

# Die Bedeutung der räumlichen Separation von Schallquellen in der akustischen Suche

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der  
Mathematisch-Naturwissenschaftlichen Fakultät der  
Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Gunnar Regenbrecht

aus Bedburg-Hau

Oktober 2012

Aus dem Institut für Experimentelle Psychologie  
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der  
Mathematisch-Naturwissenschaftlichen Fakultät der  
Heinrich-Heine-Universität Düsseldorf

Referentin: PD Dr. Susanne Mayr

Korreferent: Prof. Dr. Axel Buchner

Tag der mündlichen Prüfung: 7. Dezember 2012

# Inhaltsverzeichnis

1. Zusammenfassung .....	5
2. Abstract .....	7
3. Einleitung.....	8
3.1. Hinweise zur Segregation.....	9
3.2. Anwendungshintergrund.....	11
3.3. Beschreibung des Separationseffektes .....	12
4. Erklärungsmechanismen des Separationseffektes .....	14
4.1. Binaurale Hinweise der Lokalisation.....	14
4.2. Monaurale Hinweise der Lokalisation.....	16
4.3. Informatorische Maskierung und Aufmerksamkeitsausrichtung .....	19
4.4. Richtungshinweise und ihr Einfluss auf den Separationseffekt .....	21
4.5. Fazit.....	26
5. Akustische Suche in alltagsnäheren Situationen .....	28
5.1. Anzahl der akustischen Reize .....	29
5.2. Sprache versus Umweltgeräusche .....	32
5.3. Vorhersagbarkeit der Position einer Schallquelle.....	34
5.4. Die Untersuchungen von Eramudugolla et al. (2008).....	36
5.5. Offene Fragen und Motivation der eigenen Untersuchungen.....	37
6. Eigene Untersuchungen .....	41
6.1. Untersuchung 1 - Separationseffekt im Raum mit Reflexionen.....	41
6.1.1. Methode.....	42
6.1.2. Ergebnisse .....	47
6.1.3. Diskussion .....	51

6.2. Untersuchung 2 - Separationseffekt im Raum ohne Reflexionen.....	56
6.2.1. Methode .....	56
6.2.2. Ergebnisse .....	57
6.2.3. Diskussion .....	60
6.3. Untersuchung 3 - Separationseffekt mit akustischem Hinweis.....	62
6.3.1. Methode .....	63
6.3.2. Ergebnisse .....	64
6.3.3. Diskussion .....	68
6.4. Untersuchung 4 - Separationseffekt und Ähnlichkeit von Reizen.....	70
6.4.1. Methode .....	73
6.4.2. Ergebnisse .....	75
6.4.3. Diskussion .....	77
7. Allgemeine Diskussion .....	79
8. Literaturverzeichnis .....	86

# 1. Zusammenfassung

Schallquellen von verschiedenen räumlichen Positionen können besser wahrgenommen werden, als wenn sie von einer gemeinsamen räumlichen Position gespielt werden. Dieser Effekt der räumlichen Separation wurde vor allem für Sprachreize mit vorhersehbarem Ort und wenigen Distraktoren beschrieben. In den Untersuchungen dieser Arbeit wurden bis zu acht naturalistische Geräusche über bis zu acht im Kreis um die Teilnehmer angeordnete Lautsprecher dargeboten. Dabei wurden in der Bedingung mit Separation verschiedene Lautsprecher und in der Bedingung ohne Separation genau ein Lautsprecher zufällig für die Wiedergabe der Geräusche ausgewählt. In den Untersuchungen 1 bis 3 konnte die Präsentationsdauer 250, 500, 1000 oder 1500 ms betragen. Die Aufgabe bestand in jedem Durchgang darin, die An- oder Abwesenheit eines zufällig ausgewählten Zielgeräusches in einem Gemisch aus vier bis acht Geräuschen rückzumelden. In einem Raum mit Reflexionen in Untersuchung 1 konnte ausschließlich für sechs gleichzeitig präsentierte Geräusche und lediglich in der Sensitivität ein Vorteil räumlich separierter Schallquellen gefunden werden – bei vergleichbar langen Reaktionszeiten. Untersuchung 2 wurde – wie alle weiteren Untersuchungen – in einem reflexionsarmen Raum durchgeführt und zeigte einen generellen Vorteil der Separation sowohl in der Sensitivität als auch in den Reaktionszeiten. In Untersuchung 3 wurde der Zielhinweis akustisch statt visuell dargeboten. Ein Vorteil räumlicher Separation zeigte sich ausschließlich in den Reaktionszeiten und auch nur bei sechs gleichzeitig präsentierten Geräuschen oder bei abwesendem Ziel, während die dazugehörigen Sensitivitäten sich nicht unterschieden. In keiner der drei Untersuchungen konnte ein Einfluss der Präsentationsdauer auf den Effekt der räumlichen Separation gefunden werden. In Untersuchung 4 mit visuellem Zielhinweis wurde schließlich die Ähnlichkeit von akustischen Reizen mit Hilfe von Musikinstrumenten variiert, die in gleicher oder verschiedener Tonhöhe dargeboten wurden. Obwohl keine Interaktion zwischen Separation und Tonhöhe gefunden werden konnte, lieferten Post-hoc-Tests nur für gleiche und nicht für verschiedene Tonhöhen einen Vorteil räumlich separierter Schallquellen. Die Ergebnisse zeigen, dass der Effekt räumlicher Separation auch in komplexen Hörsituationen mit bis zu acht naturalistischen Geräuschen, wechselndem Ziel und un-

vorhersagbarer Zielposition gefunden werden kann. Räumliche Hinweise bei akustischen Suchaufgaben scheinen aber vor allem in Situationen genutzt werden, in denen spektrotemporale Hinweise alleine nennenswert weniger hilfreich sind als beide Arten von Hinweisen zusammen.

## 2. Abstract

Perceiving sound sources from different spatial positions is easier than perceiving sound sources played from one common spatial position. This effect of spatial separation was specified primarily for speech stimuli presented from a predictable location and involving only few distractors. In the studies of this thesis, up to eight naturalistic sounds were presented through speakers arranged in a circle around the participants. Different randomly selected speakers were used in the spatially separated condition, whereas the sounds were played by one randomly selected common speaker in the condition without separation. The durations of the sounds in studies 1 to 3 could be 250, 500, 1000, or 1500 ms, and the task in each trial was to indicate the presence or absence of a randomly selected target sound within the mix of four, six or eight different sounds. Study 1 was conducted in a room with reflections and provided an advantage of separation solely in the sensitivity and with only six simultaneously presented sounds (accompanied by comparable reaction times). For study 2 and all further studies a semi-anechoic chamber was used. Under these circumstances, a general advantage of spatial separation in terms of higher sensitivities and faster reaction times could be observed. In study 3 the target cue was presented acoustically instead of visually. An advantage of spatial separation could be found exclusively in the reaction times and only when six sounds were presented or when the target was absent, while the sensitivity remained the same. In these three studies with varying sound durations, no influence of this variable on the separation effect became apparent. In study 4, a visual target cue was employed, and the similarity between the stimuli was varied by using same or different pitch stimuli. Although there was no interaction between separation and pitch, post hoc tests showed a significant advantage of spatially separated sound sources only for same but not for different pitch stimuli. Taken as a whole, the results demonstrate that the effect of spatial separation could also be found in complex listening settings featuring up to eight simultaneous naturalistic sounds, a changing target, and an unpredictable target location. In auditory search, however, spatial cues seem to be used mainly in situations in which spectro-temporal cues were noticeably less helpful than both types of cues together.

