Unterdrückung des Störschalls, der von der Seite und von hinten kommt.
- Verbessert den SNR deutlich.



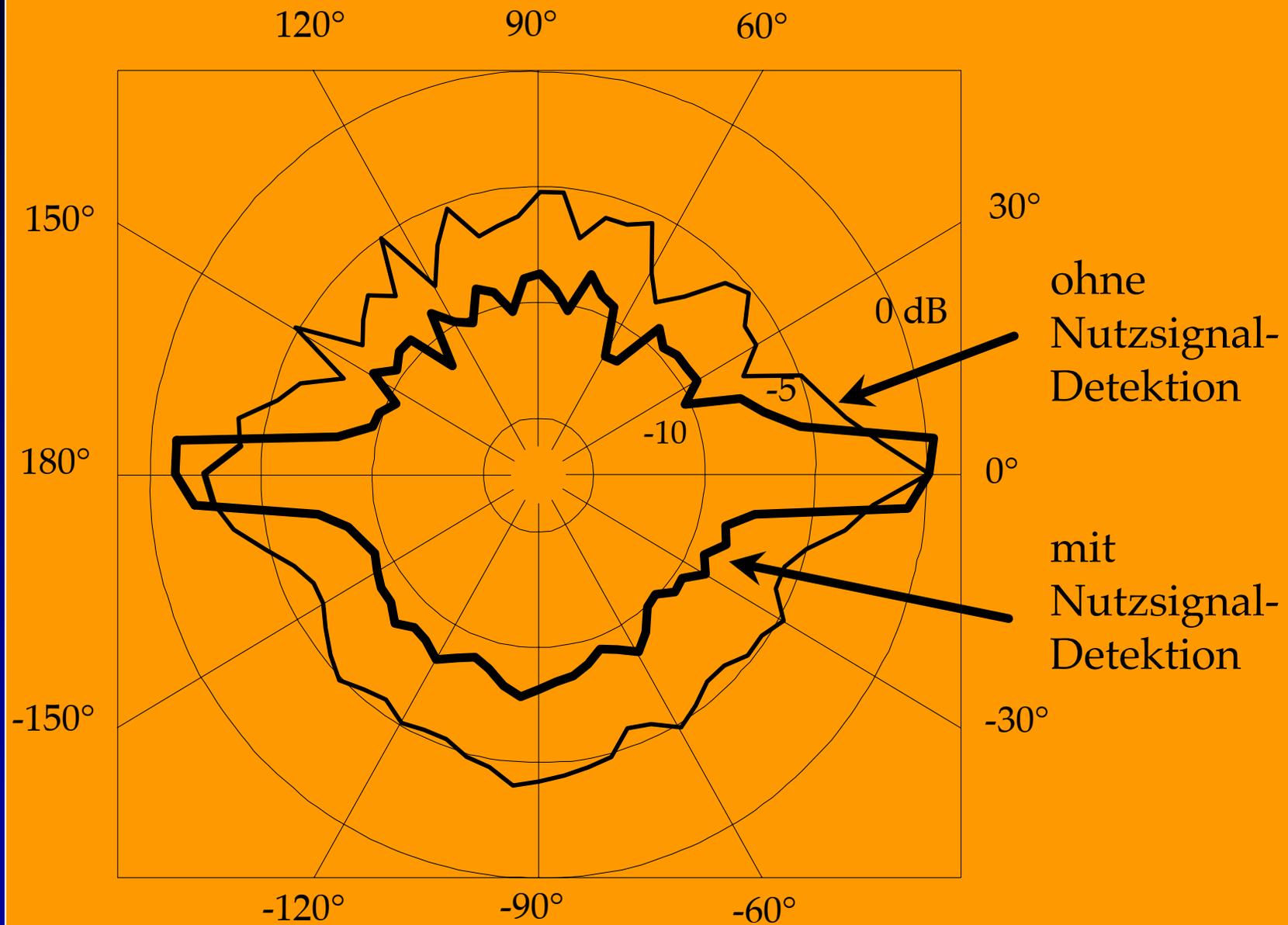
# Beamformer-Prinzip-Schema



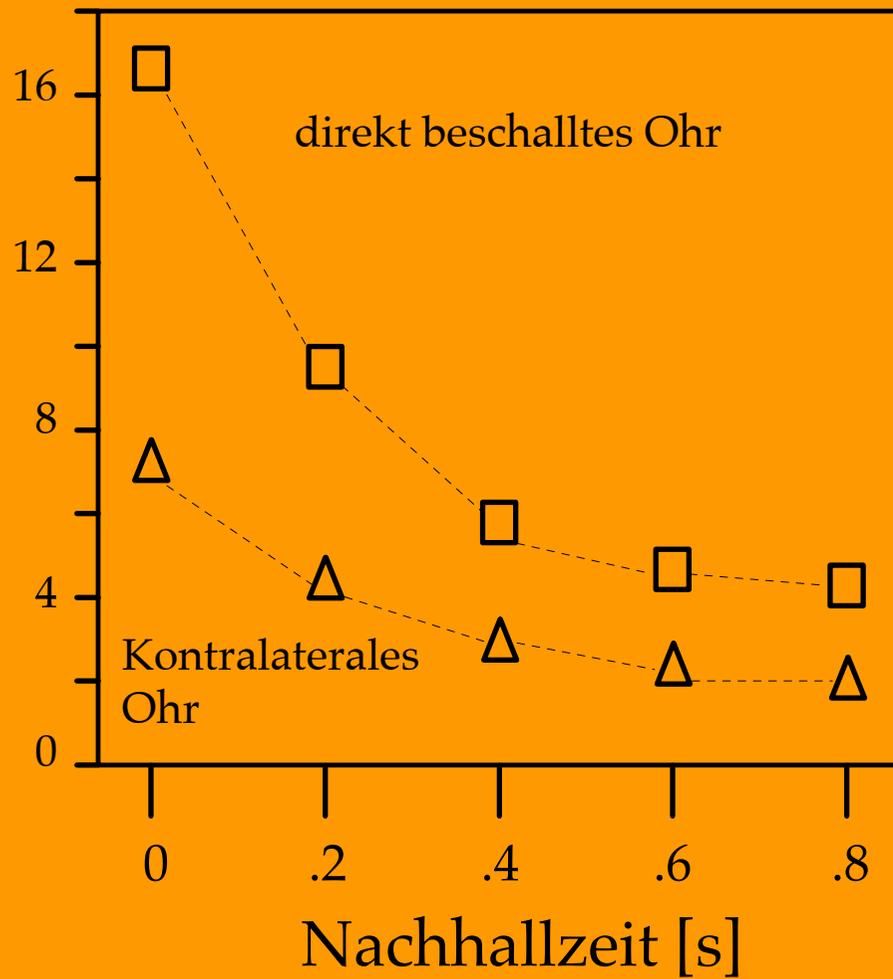
# Adaptiver Beamformer und Raum-Uebertragungsfunktionen



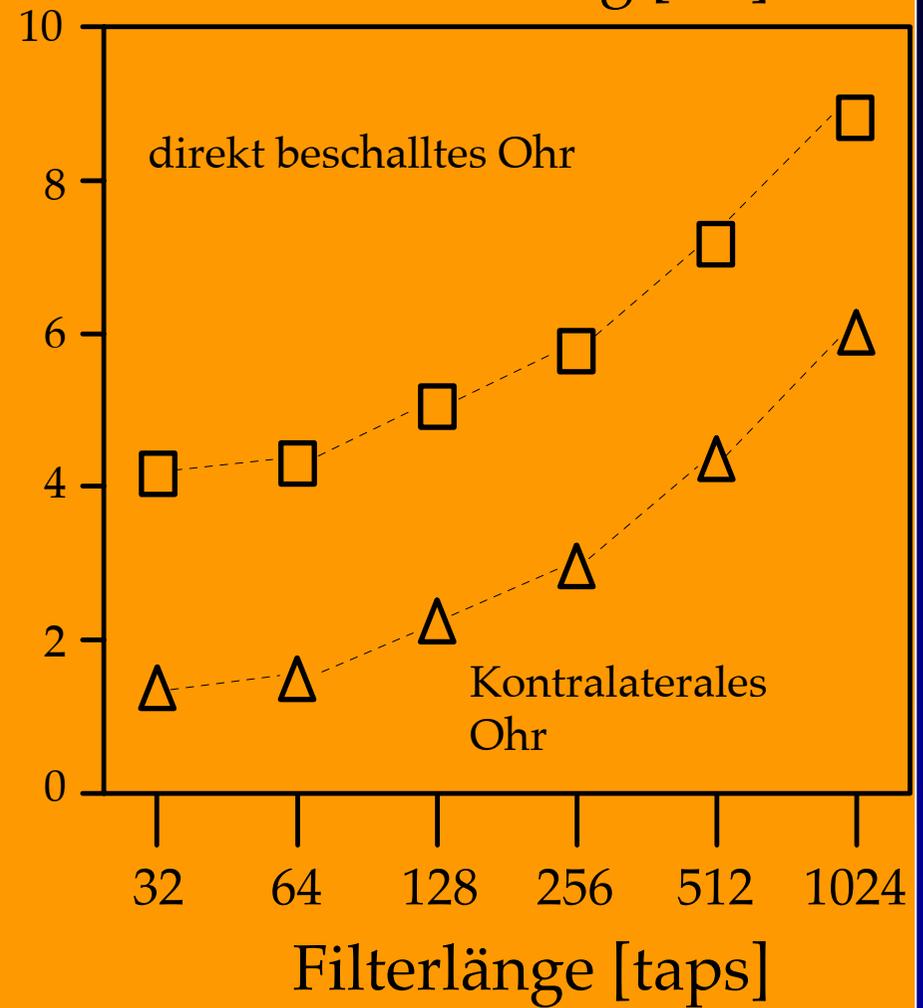
# Direktivitäts-Muster



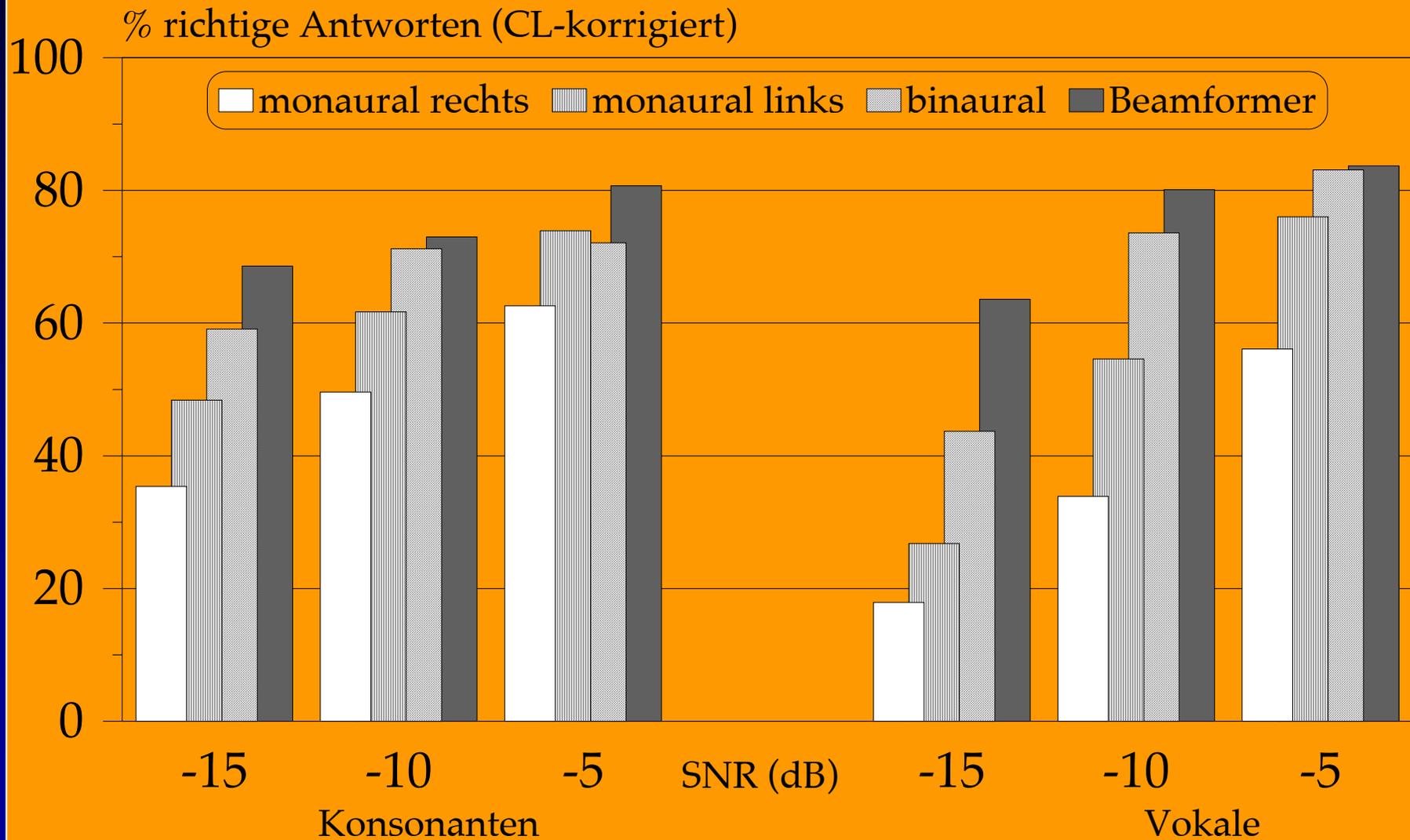
### SNR Verbesserung [dB]



### SNR Verbesserung [dB]



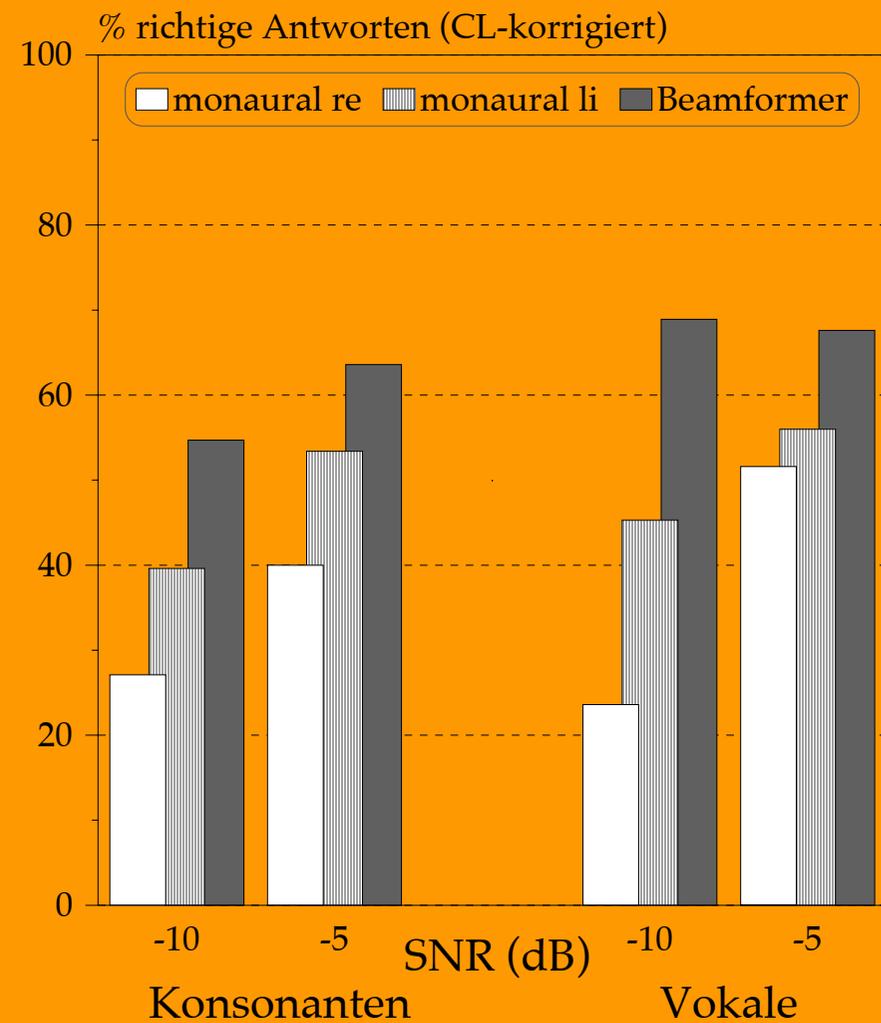
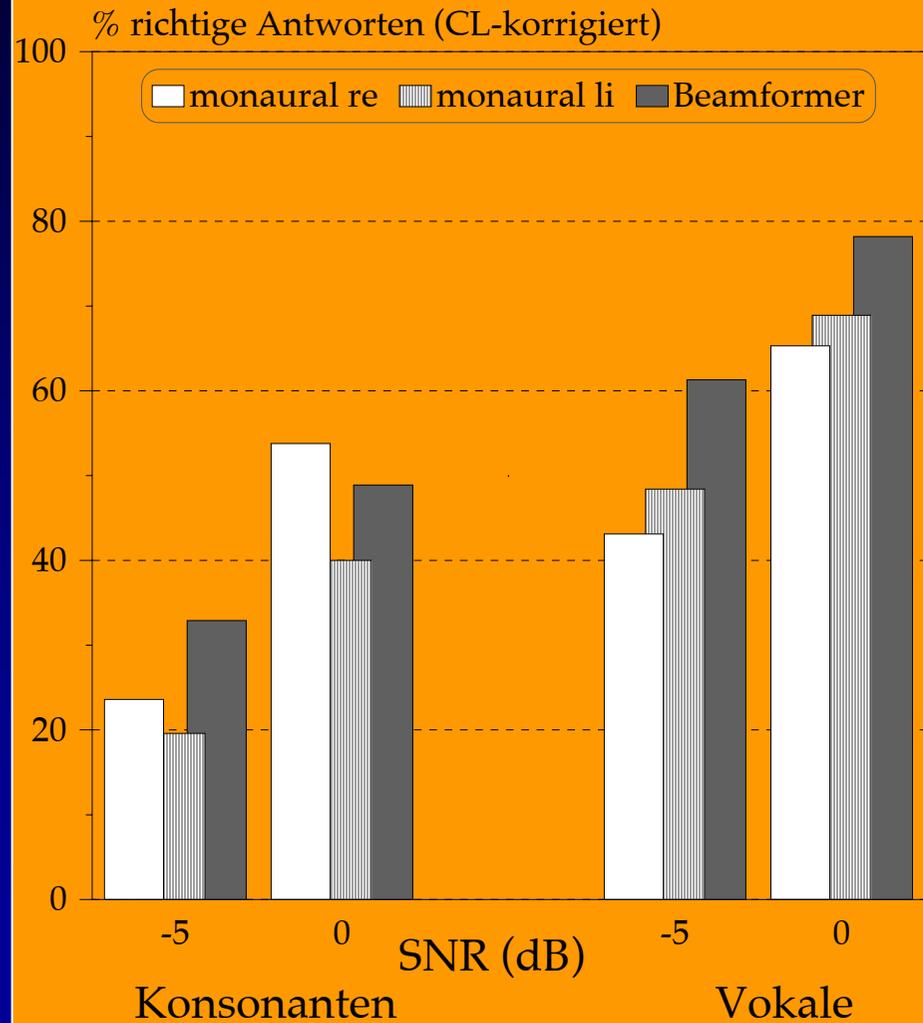
# Normalhörende Versuchspersonen (n=9)