### 3. Einleitung

In komplexen akustischen Situationen erreichen Geräusche das Ohr nicht einzeln, sondern als Gemisch mit anderen Geräuschen und müssen mit Hilfe verschiedener Hinweise separiert werden, um als einzelne Geräusche wahrgenommen werden zu können. Laut Bregman (1993) sind bei der Analyse auditorischer Szenen Hinweise wie Harmonizität, Lokation und zeitliche Eigenschaften entscheidend.<sup>1</sup> Harmonische Frequenzen, d.h. solche mit ganzzahligem Frequenzverhältnis, werden einander zugeordnet wahrgenommen – genauso wie Frequenzen, die von gemeinsamen Orten stammen (oder zumindest als solche wahrgenommen werden), oder welche, die gemeinsame Lautstärkemodulationen besitzen (z.B. zeitgleich beginnen bzw. enden). Da ein einzelner dieser Hinweise nicht immer zu richtigen Schlussfolgerungen führen muss, wird angenommen, dass mehrere Hinweise zur Gruppierung von Frequenzen und damit zur Bildung von akustischen Objekten beitragen.

In dieser Arbeit wird untersucht, ob und unter welchen Umständen das Erscheinen von Geräuschen an *verschiedenen* Orten einen Vorteil für das Erkennen dieser Geräusche hat – im Vergleich zu Geräuschen, die einem *gemeinsamen* Ort entspringen. Dabei werden gegebenenfalls Einflüsse der Raumakustik, visueller oder akustischer Hinweise und der Ähnlichkeit zwischen Ziel- und Distraktorgeräuschen auf diesen Vorteil räumlich separierter Schallquellen aufgeklärt. Im nächsten Abschnitt wird zunächst auf nicht-räumliche Hinweise eingegangen, mit Hilfe derer einzelne Geräusche aus einem Gemisch extrahiert werden können. Anschließend wendet sich die Arbeit insbesondere den räumlichen Hinweisen bei der Analyse auditorischer Szenen zu. Hierzu wird der Anwendungshintergrund räumlich separierter Darbietung von akustischen Ereignissen beschrieben, woraufhin im vierten Kapitel auf Erklärungsmechanismen für den Erkennungsvorteil räumlich separierter Schallquellen und auf die damit zusammenhängenden Lokalisationshinweise eingegangen wird. Das fünfte Kapitel erläutert – auch anhand empirischer Befunde – bisher wenig untersuchte Aspekte komplexer, alltagsnaher akustischer Szenarios, während

---

<sup>1</sup> Existiert Vorwissen über die gehörten Geräusche, so können auch gelernte Schemata automatisch aktiviert oder willkürlich verwendet werden.

im sechsten Kapitel Untersuchungen vorgestellt werden, die auf Bedingungen eingehen, unter denen räumliche Hinweise für die Suche von Ereignissen in solchen komplexen, alltagsnahen akustischen Szenarios genutzt werden.

### 3.1. Hinweise zur Segregation

Zur Extraktion relevanter akustischer Information aus einem Gemisch von mehreren Geräuschen können neben räumlichen Hinweisen auch eine Reihe nicht-räumlicher Hinweise genutzt werden. Da diese Hinweise häufig auch in Untersuchungen, die sich mit der Nutzung räumlicher Hinweise beschäftigen, eine Rolle spielen, werden in diesem Abschnitt Schwankungen in der Lautstärke, Formanten und Harmonizität als solche nicht-räumlichen Hinweise erläutert, bevor sich der nächste Abschnitt wieder der räumlichen Separation von Schallquellen zuwendet.

Zielreize können besser wahrgenommen werden, wenn der Distraktor eine schwankende statt einer gleichbleibenden Lautstärke (bei gleicher Durchschnittslautstärke) besitzt (Festen & Plomp, 1990), wie es beispielsweise bei Sprache der Fall ist. Dieses Hören in den Lücken (englisch *dip listening* oder *listening in the gaps*) genannte Phänomen erleichtert die Wahrnehmung eines Zielreizes, da sich das Zielreiz-Distraktor-Lautstärkeverhältnis oder auch Signal-Rausch-Verhältnis (englisch *signal to noise ratio*) in kurzen Zeitabschnitten verbessert. Miller und Licklider (1950) untersuchten das Hören in den Lücken, indem sie Sprache zusammen mit weißem Rauschen präsentierten, das verschieden oft in der Sekunde ein- und ausgeschaltet wurde. Die Sprachverständlichkeit war am höchsten, wenn die Unterbrechungsrate bei etwa 10 Hz lag. Dabei spielte es keine große Rolle, ob die Unterbrechungen in regelmäßigen oder zufälligen Abständen erfolgten. Festen und Plomp (1990) untersuchten zusätzlich, ob eine komplexere Amplitudenmodulation, etwa unabhängig für verschiedene Frequenzbänder, zu einer geringeren Maskierung führt als eine einfache Amplitudenmodulation des gesamten Signals. Dazu erstellten sie ein konstantes Rauschen, ein einfach amplitudenmoduliertes Rauschen und ein Rauschen, dessen Amplitudenmodulation für hohe und tiefe Frequenzen unabhängig war – ähnlich der Lautstärkehüllkurve von Sprache für diese Frequenzen. Zusammen mit dem konstanten Rauschen erhielten sie für Sprache eine Verständlichkeitsschwelle von  $-4.7$  dB. Im Vergleich dazu

betrug die Schwelle bei einfach amplitudenmoduliertem Rauschen  $-8.4$  dB, bei Rauschen mit zweibandiger Modulation  $-9.5$  dB und bei Sprache als Maskierer  $-11.4$  dB. Die Wahrnehmung von Sprache scheint also um so leichter zu fallen, je komplexer sich die Amplitudenmodulation des Maskierers gestaltet.

Formanten sind Frequenzbereiche mit hoher Schallenergie und entstehen zum Beispiel durch Resonanzen des Vokaltrakts des Menschen (Mundraum, Rachen und Nasenhöhlen) oder des Korpus eines Musikinstruments. Laut Darwin, Brungart und Simpson (2003) können zwei Sprachsignale einer Person leichter getrennt werden, wenn die Länge des Vokaltraktes bei einem Signal in einer Simulation um mindestens den Faktor 1,08 verlängert wird. Die Effektivität einer Trennung von Sprache bei verschiedenen Geschlechtern wird jedoch nicht erreicht. Eine Kombination aus Formant und Grundfrequenz als Hinweise zur Trennung von Sprachsignalen ist zudem vorteilhafter als die Summe der Einzelleffekte erwarten lässt (Darwin et al., 2003).