# Hörbehinderte Versuchspersonen (n=6)

## Cafeteria Noise

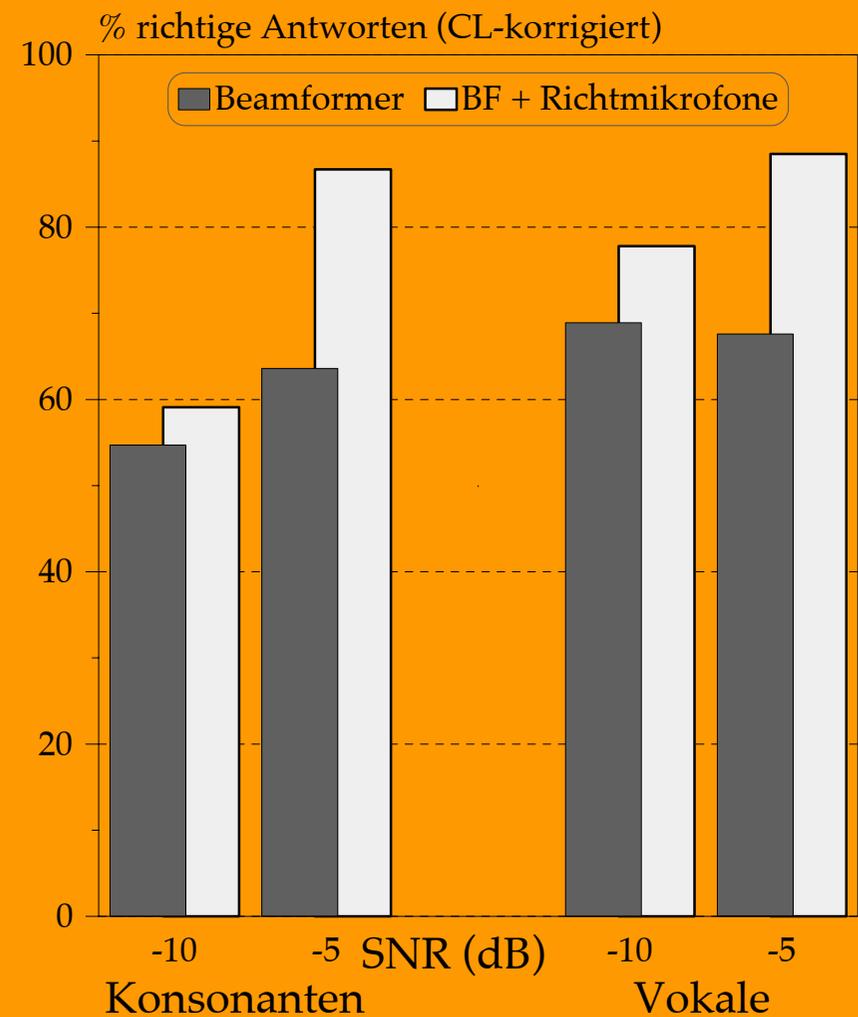
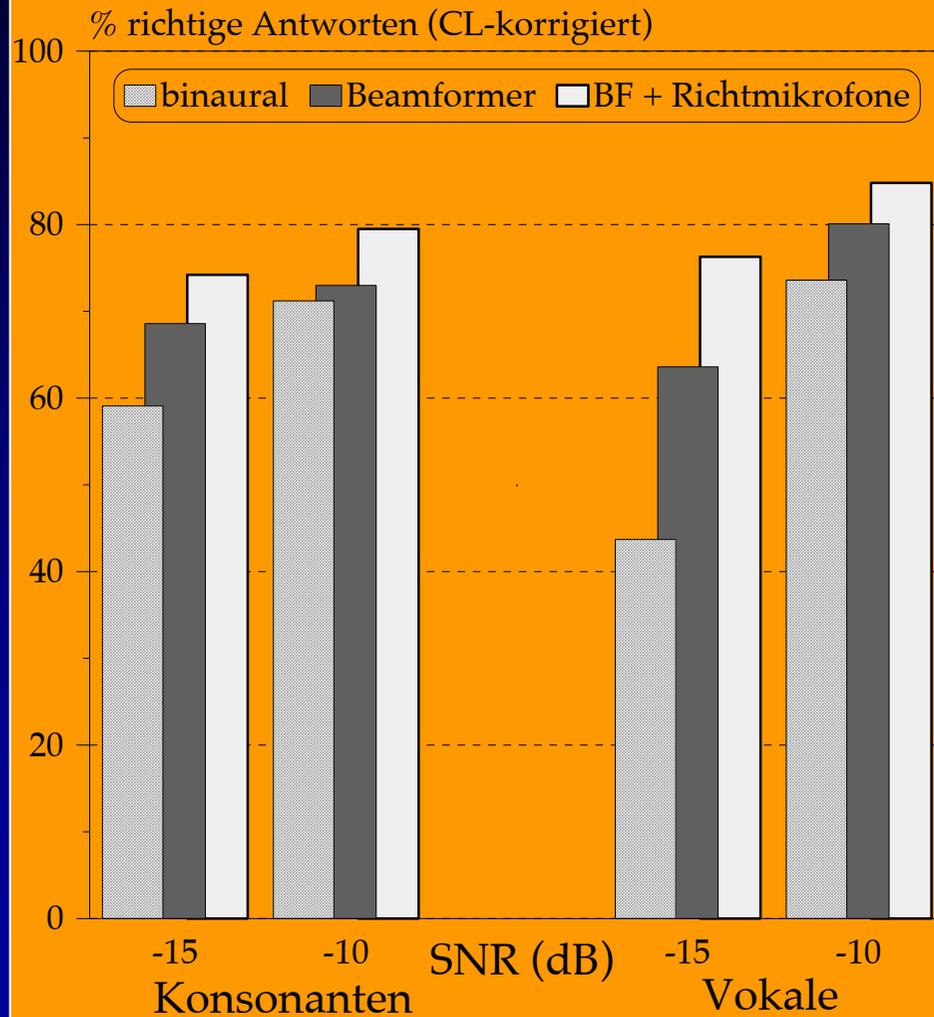
## Speech Spectrum Noise



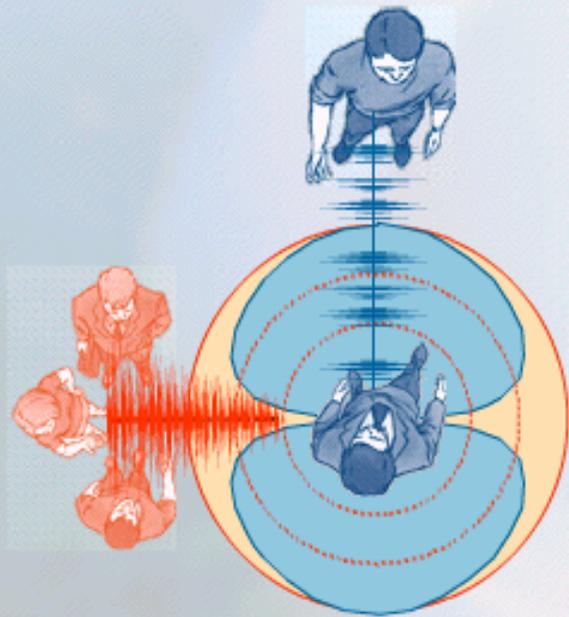
# Einfluss von Richtmikrofonen

Normalhörende Versuchspersonen (n=9)

Hörbehinderte Versuchspersonen (n=6)



|dAZ|  
Adaptive digital Audio Zoom

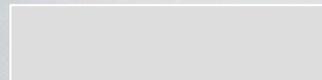


## Adaptive digital AudioZoom

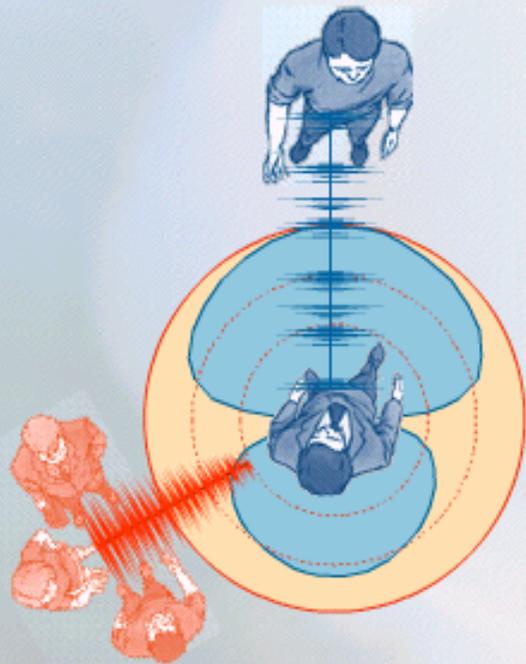
### Ortung des Störgeräusches

dAZ sucht die lauteste Störquelle und passt die Richtcharakteristik automatisch zur bestmöglichen Störgeräuschunterdrückung an

Unabhängig davon, ob sich der Zuhörer bewegt oder sich die Störgeräuschquelle ändert



|dAZ|  
Adaptive digital Audio Zoom



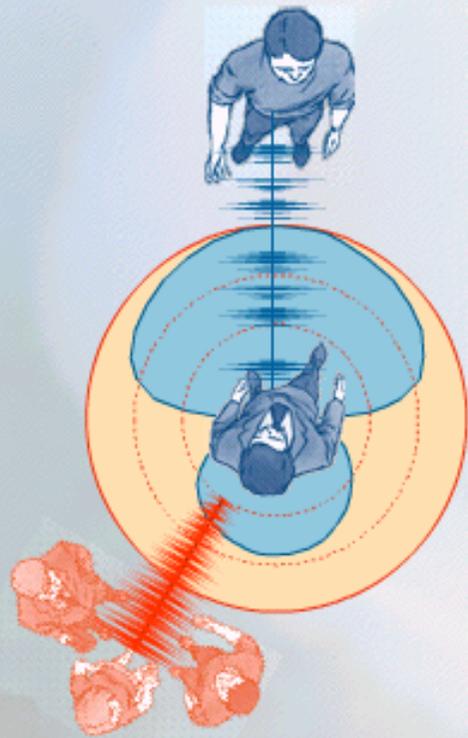
## Adaptive digital AudioZoom

### Ortung des Störgeräusches

dAZ sucht die lauteste Störquelle und passt die Richtcharakteristik automatisch zur bestmöglichen Störgeräuschunterdrückung an

Unabhängig davon, ob sich der Zuhörer bewegt oder sich die Störgeräuschquelle ändert

|dAZ|  
Adaptive digital Audio Zoom

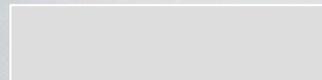


## Adaptive digital AudioZoom

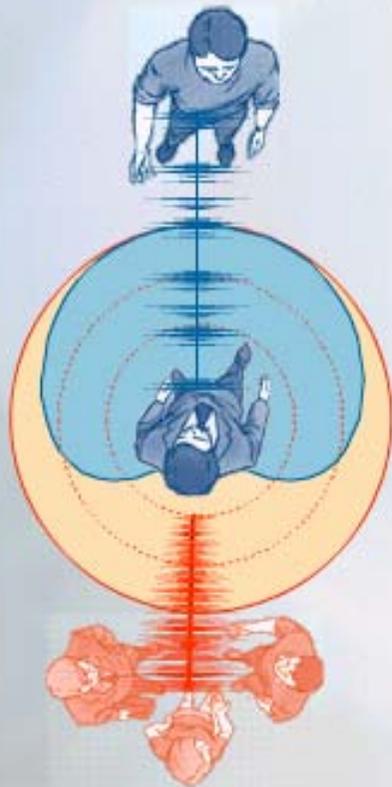
### Ortung des Störgeräusches

dAZ sucht die lauteste Störquelle und passt die Richtcharakteristik automatisch zur bestmöglichen Störgeräuschunterdrückung an

Unabhängig davon, ob sich der Zuhörer bewegt oder sich die Störgeräuschquelle ändert



|dAZ|  
Adaptive digital Audio Zoom

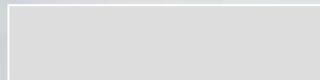


## Adaptive digital AudioZoom

### Ortung des Störgeräusches

dAZ sucht die lauteste Störquelle und passt die Richtcharakteristik automatisch zur bestmöglichen Störgeräuschunterdrückung an

Unabhängig davon, ob sich der Zuhörer bewegt oder sich die Störgeräuschquelle ändert



# Automatische Multi-Programm-Strategien - Nutzen im Alltag

- Angenehme Sprachwahrnehmung in unterschiedlichen Situationen erfordert unterschiedliche Signalverarbeitung
  - Sprache in Ruhe:
    - ⇒ omni-direktional, grosser Frequenzbereich
  - Sprache in Störlärm
    - ⇒ Richtungsmikrofon, reduzierter Tieftonbereich
- Störgeräuschunterdrückung (Dillier: Vorlesung Medizinische Akustik)

# AutoSelect-Funktion

~~Akustische  
Umgebung?~~

~~Passendes  
Hörprogramm?~~

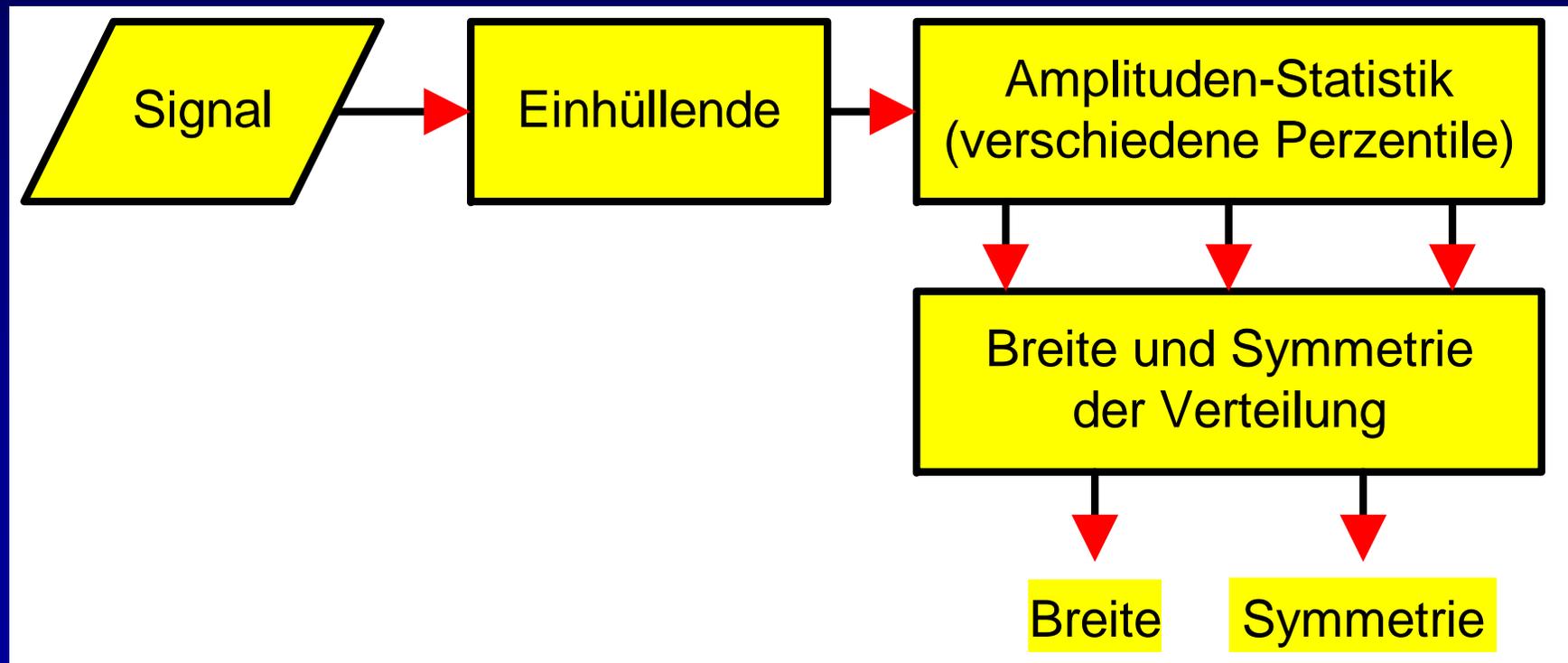


AutoSelect

Analyse der  
akustischen Umgebung  
Auf optimales  
Hörprogramm schalten

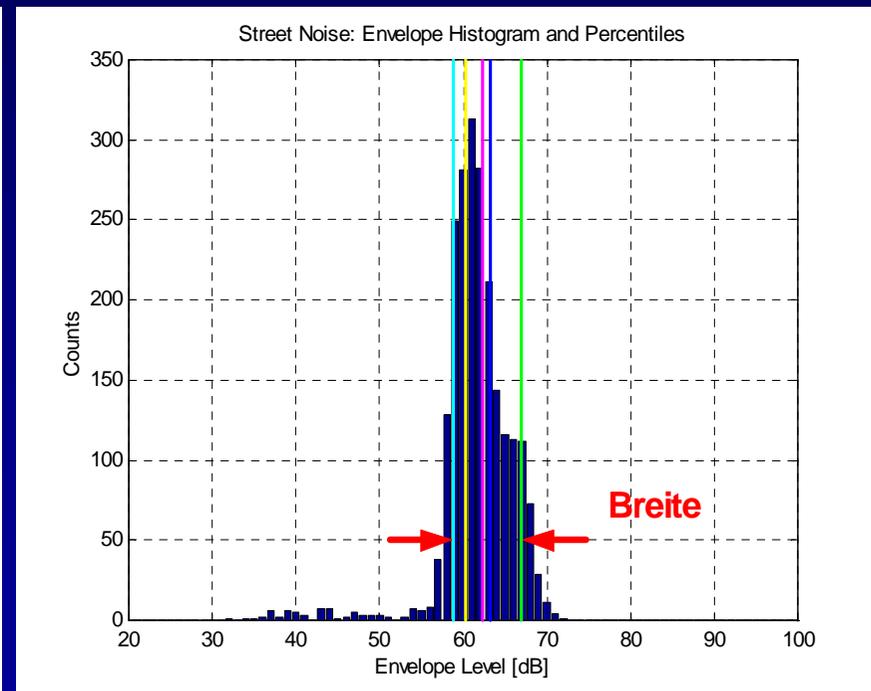
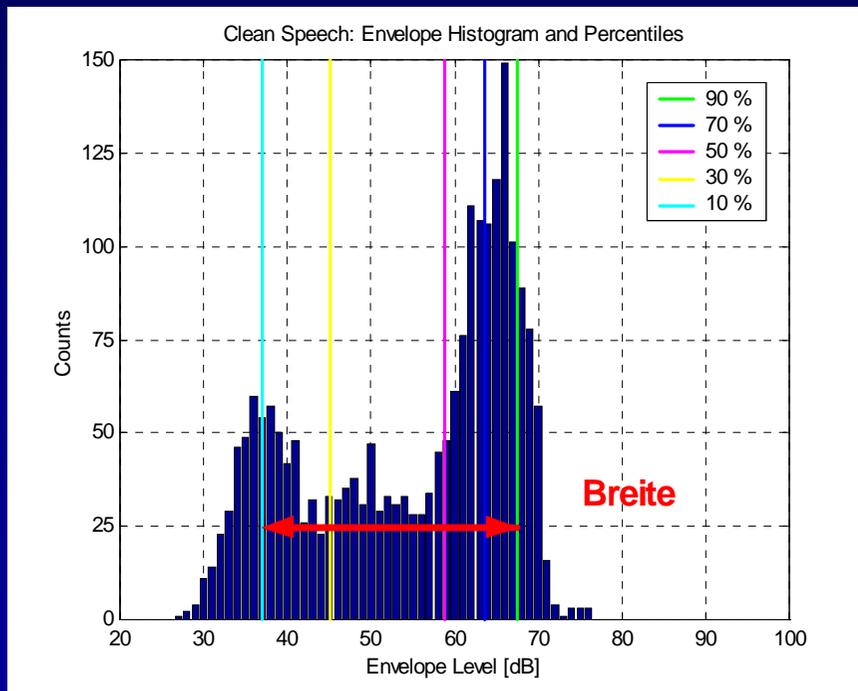
# Amplitudenstatistik

- **Merkmale:** - *Breite & Symmetrie der Verteilung*
- **Klassifizierung:** - *Sprachsignale eindeutig*



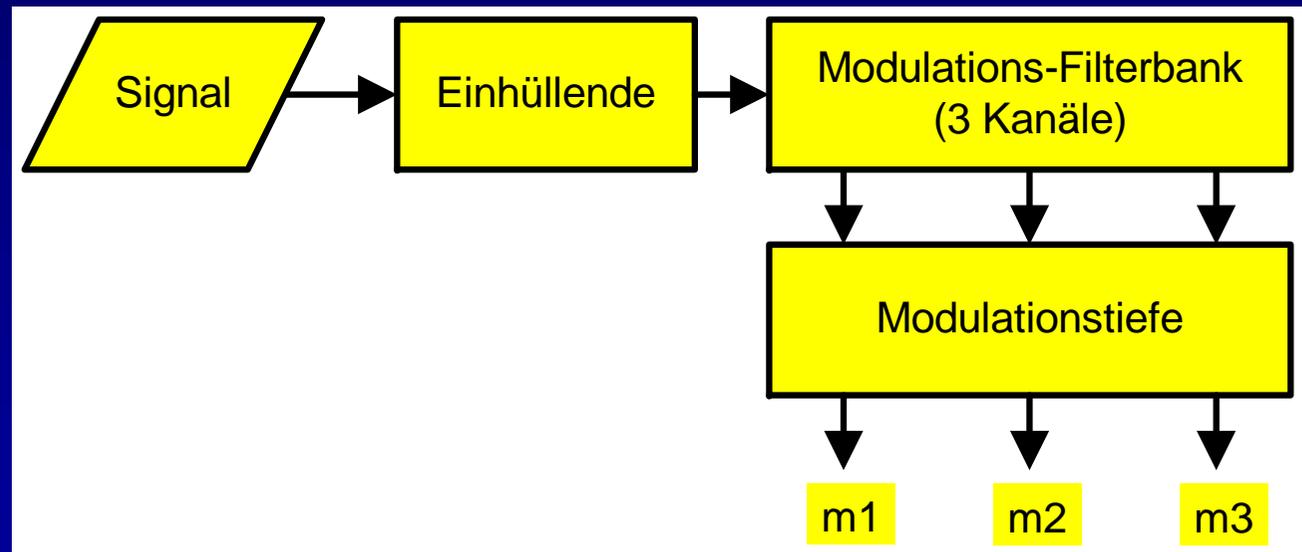
# Amplitudenstatistik

- Beschreibung der Amplitudenverteilung der Einhüllenden durch Breite und Symmetrie. Die Breite der Verteilung ist ein Mass für die zeitlichen Schwankungen des Signals.



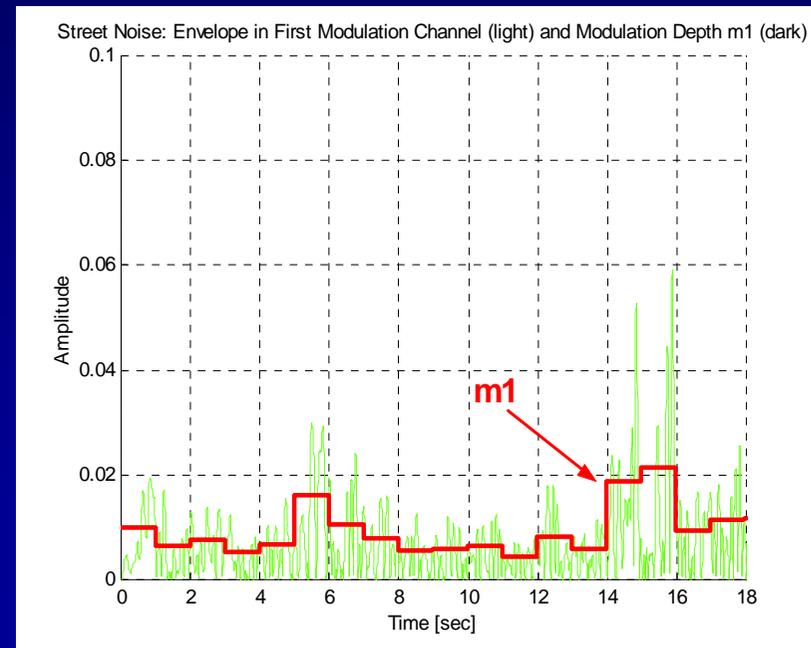
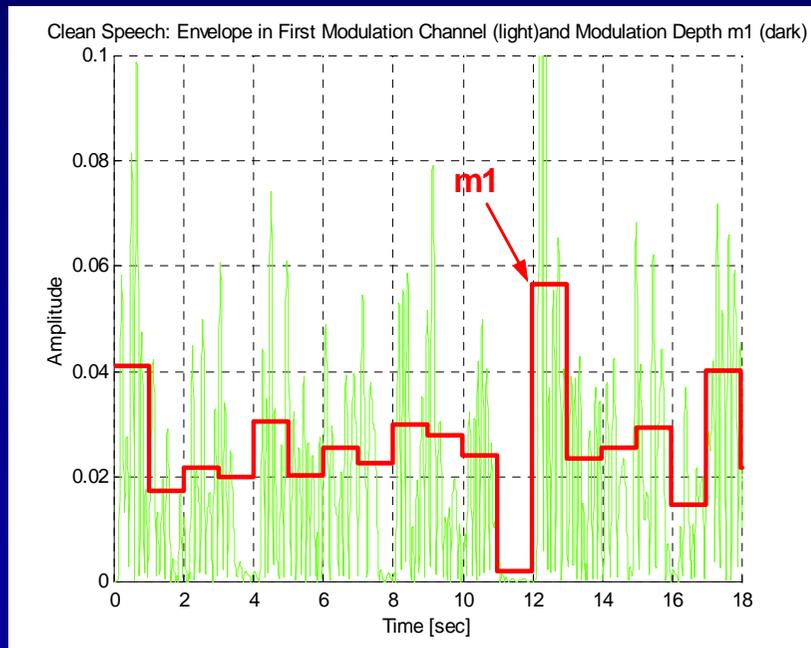
# Modulationsfrequenzanalyse