Geräusche mit harmonischen (d.h. ganzzahligen) Obertonstrukturen – wie zum Beispiel die stimmhaften Anteile von Sprache – können besser voneinander getrennt werden, wenn sie verschiedene Grundfrequenzen (und damit i.d.R. auch weitgehend verschiedene Obertonfrequenzen) besitzen. Darwin et al. (2003) präsentierten zwei Sätze eines Sprechers<sup>2</sup> gleichzeitig, wobei sie die Differenz der Grundfrequenzen auf bis zu einer Oktave erhöhten. Es zeigte sich, dass ein Unterschied von zwei Halbtönen oder mehr zu einer besseren Trennung der beiden Sprachsignale führte. Eine ähnlich gute Trennung kann beobachtet werden, wenn Sprecher auch ohne Manipulation eine hohe Variabilität in der Intonation aufweisen. Unter Umständen kann dann mit Hilfe von Änderungen der Grundfrequenz (oder der Vokaltraktlänge) keine weitere Verbesserung erzielt werden. Die Trennung von Sprachsignalen allein anhand von Grundfrequenzen gelingt jedoch nicht so gut wie eine Trennung eines Gemisches aus männlicher und weiblicher Sprache.

---

<sup>2</sup> In dieser Arbeit wird das generische Maskulinum verwendet, so dass auch weibliche Personen gemeint sind.

### 3.2. Anwendungshintergrund

Bei räumlich separierten Schallquellen im Vergleich zu akustischer Information, die von einem Ort stammt, können in verschiedenen Arbeitssituationen wie Telefonkonferenzen, Notrufzentralen, Cockpits und Leitstellen in der Luft- und Raumfahrt bessere Leistungen beobachtet werden. So konnte Begault (1999) zeigen, dass die Sprachverständlichkeitsschwelle für Sprache, die in einer virtuellen akustischen Anzeige räumlich separiert von sprachähnlichem Rauschen präsentiert wurde, niedriger ausfiel als bei nicht separierter Sprache. Auch Abouchacra, Breitenbach, Mermagen und Letowski (2001) lieferten Evidenz für einen Vorteil räumlicher Separation von Sprache in einer simulierten militärischen Umgebung mit Hintergrundlärm. Die Teilnehmer waren Soldaten und trugen während der Untersuchung einen modifizierten binauralen Helm der US-Armee, mit dessen Hilfe eine räumliche Darstellung der Sprachreize basierend auf kopfbezogenen Übertragungsfunktionen (*head related transfer functions*)<sup>3</sup> erfolgen konnte. Die Autoren sehen Anwendungsfelder einer Technik mit räumlich separierter Darstellung von Schallreizen in strategischen Betriebszentralen, in der Überwachung von komplexen Sicherheitssystemen und in der Aufsicht der Luftfahrt. Arrabito (2006) fand eine verbesserte Detektionsleistung von Zielen durch eine dreidimensionale akustische Anzeige bei Personen, die ein Sonar bedienen. Eine Anwendungsmöglichkeit bestehe in der Verarbeitung passiver Sonarsignale, die mit Hilfe von dreidimensionalen akustischen Anzeigen zu früheren Entdeckungen feindlicher Ziele führen könnten. Weitere Untersuchungen belegen den Vorteil räumlich separierter Schallquellen in dreidimensionalen akustischen Anzeigen für Detektionsaufgaben im Flugsimulator (McAnally & Martin, 2007) und bei der Simulation von Luftverkehrüberwachungen (MacDonald, Balakrishnan, Orosz, & Karplus, 2002).

---

<sup>3</sup> Mit Hilfe von kopfbezogenen Übertragungsfunktionen werden von dem Ort der Schallquelle abhängige Einflüsse von Körperteilen wie Kopf, Ohrmuschel, Schultern und Rumpf simuliert. Dadurch ist virtuell nicht nur eine laterale Positionierung von Schallquellen möglich, sondern auch eine im gesamten dreidimensionalen Raum, so dass etwa Funksprüche, die nicht per se eine räumliche Position besitzen, auf eine ähnliche Weise wie in einer natürlichen Hörsituation an verschiedenen Orten im Raum dargestellt werden können.

### 3.3. Beschreibung des Separationseffektes

Im Gegensatz zur visuellen Modalität, in der die räumlichen Positionen von Reizen bereits auf Rezeptorebene direkt neuronal encodiert werden können (retinotope Informationsverarbeitung), können räumliche Positionen von Reizen in der auditorischen Modalität erst im weiteren Verlauf von Verarbeitungsprozessen aus den Signalen der in der Cochlea befindlichen Rezeptorzellen extrahiert werden. Die Sensitivität letzterer ist nämlich frequenzspezifisch (tonotope Informationsverarbeitung; Wolfe et al., 2009). Die räumliche Information von Schallquellen muss also auf höherer Ebene aus der neuronalen Repräsentation extrahiert werden. Auf die für die Richtungswahrnehmung relevanten Hinweise wird in Kapitel 4 noch näher eingegangen.

Von einer Maskierungsreduktion durch räumliche Separation oder räumlicher Demaskierung (englisch: *spatial release from masking* oder *spatial unmasking*) spricht man, wenn die Maskierung bei Schallquellen von *verschiedenen* Orten geringer ist als bei Schallquellen von nur *einem* Ort (z.B. Ebata, 2003). Eine dadurch erzielte Verbesserung der Leistung in Lokalisations-, Detektions- oder Aufgaben des Sprachverstehens wird im Folgenden als Effekt oder Vorteil räumlicher Separation oder kurz als Separationsvorteil oder -effekt bezeichnet. Die binaurale Differenz im Maskierungspegel (*Binaural Masking Level Difference*) gibt hingegen an, um wie viel der Pegel eines Signals verringert werden kann, wenn mit zwei anstatt nur mit einem Ohr – oder mit identischem Signal an beiden Ohren – gehört und dabei die gleiche Leistung erreicht wird (z.B. Stern & Trahiotis, 1995). Sind die Signale an beiden Ohren identisch, so spricht man von einer diotischen Hörsituation. Unterscheiden sich die Signale – wie fast immer unter natürlichen Bedingungen – nennt man die Hörsituation dichotisch.

Der Effekt räumlicher Separation ist jedoch abhängig davon, wie eine räumliche Separation relativ zum Hörer ausgerichtet ist: Eine räumliche Separation, bei der sich die Reize links und rechts der hörenden Person befinden, führt zu einem größeren Effekt als eine Vorne-hinten-Separation (Dirks & Wilson, 1969). Befindet sich der Zielreiz vorne, so erhält man den größten Separationseffekt nicht bei maximalem Winkel (180°, d.h. hinten), sondern in der Nähe der maximalen lateralen Separation (ca. -90° oder 90°, d.h. links oder

rechts; Bronkhorst & Plomp, 1988). Analog dazu ist die Links-rechts-Lokalisation im Allgemeinen auch genauer als die Vorne-hinten-Lokalisation (z.B. Giguere & Abel, 1993). Diese Ergebnisse stehen im Einklang mit der Annahme, dass vor allem interaurale Pegel- und Zeitdifferenzen zu einem Separationseffekt beitragen und die für die Vorne- oder Hinten-Lokalisation entscheidenden spektralen Hinweise eine geringere Rolle spielen.

Im vierten Kapitel werden nun die Erklärungsmechanismen für den Separationseffekt erläutert, wozu zunächst auf die verschiedenen Lokalisationshinweise in der horizontalen Ebene eingegangen wird. Anschließend wird im fünften Kapitel der Separierungseffekt in alltagsnäheren, vor allem komplexeren Situationen thematisiert, wie sie auch in den Untersuchungen dieser Arbeit betrachtet wurden, bevor im sechsten Kapitel die Intention der vorgestellten Untersuchungen erläutert wird.

## 4. Erklärungsmechanismen des Separationseffektes

Der Vorteil räumlich separierter Schallquellen beim Identifizieren von Zielreizen geht auf mehrere Mechanismen zurück, die unterschiedlich gut verstanden und beschrieben worden sind (Shinn-Cunningham, 2005). Zunächst wird in den Abschnitten 4.1 und 4.2 auf die an der Wahrnehmung separierter Lokationen beteiligten Lokalisationshinweise selbst eingegangen, bevor in Abschnitt 4.3 das Phänomen der informativen Maskierung und die Rolle der Aufmerksamkeitsausrichtung als weitere, weniger gut erforschte Erklärungsansätze des Separationseffektes behandelt werden. Daraufhin wird der Zusammenhang dieser Mechanismen zum Separationseffekt in Abschnitt 4.4 aufgezeigt, bis in Abschnitt 4.5 schließlich Schlussfolgerungen gezogen werden.

### 4.1. Binaurale Hinweise der Lokalisation

Die Bestimmung der lateralen azimuthalen Richtung von akustischen Ereignissen erfolgt wesentlich mit Hilfe von zwei binauralen Hinweisen: der interauralen Pegeldifferenz (*interaural level difference*) und der interauralen Zeitdifferenz (*interaural time difference*). Interaurale Pegel- und Zeitdifferenzen treten bei Schallquellen auf, die links oder rechts der Medianebene einer Person liegen. Während diese Differenzen also bei Schallquellen von mittig vorne oder hinten (und auch von oben oder unten) nicht auftreten, erhöhen sich interaurale Pegel- und Zeitdifferenzen – etwas vereinfacht gesagt – mit zunehmend lateraler gelegenen Schallquellen (siehe auch Colburn, Shinn-Cunningham, Kidd, & Durlach, 2006). Kommt beispielsweise ein Ton von links, so erreicht der Schall das linke Ohr früher und auch mit einem höheren Pegel als das rechte Ohr. Eine erste Verarbeitung von interauralen Pegel- bzw. Zeitdifferenzen im auditorischen Nervensystem findet in der lateralen bzw. medialen oberen Olive statt (Wolfe et al., 2009).

Zur Entstehung von interauralen Pegeldifferenzen tragen mindestens zwei Mechanismen bei: Einerseits wird Schall durch den akustischen Kopfschatten in seinem Pegel abgeschwächt. Andererseits entstehen interaurale Pegeldifferenzen auch dadurch, dass die Strecken von einer lateralen Schallquelle zu den beiden Ohren verschieden lang sind und

der Schallpegel mit der Distanz abnimmt. Letztere Differenzen können jedoch nur bis zu einem Abstand von ein oder zwei Metern wahrgenommen werden, weil darüber das Verhältnis zwischen den beiden Strecken nahe bei eins liegt, was zu sehr kleinen interauralen Pegeldifferenzen führt. Im Nahbereich sind interaurale Pegeldifferenzen folglich nicht nur von der Richtung, sondern auch von der Entfernung der Schallquelle abhängig (Shinn-Cunningham, Santarelli, & Kopco, 2000).

Pegeldifferenzen zwischen den Signalen an den beiden Ohren des Menschen können im Freifeld (d.h. ohne Reflexionen) bis zu etwa 20 dB betragen – jedoch vor allem für hohe Frequenzen, da die Kopfabschattung hohe Frequenzen stärker betrifft als niedrige (Bronkhorst & Plomp, 1988). Allerdings nehmen die Pegeldifferenzen nicht einfach nur mit steigenden Frequenzen zu, sondern können (auf Grund von Reflexionen durch Körperteile wie Schultern, Brust und Rücken) nicht monotone Verläufe aufweisen, was Macpherson und Middlebrooks (2000) dazu veranlasst haben mag, den Begriff interaurales Pegelspektrum (*interaural level spectrum*) einzuführen. Sie fanden jedoch keinen Unterschied in der Beurteilung von lateralen Winkeln, wenn das detaillierte interaurale Pegelspektrum implementiert wurde anstatt einer (für alle Frequenzen konstanten) interauralen Gesamtpegeldifferenz, die im Ergebnis als effektiver Hinweis bestätigt werden konnte. Pegeldifferenzen scheinen also weitgehend auch genutzt werden zu können, wenn keine naturgetreue Frequenzabhängigkeit vorliegt.

Interaurale Zeitdifferenzen entstehen, wenn die Ohren einer Person verschieden weit von einer Schallquelle entfernt sind. Die dabei auftretenden interauralen Zeitdifferenzen können bis zu 650  $\mu$ s betragen (Culling & Akeroyd, 2010). Das menschliche Gehör ist dabei vor allem sensitiv für Zeitdifferenzen im tieffrequenten Signalen (Colburn et al., 2006). Bei Sinustönen mit Frequenzen oberhalb von 1500 Hz werden keine interauralen Zeitdifferenzen wahrgenommen; interaurale Zeitdifferenzen von 650  $\mu$ s würden hier auch zu einer Verschiebung von mehreren Wellenzügen führen, die ein Zuordnen der Phasen zu den Signalen auf dem linken und rechten Ohr uneindeutig macht. Bei breitbandigen hochfrequenten Reizen kann jedoch die Zeitdifferenz zwischen den an den beiden Ohren anliegenden Lautstärkehüllkurven der Reize genutzt werden. Unter natürlichen Bedingungen fallen die interauralen Zeitdifferenzen für hohe Frequenzen zwar um etwa 30 % kleiner

aus als für tiefe – möglicherweise weil tiefe Frequenzen eher als hohe um den Kopf herum gebeugt werden und der durchschnittliche Weg dadurch für tiefe Frequenzen länger ist. Diese Frequenzabhängigkeit scheint jedoch keine einflussreiche Eigenschaft der interauralen Zeitdifferenzen zu sein: Bronkhorst und Plomp (1988) fanden keine unterschiedlichen Sprachverständlichkeitsschwellen für Reize mit (natürlichen) frequenzabhängigen interauralen Zeitdifferenzen verglichen mit Reizen mit interauralen Zeitdifferenzen, die eine (künstlich festgelegte) konstante Zeitdifferenz für alle Frequenzen aufwiesen.

Zhang und Hartmann (2010) untersuchten die Bedeutung des interauralen Pegelspektrums für die Vorne-Hinten-Lokalisation. Dazu wurden die spektralen Details auf einem Ohr beseitigt, das interaurale Pegelspektrum aber erhalten (mit der Folge, dass auch auf dem anderen Ohr veränderte spektrale Hinweise entstanden). Jedoch verschlechterte sich auch in diesem Fall die Vorne-Hinten-Lokalisation, so dass das interaurale Pegelspektrum alleine nicht auszureichen scheint. Zhang und Hartmann untersuchten auch, ob die Vorne-Hinten-Lokalisation auf feste Frequenzbereiche angewiesen ist, fanden jedoch, dass auf Information in anderen Frequenzbändern ausgewichen werden kann, wenn die Information aus dem vermeintlich notwendigen Frequenzband beseitigt wurde. Insgesamt folgern die Autoren aus ihren Untersuchungen, dass es bei der Vorne-Hinten-Lokalisation große individuelle Unterschiede und verschiedene verwendete Strategien gibt, wobei sogar auch innerhalb einer Person mehrere Strategien eingesetzt werden können.