- **Merkmale:** - *Modulationstiefe für tiefe, mittlere und hohe Modulationsfrequenzen*
- **Klassifizierung:** - *trennt reine Sprache von Sprache+Störlärm und Geräusche ohne Sprache*

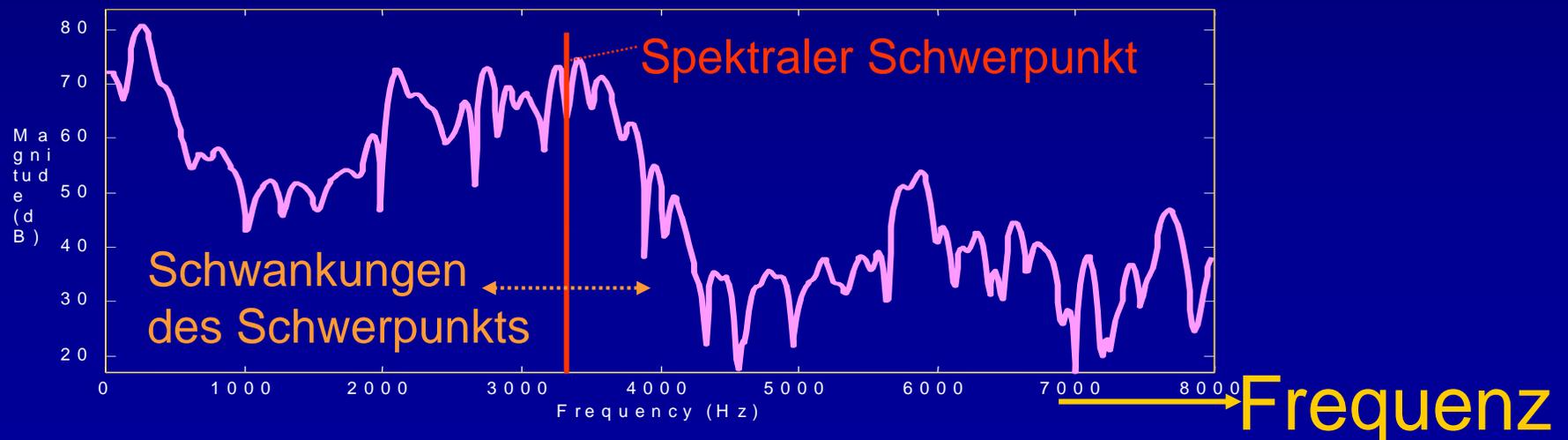
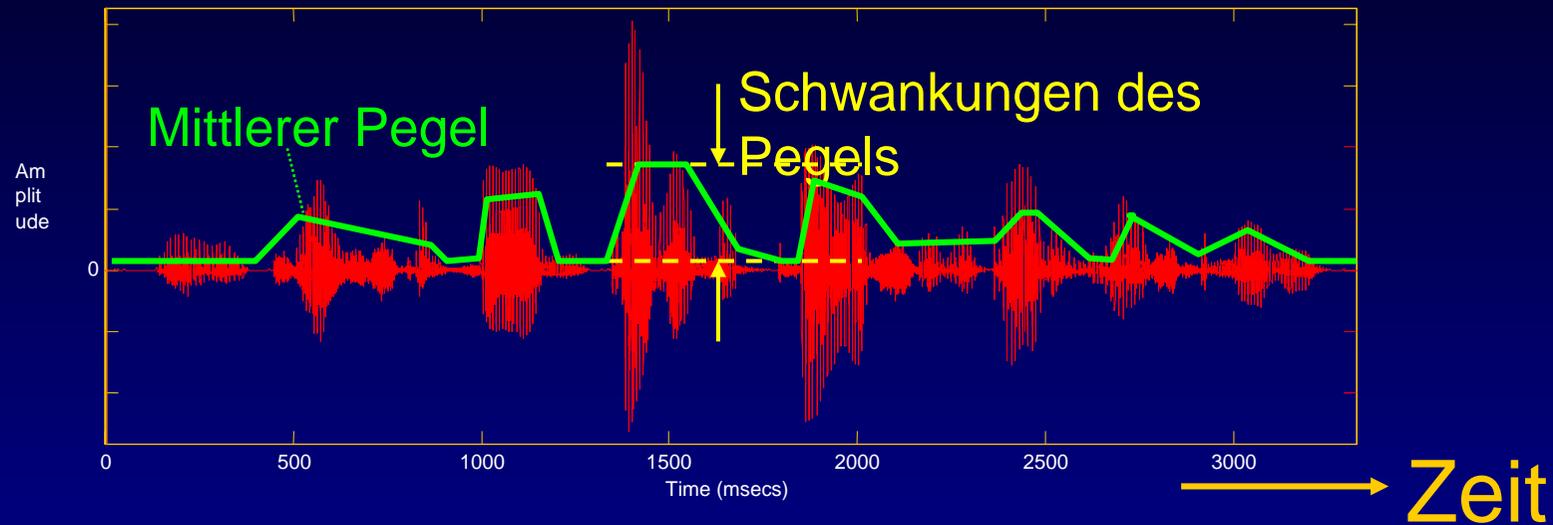


# Modulationsfrequenzanalyse

- Beschreibung der Modulationstiefe der Einhüllenden bei niedrigen, mittleren und hohen Modulationsfrequenzen. Die Stärke der Amplitudenmodulation ist ein Mass für die zeitlichen Signalschwankungen.

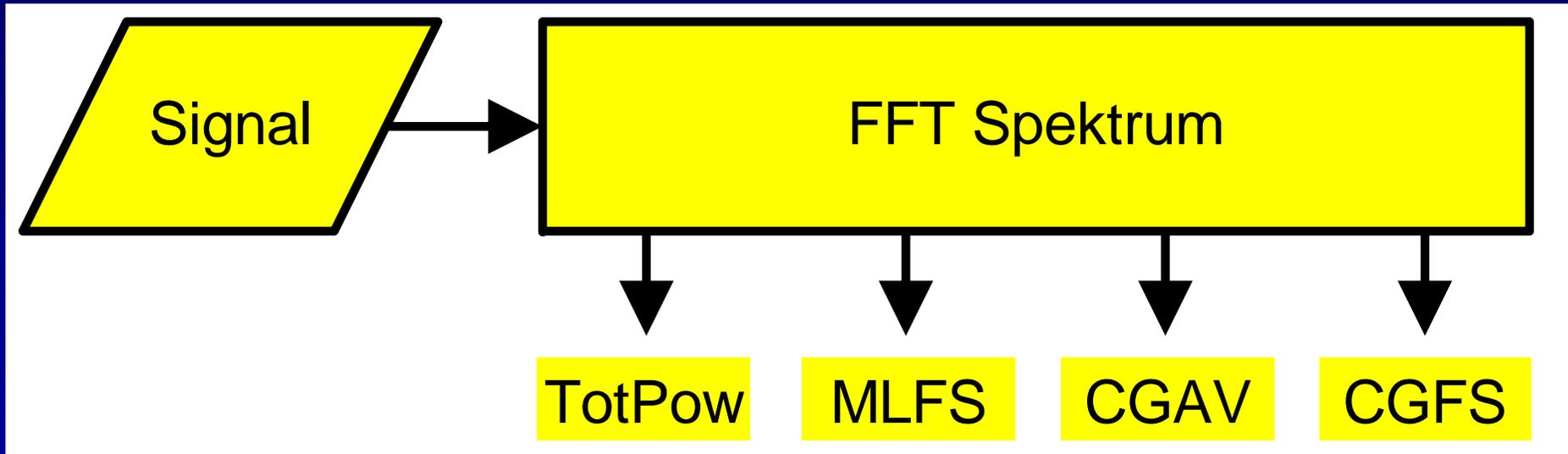


# AutoSelect: Vier Signalmerkmale



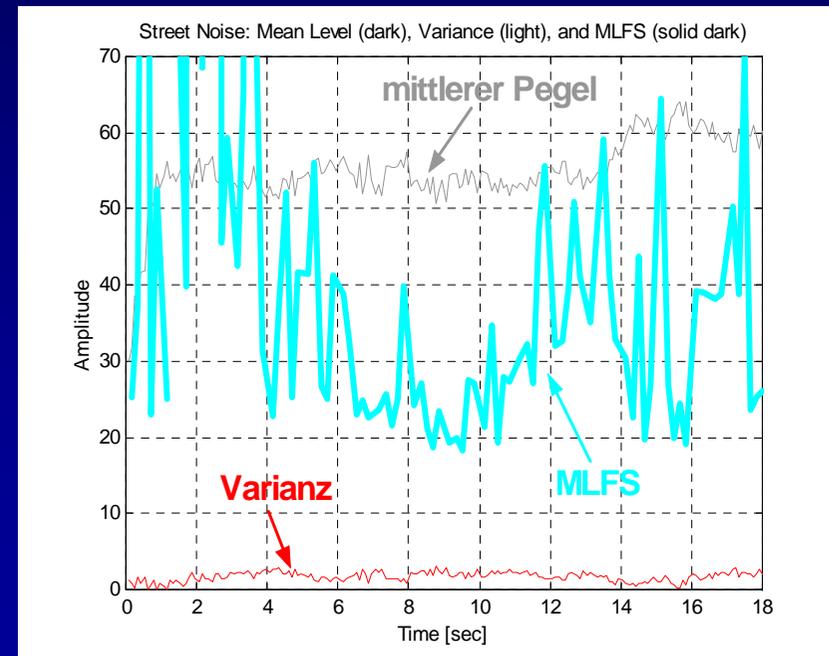
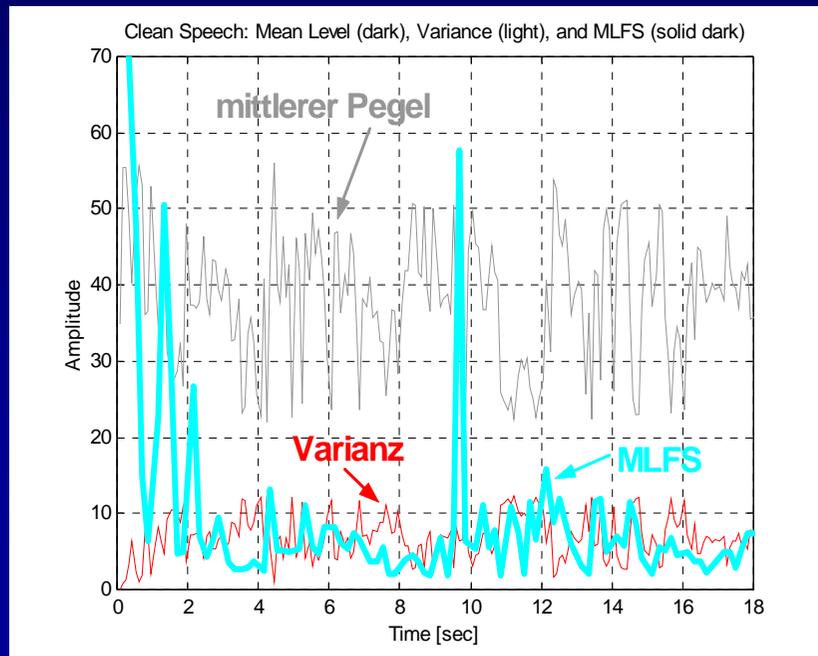
# Zeitliche & spektrale Fluktuationen und Form