## 4.2. Monaurale Hinweise der Lokalisation

Akustische Hinweise zur horizontalen Richtungsbestimmung, die nicht auf Unterschieden in den Signalen beider Ohren basieren, nennt man monaural. Monaurale Hinweise können somit auch beim Hören mit nur einem Ohr genutzt werden (siehe auch Blauert, 1997).

Ein monauraler Hinweis, die Lautstärke eines akustischen Ereignisses, kommt durch akustische Abschattung durch den Kopf der hörenden Person zustande (z.B. Culling & Akeroyd, 2010). Diese Richtungsabhängigkeit der Lautstärke beruht auf dem gleichen physikalischen Mechanismus wie die interauralen Pegeldifferenzen. Befindet sich der Zielreiz weiter links (bzw. rechts) als der Distraktor, so ergibt sich für das linke (bzw. rechte) Ohr

ein günstigeres Zielreiz-Distraktor-Lautstärkeverhältnis. Wie auch bei den interauralen Pegeldifferenzen fällt die Abschattung für hohe Frequenzen stärker aus als für niedrige Frequenzen und kann zu einem Unterschied von bis zu 20 dB zwischen den Signalen an den Ohren führen (Feddersen, Sandel, Teas, & Jeffress, 1957).

Bei Variation der horizontalen Position einer Schallquelle kommt es im Signal am Gehörgang neben Zeit- und Pegelveränderungen auch zu Veränderungen des Spektrums, die auch monaural ausgewertet werden können. Den Einfluss dieser Veränderungen im Spektrum auf die Wahrnehmung der Schallrichtung untersuchten Macpherson und Middlebrooks (2000): Sie präsentierten Geräusche über Kopfhörer und ließen ihre Teilnehmer Kopfdrehbewegungen hin zu den virtuellen Positionen der Geräusche durchführen. Während die Zeit- und Pegeldifferenzen zwischen den Ohren zu einem Ort A passten, wurde das Spektrum auf dem (dem Zielreiz) zugewandten Ohr wie für einen Ort B präsentiert, der um 30 oder 60° von Ort A abwich. Auf dem abgewandten Ohr wurde ein Spektrum präsentiert, das zu einer interauralen Spektrumsdifferenz führte, die bei Ort A entstanden wäre. (Dieses Spektrum sollte also nicht wie das der zugewandten Seite zu Ort B passen, da sonst auch eine interaurale Spektrumsdifferenz wie bei Ort B entstanden wäre und somit eine Konfundierung von Spektrum und interauraler Spektrumsdifferenz vorgelegen hätte.) Monaurale spektrale Hinweise erwiesen sich jedoch als wenig bedeutend für die Beurteilung des lateralen Winkels, da die Personen den Zielreiz weitgehend an der Position lokalisierten, die durch Zeit- und Pegeldifferenzen charakterisiert war.

Obwohl zumindest ein Teil der Menschen, die seit ihrer Geburt mit nur einem Ohr hören, oft recht genau (mit Fehlern von durchaus unter 10°) lateral lokalisieren kann, scheinen Normalhörende monaurale Hinweise weniger effektiv nutzen zu können und zeigen mit einem verschlossenen Ohr typischerweise zum offenen Ohr hin gerichtete Lokalisationsfehler zwischen 15 und 40° (Slattery & Middlebrooks, 1994), während die Fehler (in beide Richtungen) bei binauralem Hören zwischen -1° und 3° liegen. Die Autoren konnten auch keine wesentliche Verbesserung beobachten, wenn das Ohr vor der Messung bereits 24 Stunden verschlossen war.

Untersuchungen von Wightman und Kistler (1997) legen nahe, dass eine korrekte Vornehinten-Lokalisation allein auf Basis von monauralen Hinweisen kaum möglich ist. Frühere Untersuchungen zu monauralen Hinweisen, in denen ein Ohr der Teilnehmer verschlossen wurde, leiden möglicherweise unter einer Konfundierung mit interauralen Hinweisen, da die Abschirmung sowohl schwierig zu realisieren als auch nur mit Unsicherheit abzuschätzen ist und für niedrige Frequenzen im Allgemeinen auch weniger effektiv ausfällt. Um den Einfluss monauraler und auch binauraler Hinweise auf die Vorne-Hinten-Lokalisation unter Berücksichtigung eventuell unzureichender Abschirmung zu untersuchen, wurden binaurale und monaurale Richtungsbestimmungen sowohl im Freifeld (ohne und mit Verschluss eines Ohres) als auch im virtuellen Raum (beidseitige und einseitige Darbietung mit Hilfe von Kopfhörern) durchgeführt. Der Vorteil virtueller Präsentation lag darin, dass auf der nicht stimulierten Seite auch in den tiefen Frequenzen ein um mindestens 40 dB geringerer Pegel gemessen werden konnte. Während die Winkel einer Schallquelle in der binauralen Bedingung für Freifeld- und virtuelle Präsentation ähnlich genau beurteilt wurden, schien die Genauigkeit in der monauralen Bedingung bei virtueller Präsentation geringer als bei Freifeldpräsentation und auch nicht sichtbar oberhalb des Zufallsniveaus auszufallen. Zwar berichten Wightman und Kistler (1997) keine Inferenzstatistik, jedoch stehen die Daten zumindest deskriptiv im Einklang mit der Interpretation, dass erfolgreiche Vorne-Hinten-Lokalisation im Freifeld auf binaurale Hinweise (und damit auf unvollkommenes Verschließen des einen Ohres in manchen Studien) zurückgeht. Dies wird auch in einer weiteren Untersuchung mit variierender Abdämpfung des Signals auf einem Ohr um 10, 20, 30 und 40 dB bestätigt: Mit zunehmender Dämpfung stimmten geschätzte und tatsächliche Winkel dem Erscheinungsbild nach immer weniger gut überein.

Auch die Untersuchungen von Zhang und Hartmann (2010) lassen eine zuverlässige Vorne-Hinten-Lokalisation alleine anhand von monauralen Hinweisen abwegig erscheinen. Die Autoren verwendeten eine Technik, die den virtuellen Raum über nahe am Ohr positionierte Lautsprecher simuliert, da Kunstkopfmessungen mit Kopfhörern – abhängig von der genauen Position der Kopfhörerpolster auf dem Kopf – Standardabweichungen von bis zu 10 dB in den Pegelmessungen aufweisen können (vor allem bei Frequenzen von etwas weniger als 10 und 15 kHz; Kulkarni & Colburn, 2000). Um den Einfluss monauraler

Hinweise zu untersuchen, beseitigten Zhang und Hartmann (2010) die spektralen Details im Signal des einen Ohrs, so dass keine interauralen Spektrumsdifferenzen mehr für die Vorne-hinten-Lokalisation vorlagen, aber gleichzeitig auch keine so starke Lateralisierung auftrat wie beim Verschließen eines Ohrs. Das Signal auf dem anderen Ohr inklusive der spektralen Hinweise wurde beibehalten, führte jedoch nur zu einer ungenaueren Vorne-Hinten-Lokalisation – bei einzelnen Personen sogar nur auf Zufallsniveau. Die Autoren schlussfolgern, dass sowohl monaurale spektrale Hinweise als auch das interaurale Pegelspektrum für eine genaue Lokalisation wichtig sind.