- **Merkmale:** - *mittlerer Pegel, Pegelschwankungen, spektraler Schwerpunkt, Schwankungen des spektralen Schwerpunkts*
- **Klassifizierung:** - *reine Sprache eindeutig*

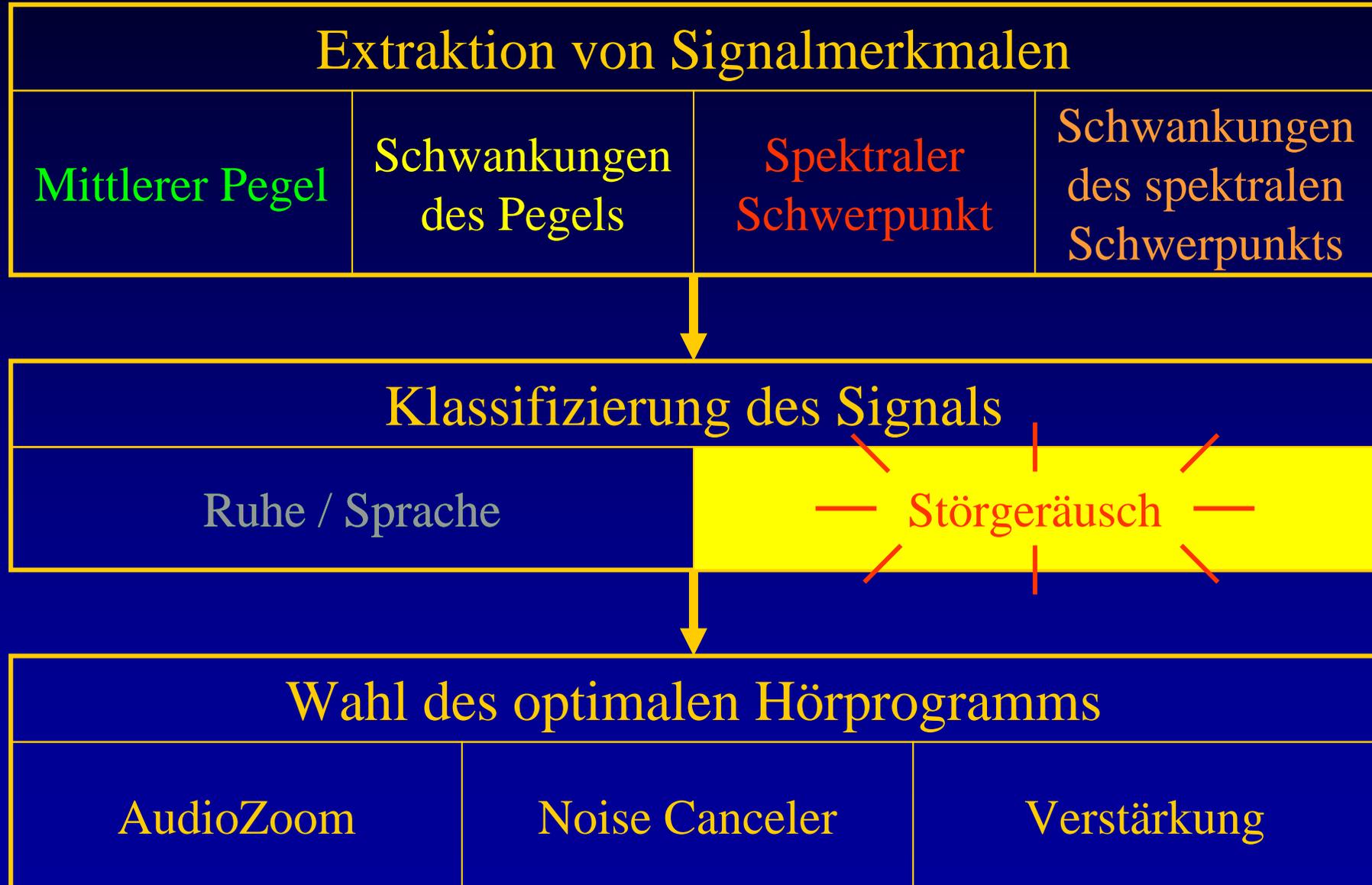


# Zeitliche & spektrale Fluktuationen und Form

- Beschreibung der Verteilung der Amplituden durch die Pegelschwankungen MLFS. MLFS ist umgekehrt proportional zur Varianz der zeitlichen Amplituden und ist damit ein Mass für die zeitlichen Schwankungen des Signals



# AutoSelect-Funktion



# Grenzen der AutoSelect-Funktion

- Robuste Erkennung von Sprachsignalen
- Musiksignale werden oft als Störgeräusch interpretiert.

## Aktuelle Forschung

- Extraktion weiterer Merkmale
- Intelligenter Mustererkenner
  - Mehr Klassen unterscheidbar:
    - Musik
    - Sprache im Störgeräusch
    - Verschiedene Störgeräusche
    - Alarmsignale

# Klassifizierung von Geräuschen

- Alle Methoden trennen auf robuste Weise Sprachsignale / Nicht-Sprachsignale
- Eine differenzierte Unterscheidung verschiedener akustischer Umgebungen ist nicht möglich
- Eine Kombination bringt keine verbesserte Klassifizierungsleistung
- Breite des Amplitudenhistogramms, Modulationstiefe und Pegelschwankungen sind alles Masse für zeitliche Schwankungen der Signalamplitude, also äquivalente Merkmale

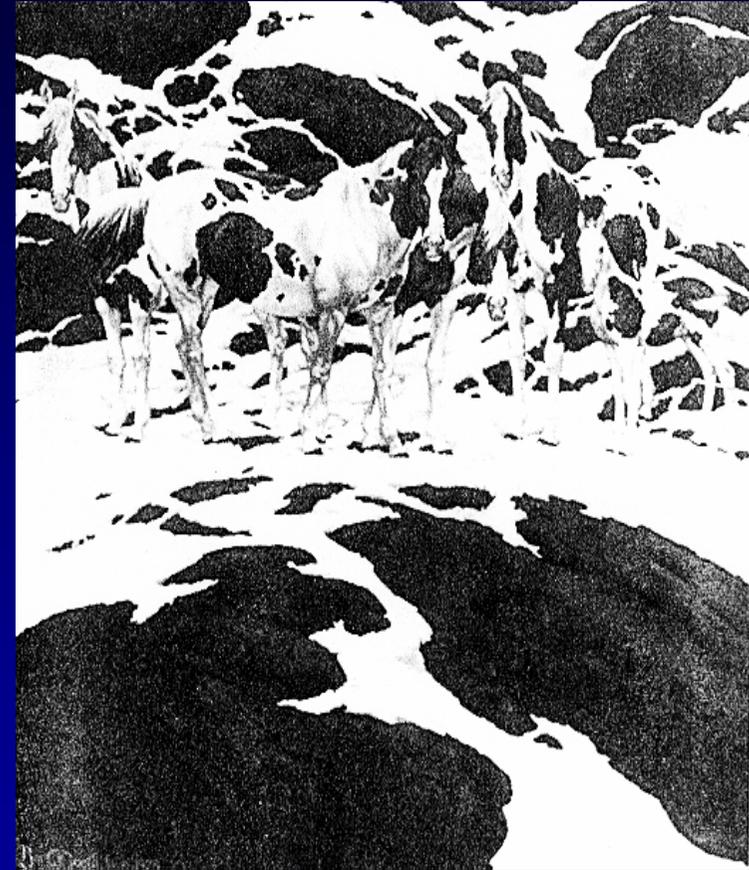
# Wie klassifiziert das auditorische System Klänge, Geräusche ?

- Die untersuchten Klassifizierer haben klare Limiten, welche unsere Ansprüche nach intelligenten Hörgeräten noch nicht erfüllen können.
- Wir können natürlich noch beliebig viele physikalische Signaleigenschaften sammeln und in den Entscheidungsprozess integrieren, doch welche Merkmale sind für die Wahrnehmung fundamental ? - Neue Wege sind zu beschreiten, lassen wir uns inspirieren ...
  - Auditorische Szenenanalyse, Wahrnehmung von akustischen Objekten

# Figur-Hintergrund Problem

*... woher weiss das auditorische System welche Komponenten zu welchem Ereignis, zu welcher Schallquelle gehören ?*

- (Auditorische) Bilder basieren auf zwei Komponenten:
  - *Fusion, auditorische Integration oder Gruppierung*
  - *Trennung*

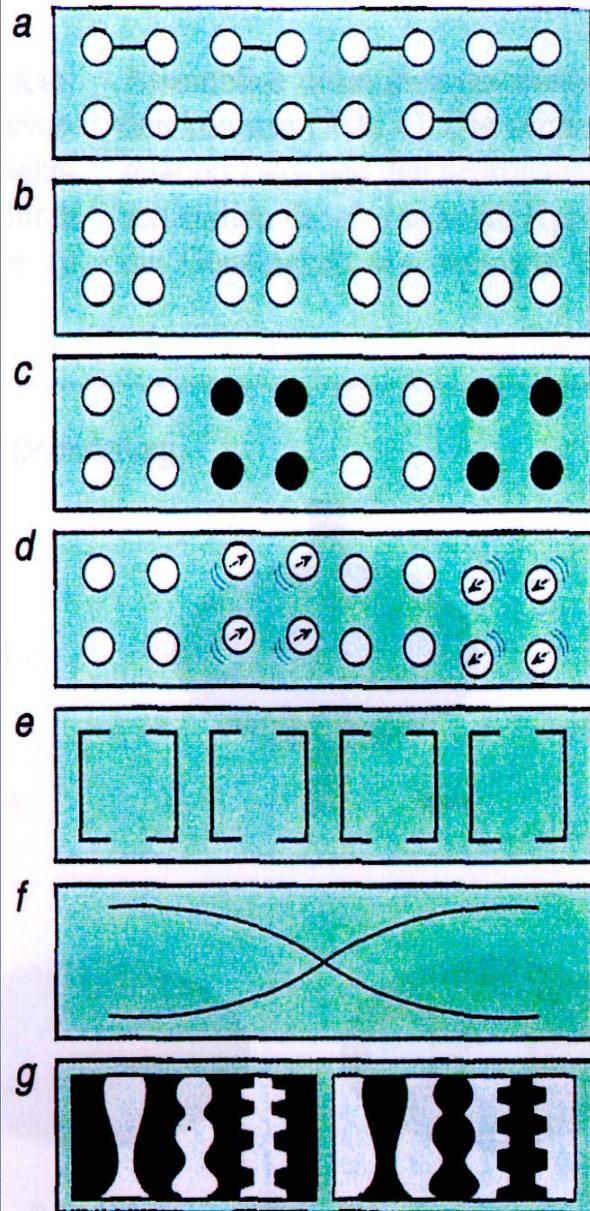


# Wahrnehmung auditorischer Objekte

- ... einzelne Instrumente werden heraus gehört, wir können also einzelne Schallquellen trennen und akustische “Objekte” wahrnehmen
- **Mechanismen ?**

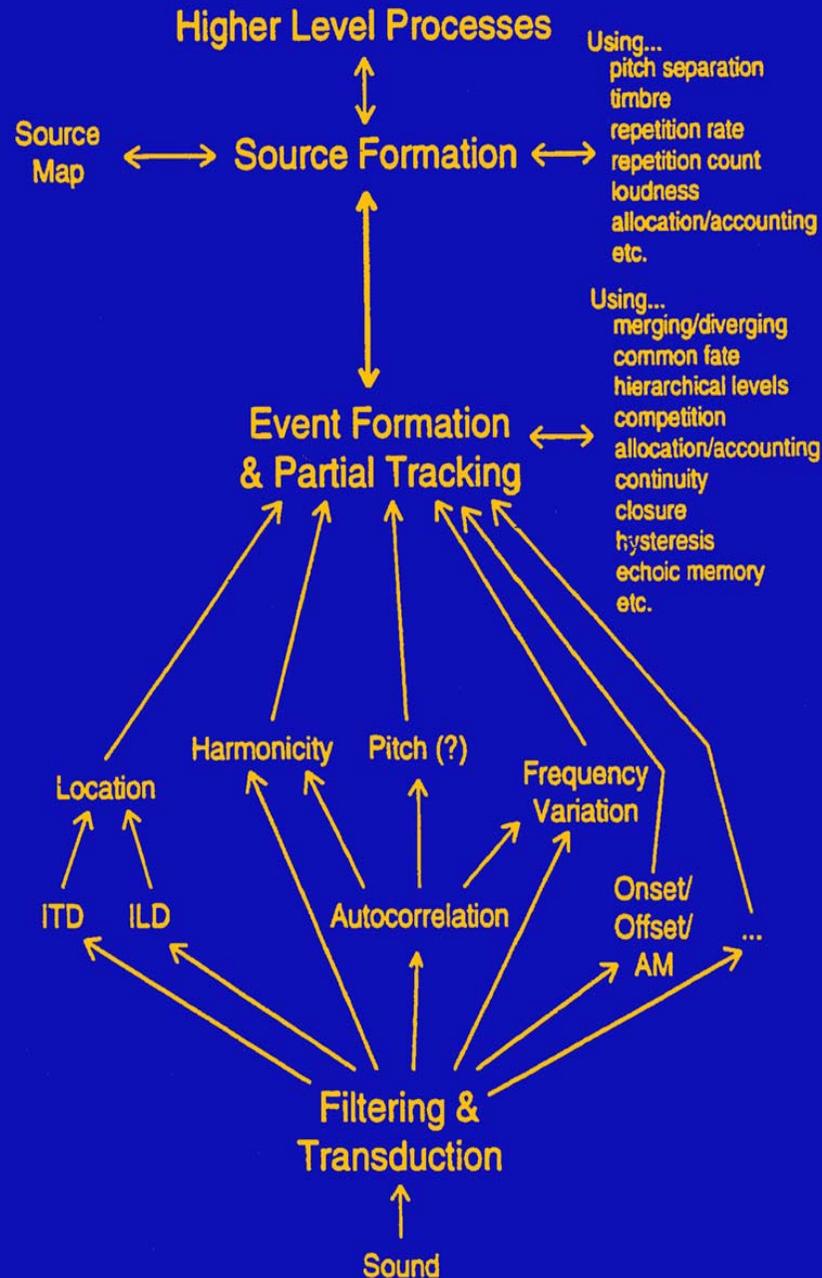


# Gestalt - Prinzipien



- (a) *Kontinuität* verbundene Regionen eines Bildes werden als Einheit gruppiert
- (b) *Nähe* in der Nähe liegende Elemente werden gruppiert
- (c) *Ähnlichkeit* Gruppierung ähnlicher Elemente
- (d) *gleiches Schicksal* Elemente welche sich räumlich und zeitlich kohärent verändern
- (e) *Abgeschlossenheit* bilden Elemente eine geschlossene Kontour, werden sie gruppiert.
- (f) *glatter Übergang* glatte Übergänge statt abrupte Veränderungen, die kreuzenden Linien werden hier gesehen und nicht zwei sich berührende Linien, Körper
- (g) *Symmetrie* symmetrisch eingefasste Gebiete erscheinen als gruppiertes Bild

# Stufen der Wahrnehmung auditorischer Objekte

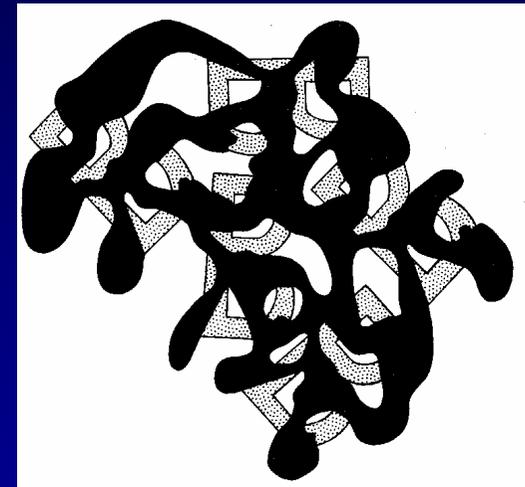


- Gruppierungsverfahren:
  - primitives Gruppieren
  - Gruppieren nach Schema
  - Gestalt-Prinzipien
- Bildung von Ereignissen
- Bildung von Quellen

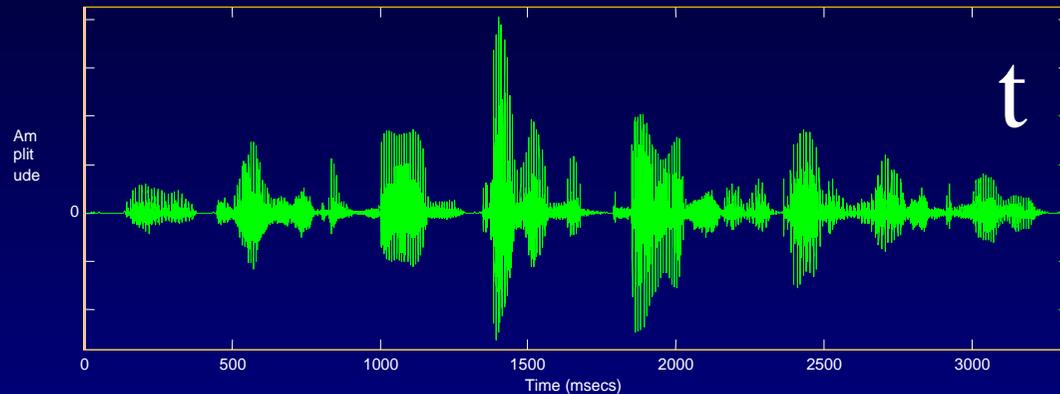
# Gruppierung: Ereignis-Bildung

Partialtöne eines Klangs zeitlich gruppieren

- Wichtigste Prinzipien:
  - Gemeinsames Schicksal  
Common Onset, AM, FM
  - Kontinuität  
Kurze Unterbrüche werden ignoriert
  - Geschlossenheit  
Perzeptive Addition von nicht vorhandenen Teilen



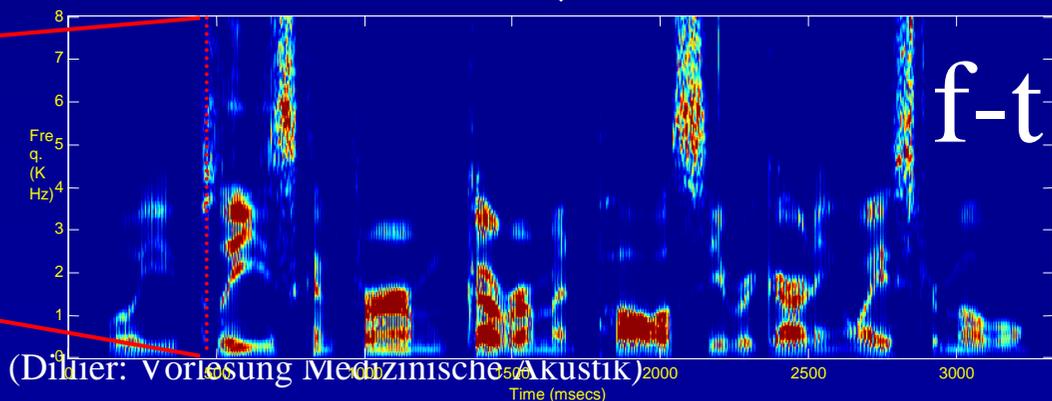
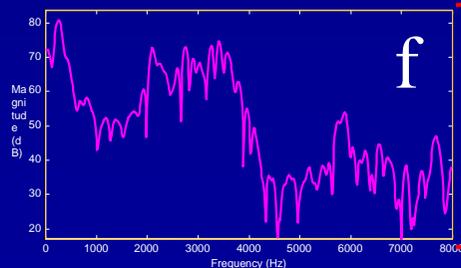
# Merkmale: Spektrale Trennung



Zeitsignal



Filterbank



Störgeräuschunterdrückung (Differenzierung)