#### 4.3. Informatorische Maskierung und Aufmerksamkeitsausrichtung

Informatorische Maskierung (oder zentrale Maskierung – englisch *informational masking* oder *central masking*) und energetische Maskierung werden häufig durch Gegenüberstellung von einander abgegrenzt (Durlach et al., 2003). Zu energetischer Maskierung (oder peripherer Maskierung – englisch *energetic masking* oder *peripheral masking*) kommt es, wenn in den Frequenzbändern des auditiven Systems jeweils nicht nur die Aktivität des Signals, sondern auch des Maskierers vorzufinden ist (Moore, 2004; Richards & Kidd, 2009). Obwohl es verschiedene Auffassungen hinsichtlich des Mechanismus der Maskierung gibt, ist energetische Maskierung vereinfacht gesagt umso größer, je höher der Pegel des Maskierers ist und je mehr der Maskierer dem Signal hinsichtlich des Frequenzspektrums ähnelt. Informatorische Maskierung bezeichnet demgegenüber alle Maskierung, die nicht als energetische Maskierung verstanden werden kann und vermutlich auf späteren Verarbeitungsstufen stattfindet.

Unsicherheit eines Reizes, d.h. ein von den Erwartungen abweichender oder nicht genau vorhersagbarer Reiz, ist zwar weder nötig noch ausreichend, um informatorische Maskierung auszulösen, wird aber als einer der beiden entscheidenden modulierenden Faktoren angesehen (Durlach et al., 2003). Der zweite Faktor ist die Ähnlichkeit zwischen Ziel- und Distraktorreiz, wobei eine objektive Definition von Ähnlichkeit Schwierigkeiten bereitet und der Zielreiz auch erkannt werden muss, wenn er sich mit dem Distraktor überlappt. Laut Durlach et al. wird Ähnlichkeit durch vergleichbare Eigenschaften hervorgerufen wie die Gruppierung innerhalb der auditorischen Szenenanalyse nach Bregman (1990). Verei-

facht gesagt tritt informatorische Maskierung vermehrt auf, wenn Verwechslungsgefahr zwischen Ziel und Distraktor besteht.

Informatorische Maskierung kann individuell sehr unterschiedlich stark ausgeprägt sein. Während eine Distraktor-Unsicherheit in einer Situation mit großer Ähnlichkeit zwischen Ziel und Distraktor bei manchen Personen zu einer Schwellenerhöhung um bis zu 40 dB führen kann, gibt es auch Personen, die nicht sensitiv gegenüber einer solchen Unsicherheit zu sein scheinen (Durlach et al., 2003). Oxenham, Fligor, Mason und Kidd (2003) verglichen Musiker und Nicht-Musiker im Hinblick auf informatorische Maskierung und fanden eine geringere Beeinträchtigung bei Musikern. Nach Meinung der Autoren geht dieses Ergebnis auf Unterschiede in der Fähigkeit zum analytischen Hören zurück.

Als ein Phänomen, das im Zusammenhang mit informatorischer Maskierung steht, kann die akustische Aufmerksamkeit betrachtet werden. Spence und Driver (1994) untersuchten Aufmerksamkeitsmechanismen in Aufgaben der akustischen Wahrnehmung, die auch an der Entstehung des Effektes räumlich separierter Schallquellen beteiligt sein könnten. In einer Reihe von Untersuchungen zeigten die Autoren, dass verdeckte Ausrichtung der Aufmerksamkeit auch in der auditiven Modalität zu einer schnelleren Bearbeitung von Wahlreaktionsaufgaben führen kann. Die teilnehmenden Personen sollten unter anderem beurteilen, ob Zielgeräusche, die auf jeder Seite auftreten konnten, eher oben oder unten präsentiert wurden. Zeitlich vor den zu beurteilenden Reizen wurde ein akustischer, uninformativer (d.h. räumlich vom Zielreiz unabhängiger) Hinweisreiz entweder links oder rechts dargeboten. Erschien der Hinweisreiz auf der gleichen Seite wie der Zielreiz (valide Durchgänge), so zeigten die Personen bessere Leistungen, als wenn der Hinweisreiz auf der anderen Seite erschien (invalide Durchgänge). Dies legt den Schluss nahe, dass exogene Mechanismen zur Ausrichtung der Aufmerksamkeit in der auditiven Modalität existieren. Außerdem wurden endogene Mechanismen untersucht, indem der Hinweisreiz in 75% der Durchgänge auf der gleichen Seite wie der Zielreiz und in 25% auf der anderen Seite auftrat. Die Aufgabe war in dieser Untersuchung zu beurteilen, ob der Zielreiz eher von vorne oder hinten präsentiert wurde. Auch hier war die Leistung in der validen Bedingung besser als in der invaliden.

Dass die horizontale Ausrichtung der Aufmerksamkeit zu einem besseren Erkennen eines akustischen Zielreizes führen kann, konnten auch Kidd, Arbogast, Mason und Frederick (2005) zeigen. Dazu präsentierten sie drei zufällig ausgewählte Zielsätze, deren am Ende gesprochene Wörter wiedergegeben werden sollten. War der Ort des Zielsatzes bekannt, so zeigten die Teilnehmer bessere Leistungen, als wenn der Ort unbekannt war – unabhängig davon, ob ein Hinweiswort im Zielsatz vor oder nach einem Durchgang bekannt gegeben wurde.

Ein Vorteil durch räumliche Ausrichtung von Aufmerksamkeit ist auch in Situationen mit räumlicher Separation wahrscheinlich, in denen keine oder wenig Maskierung auftreten kann, da die Reize hintereinander präsentiert werden. So konnte Baldis (2001) zeigen, dass eine räumliche Separation verschiedener Sprecher bei Telefon- und Videokonferenzen das Quellengedächtnis, das subjektive Behalten von geäußerten Meinungen zu einem kontroversen Thema und die Präferenz der Teilnehmer positiv beeinflussen kann. Auch bei Doll und Hanna (1995) waren die Reize über einen längeren Zeitraum verteilt und traten meist nicht gleichzeitig auf, so dass eine energetische Maskierung unwahrscheinlich ist. Dennoch finden auch sie einen Vorteil räumlich separierter Präsentation der akustischen Sequenzen bei der Detektion von Zielreizen. Nach der Interpretation der Autoren ergibt sich der Vorteil durch präattentive parallele Verarbeitung der Reize, die daher als verschiedene Objekte wahrgenommen werden.

#### 4.4. Richtungshinweise und ihr Einfluss auf den Separationseffekt

Im Folgenden wird die Rolle von interauralen Pegel- und Zeitdifferenzen, monauralen Hinweisen und informatorischer Maskierung im Zusammenhang mit dem Effekt der räumlichen Separation beschrieben. In den zwei Untersuchungen, die hierzu näher betrachtet werden (Bronkhorst & Plomp, 1988; Culling, Hawley, & Litovsky, 2004), wurden die Positionen der Reize mittels kopfbezogener Übertragungsfunktionen simuliert, gegebenenfalls interaurale Pegeldifferenzen oder interaurale Zeitdifferenzen aus den Signalen entfernt und diese per Kopfhörer präsentiert.

Bronkhorst und Plomp (1988) bestimmten Sprachverständlichkeitsschwellen für von vorne ( $0^\circ$ ) präsentierte Sätze, die mit spektral ähnlichem Rauschen aus Richtungen zwischen  $30^\circ$  und  $150^\circ$  in der horizontalen Ebene überlagert waren. Ein Vergleich von Bedingungen, in denen die Richtung des Rauschens durch interaurale Pegeldifferenzen, interaurale Zeitdifferenzen oder beides dargestellt wurde, zeigte, dass sowohl interaurale Pegeldifferenzen als auch interaurale Zeitdifferenzen wesentliche Beiträge zum Separationseffekt liefern. Der Separationseffekt in der Bedingung mit beiden interauralen Differenzen fiel jedoch etwas geringer aus als die Summe der Effekte in den Bedingungen mit nur interauralen Pegeldifferenzen oder interauralen Zeitdifferenzen. Diese Unter-Additivität von Effekten durch interaurale Pegel- und Zeitdifferenzen konnte in einer weiteren Untersuchung der Autoren bestätigt werden (Bronkhorst & Plomp, 1992). Interaurale Pegel- und Zeitdifferenzen scheinen folglich zu interagieren. Dies ist zumindest im Extremfall leicht einzusehen, nämlich wenn auf der einen Seite kein Signal präsentiert wird und somit große interaurale Pegeldifferenzen vorliegen. In dieser Situation könnten eventuelle zusätzliche interaurale Zeitdifferenzen offensichtlich gar nicht dargestellt oder wahrgenommen werden.

Auch Culling, Hawley und Litovsky (2004) untersuchten den Einfluss von interauralen Pegel- und Zeitdifferenzen auf die Sprachverständlichkeitsschwelle mit Hilfe simulierter Stimuluspositionen. Dabei wurde als Zielreiz Sprache von vorne ( $0^\circ$ ) präsentiert und drei Distraktoren (Sprache oder spektral ähnliches Rauschen) mit unterschiedlicher Anordnung. In einer Bedingung kamen die Distraktoren ebenfalls von vorne (Bedingung  $0^\circ 0^\circ 0^\circ$ ), in einer anderen von beiden Seiten der Personen ( $-30^\circ 60^\circ 90^\circ$ ), in einer weiteren von verschiedenen Orten auf einer Seite ( $30^\circ 60^\circ 90^\circ$ ) und in einer letzten Bedingung von einem Ort auf einer Seite ( $90^\circ 90^\circ 90^\circ$ ). Die räumliche Position der Distraktoren wurde dabei entweder durch interaurale Pegeldifferenzen, Zeitdifferenzen oder beides simuliert. Obwohl die Effekte der räumlichen Konfiguration in der Bedingung mit sprachähnlichem Rauschen kleiner ausfielen (und die Sprachverständlichkeitsschwellen höher waren) als in der Bedingung mit Sprache, führten beide Distraktorarten zu einem ähnlichen Ergebnismuster: Wenn interaurale Zeitdifferenzen oder sowohl Zeit- als auch Pegeldifferenzen verwendet wurden, führten alle räumlich separierten Konfigurationen zu einem Vorteil

gegenüber der Konfiguration mit allen Distraktoren von vorne. In Bedingungen, die nur Pegeldifferenzen enthielten, konnte ein Separationsvorteil jedoch lediglich für die Konfigurationen mit Distraktoren auf nur einer Seite gefunden werden. Wenn sich die Distraktoren jedoch auf beiden Seiten befanden, blieb der Separationsvorteil aus.

Während der Beitrag der interauralen Pegeldifferenzen zum Separationseffekt in Wesentlichen durch ein besseres Zielreiz-Distraktor-Lautstärkeverhältnis erklärt wird, nimmt man an, dass interaurale Zeitdifferenzen mit Hilfe von Neuronen ausgewertet werden, die ähnlich einer Kreuzkorrelation in schmalen Frequenzbändern arbeiten (Shinn-Cunningham, 2005). Befinden sich Ziel und Distraktor an der gleichen Position, so liefert eine Kreuzkorrelation der Signale von den beiden Ohren eine konstante Zeitdifferenz. Bei unterschiedlichen Positionen von Ziel und Distraktor ergibt sich jedoch in einem Frequenzband, das von beiden Reizen aktiviert wird, eine schwankende Zeitdifferenz. Dies kommt dadurch zustande, dass die Signale der beiden Ohren sich bei lateral separierten Schallquellen nicht nur hinsichtlich Zeit und Pegel unterscheiden, sondern auch verschiedene Wellenformen enthalten, da die Schallquellen in der Regel auch verschiedene Wellenformen besitzen. Auf diese Weise können Frequenzbereiche des Zielreizes unter Umständen auch dann detektiert und spektrotemporale Information gewonnen werden, wenn der Zielreiz bei monauralem Hören vollständig von dem Distraktor maskiert würde.

Weitere Untersuchungen (Bronkhorst & Plomp, 1992; Hawley, Litovsky, & Culling, 2004) gingen der Frage nach, welche Rolle monaurale Hinweise beim Effekt der räumlichen Separation spielen. In diesen Studien sollten die Teilnehmer von vorne präsentierte Sprache beachten und Distraktoren an unterschiedlichen Orten ignorieren. Neben der Anzahl und Anordnung der Distraktoren wurde auch variiert, auf wie vielen Ohren die Darbietung erfolgte (monaural oder binaural). In den monauralen Bedingungen enthielten die Stimuli jedoch sowohl spektrale Hinweise als auch Hinweise bezüglich des Zielreiz-Distraktor-Lautstärkeverhältnisses, das an einem der beiden Ohren anlag. Wegen ihrer geringen Bedeutung bei der Lokalisation finden die spektralen Hinweise in der Interpretation der Ergebnisse jedoch generell oft keine Berücksichtigung. Als entscheidend für den Vorteil durch monaurale Hinweise wird hingegen das bessere Signal-Rausch-Verhältnis ange-

nommen, das entsteht, wenn der Zielreiz sich näher am Ohr befindet (oder weniger stark durch den Kopf abgeschattet wird) als die Maskierer.