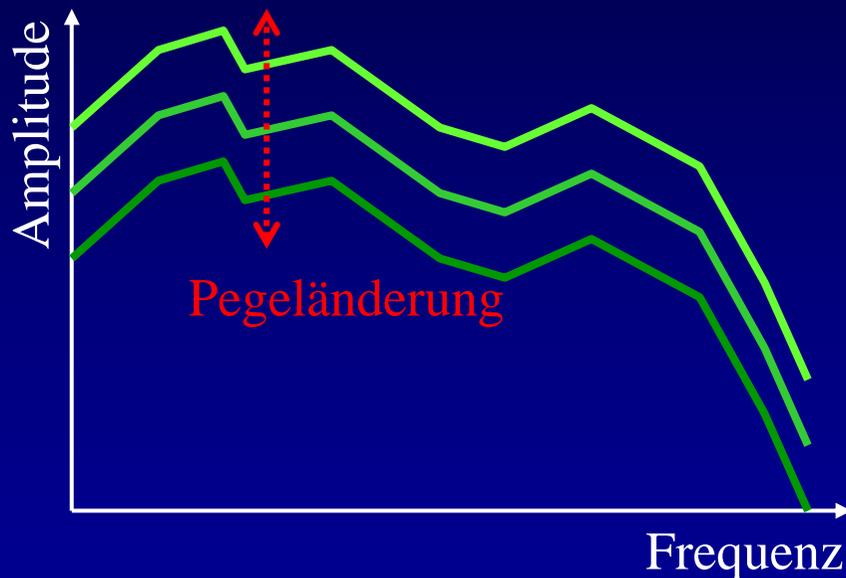
Differenzierung

2000

2500

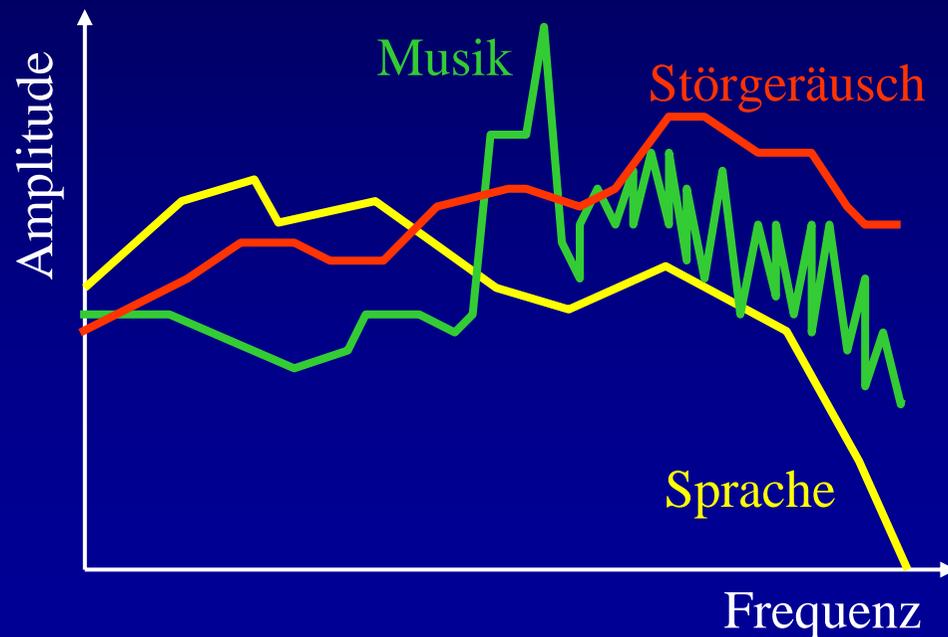
3000

# Features: Spektrales Profil



Spektrales Profil  
*pegelunabhängig* für  
viele Klänge →  
relative Amplituden-  
Differenzen werden  
betrachtet.

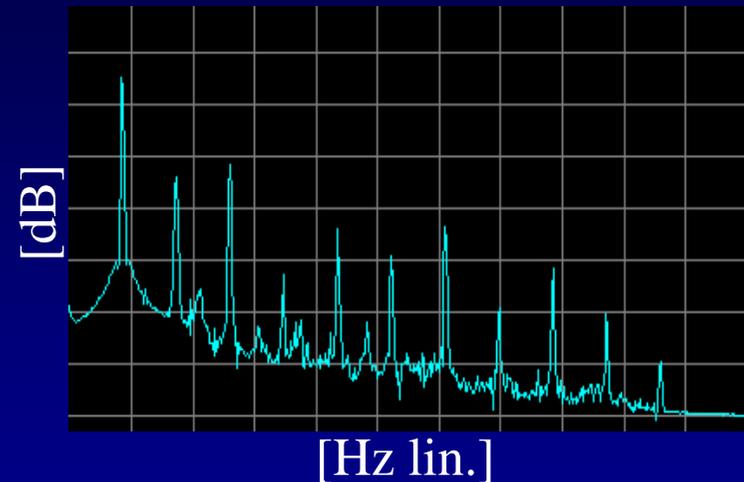
# Features: Spektrales Profil



- Unterschiedliche Klänge haben meist unterschiedliche spektrale Profile
- Grosse Varianz innerhalb Klassen möglich
- Spektrales Profil bestimmt *Klangfarbe* → *Timbre*?

# Features: Harmonizität

- Harmonizität gibt an, inwiefern einzelne Partialtöne im Signal in eine harmonische Serie fallen
- Viele natürliche Klangquellen produzieren harmonische Partialtöne



- Harmonische einzelner Quellen formen ein einzelnes auditorisches Objekt: *Pitch*
- Hervorragendes Beispiel für *simultane* Gruppierung

# Wahrnehmung auditorischer Objekte

... *Primitive Gruppierung findet auf verschiedenen hierarchischen Stufen statt und formiert Ereignisse, Quellen:*

- Die Bildung von **Ereignissen** basiert hauptsächlich auf der Gruppierung spektraler Klangbestandteile in jedem Zeitpunkt.
  - *gleiches Schicksal*      Onset / Offset, Amplitudenmodulation, Frequenzmodulation
  - *Abgeschlossenheit*      hilft Klanglücken zu überbrücken
  - *glatter Übergang*      Klangbestandteile ändern das Spektrum  
nur langsam, harmonische Abfolge

# Wahrnehmung auditorischer Objekte

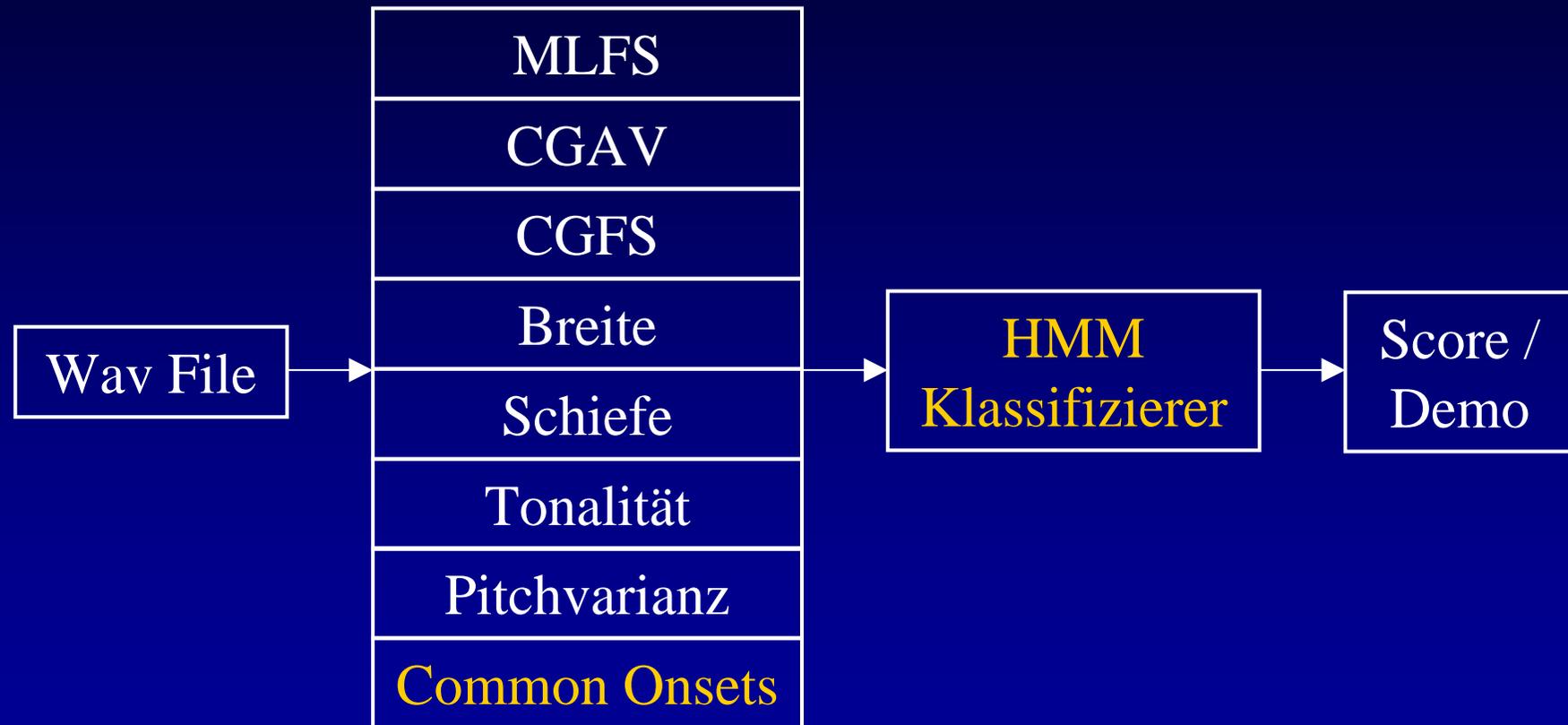
- Die Bildung von **Quellen** basiert auf der Gruppierung von zeitlich kohärenten Ereignissen.
  - *Ähnlichkeit* Töne mit gleichem Timbre kommen von einer Quelle, Instrument
  - *Nähe* aufeinanderfolgende Noten ähnlicher Frequenz (Pitch) zählen zu einer Quelle
  - *Abgeschlossenheit* stille zeitliche Lücken in Signalen sind dem auditorischen System nicht vertraut
  - *glatter Übergang* keine abrupten Übergänge im Klangcharakter

# Features für Geräuschklassifizierung

Spektrale Trennung	FFT
Amplituden-Modulation	Kates, Ludvigsen, Ostendorf
Harmonizität/Pitch	Korl: Tonalität, Pitch-Varianz
Spektrales Profil	Korl: Spektrale Welligkeit
Zeitliche Trennung, On/Offset	Büchler: Onsets
Frequenz-Modulation	?
Räumliche Trennung	?

# Geräuschklassifizierung

## Aktuelle Implementation



# Geräuschklassifizierung: Demo für 4 Klassen

- Speech
- Reverberated Speech
- Cafeteria Noise
- Speech in Cafeteria Noise
- Traffic Noise
- Classic Music
- Pan Flute
- Singing



# Schlussfolgerungen, Ausblick

- Die Rechenleistung von Hörcomputern wird in den nächsten 3 Jahren 2...3mal gesteigert, ohne dass die Chipgrösse oder der Stromverbrauch zunimmt.
- Heutige Systeme zur Klassifizierung von akustischen Ereignissen können zwar Sprache von anderen Signalen trennen, sind aber noch verbesserungsbedürftig.

# Schlussfolgerungen, Ausblick

- Das denkende, intelligente Hörgerät ist noch ein Traum. Möglicherweise helfen uns Modelle weiter, welche auf die Erkennung von auditorischen Objekten abstützen.
- Hier gibt es allerdings noch viele Unbekannte
- Stoff für weitere Forschung...

# Unter Mitarbeit von:

- Silvia Allegro
- Herbert Bächler
- Michael Boretzki
- Michael Büchler
- Mattia Ferrazzini
- Herbert Jakits
- Sascha Korl
- Volker Kühnel
- WaiKong Lai
- Stefan Launer
- Olegs Timms
- Simone Volpert