Um den Separationsvorteil zu bestimmen, der allein auf monaurale Hinweise zurückgeht, verglichen Bronkhorst und Plomp (1992) eine monaurale Bedingung mit frontalem Zielreiz und um 90° separiertem Rauschen mit einer Bedingung, in der beide Reize von vorne präsentiert wurden.<sup>4</sup> Sie fanden allein durch monaurale Hinweise einen Vorteil von annähernd 6 dB in der Sprachverständlichkeitsschwelle. Befand sich jedoch auf der anderen Seite ebenfalls ein um 90° separierter Distraktor, konnte kaum noch ein Vorteil gegenüber einer Bedingung mit allen Reizen von vorne gefunden werden (0,3 dB). Ein vergleichbar kleiner Wert wurde erreicht, wenn auf der einen Seite drei und auf der anderen Seite ein Distraktor vorhanden war – vorausgesetzt die günstigere (d.h. den drei Distraktoren abgewandte) Seite wurde präsentiert, sonst zeigte sich ein Nachteil von etwa 2 dB. Bei monauraler Präsentation einer Hörsituation von je zwei oder je drei Reizen links wie rechts konnte ebenfalls nur ein Nachteil gefunden werden – und zwar in einer Größenordnung von etwa 1 dB.

Diese Ergebnisse lassen sich gut mit einer Untersuchung von Hawley et al. (2004) vereinbaren, die Rauschen oder auch Sprache als Distraktoren verwandten. Die räumliche Separation war in den monauralen Bedingungen auch nur dann von Vorteil, wenn sich die ein bis drei Maskierer ausschließlich auf der abgewandten Seite befanden. Sobald in einer monauralen Bedingung ein Maskierer auf der Seite des hörenden Ohrs präsentiert wurde, konnte monaural kein Separationseffekt mehr gefunden werden – selbst wenn sich zwei von drei Maskierern auf der abgewandten Seite befanden. Monaurale Hinweise scheinen also keine wesentliche Rolle mehr zu spielen, wenn sich Maskierer aus Sicht des Hörers gleichzeitig links und rechts von dem Zielreiz befinden.

Arbogast, Mason und Kidd (2002) gingen der Frage nach, ob eine räumliche Separation der Schallquellen im gleichen Maße energetische und informatorische Maskierung reduziert. Dazu präsentierten sie den teilnehmenden Personen zu beachtende, modifizierte

---

<sup>4</sup> Letztere Bedingung war zwar binaural, da es keine solche monaurale Bedingung gab; jedoch sind bei einer Präsentation aller Reize von vorne auch kaum binaurale Vorteile zu erwarten (verglichen mit einer *monauralen* Präsentation aller Reize von vorne).

Sprache von vorne und einen Distraktor (modifizierte Sprache oder Rauschen) entweder auch von vorne oder von rechts. Um Distraktoren mit viel oder wenig energetischer oder auch informatorische Maskierung erstellen zu können, wurde Sprache in 15 Frequenzbänder zerlegt, und 15 Sinustöne mit den zentralen Frequenzen der Bänder wurden mit den Lautstärkehüllkurven der jeweiligen Frequenzbänder moduliert. In jedem Durchgang wurden für den Zielreiz zufällig acht Bänder ausgewählt, während für die Distraktoren in der Bedingung „verschiedene Sprachbänder“ zufällig sechs der sieben nicht für den Zielreiz verwendeten Bänder ausgewählt wurden, so dass hier höchstens eine geringe energetische Maskierung zu erwarten ist. In der Bedingung „verschiedene Rauschbänder“ wurde die Summe der sechs Bänder zusätzlich mit Gaußschem Rauschen verrechnet, was ebenso zu geringer energetischer Maskierung, aber auch zu einem dem Zielreiz unähnlichen nämlich rauschähnlichen Klangeindruck ohne Sprachverständlichkeit führte. Schließlich wurden in der Bedingung „gleiche Rauschbänder“ die gleichen acht Bänder wie für den Zielreiz verwendet und eine Faltung mit Rauschen durchgeführt, so dass hier eine starke energetische Maskierung zu erwarten ist. Die Personen sollten Tasten für die in dem Zielsatz vorkommenden Farben und Nummern drücken, wobei die Lautstärke des Distraktors von Durchgang zu Durchgang variierte.

In der Bedingung „verschiedene Sprachbänder“ zeigte sich mit etwa 18 dB ein größerer Vorteil durch die räumliche Separation als in den Bedingungen „verschiedene Rauschbänder“ und „gleiche Rauschbänder“, in denen der Separationseffekt etwa 4 bzw. 7 dB betrug. In der ersten Bedingung wurden modulierte Sinustöne nicht nur beim Zielreiz, sondern auch beim Distraktor dargeboten, was die Ähnlichkeit zwischen den auseinanderzuhaltenden beiden Signalen erhöhte und somit für höhere informatorische Maskierung spricht. In den beiden anderen Bedingungen hingegen war die Ähnlichkeit und folglich auch die informatorische Maskierung geringer, da ein rauschähnliches Signal als Distraktor diente. In der dritten Bedingung trat in stärkerem Maße energetische Maskierung auf, da Ziel- und Distraktorsignal die gleichen Frequenzbereiche umfassten; dadurch kam es jedoch nur zu einem kleineren Separationseffekt. Der Vorteil räumlicher Separation von Ziel und Distraktor wirkte sich hier bei informatorischer Maskierung also stärker aus als bei energetischer Maskierung.

Best, Ozmeral, Gallun, Sen und Shinn-Cunningham (2005) untersuchten energetische und informatorische Maskierung, indem sie Teilnehmer Vogelgesänge erkennen ließen. Die räumlich separierte oder nicht separierte Präsentation erfolgte dabei im virtuellen Raum über Kopfhörer, wobei es auch eine diotische Kontrollbedingung gab, in der das Signal des „besseren Ohrs“ (d.h. dem Ziel zugewandten Ohr) auf beiden Ohren dargeboten wurde. Als Maskierer dienten Gemische aus drei noch nicht gehörten Vogelgesängen, ein Rauschen mit dem gleichen Spektrum wie der Zielreiz oder letzteres Rauschen zusätzlich mit der Lautstärkehüllkurve eines Gemisches aus drei Vogelgesängen moduliert.

Der Vorteil der räumlichen Separation zeigte sich in allen Bedingungen, konnte jedoch bei dem Gemisch der Vogelgesänge nicht alleine durch das bessere Signal-Rauschverhältnis auf dem einen Ohr der Teilnehmer erklärt werden; hier war die Leistung in der diotischen Bedingung nämlich geringer. Da dieser Maskierer dem Zielreiz spektrotemporal am ähnlichsten war und selbst bei günstigem Zielreiz-Maskiererlautstärkeverhältnis in der nicht separierten Bedingung zu relativ vielen Fehlern führte, erscheint ein höherer Einfluss informatorischer Maskierung bei diesem Maskierer im Vergleich zu den beiden Rausch-Maskierern plausibel. Auch wenn die Ziele in dieser Studie nie als Distraktoren verwendet wurden und sich außerdem stets an der gleichen Position befanden, deuten die Ergebnisse darauf hin, dass energetische sowie informatorische Maskierung auch bei nicht sprachlichen Reizen, nämlich Vogelgesängen, durch räumliche Separation reduziert werden kann.

#### 4.5. Fazit

Sowohl interaurale Pegeldifferenzen als auch interaurale Zeitdifferenzen können also erheblich zu dem Effekt der räumlichen Separation beitragen. Interaurale Pegeldifferenzen spielen die wichtigere Rolle, wenn sich der Zielreiz weiter rechts oder links als alle Maskierer befindet. In Situation, in denen sowohl links als auch rechts des Zielreizes mindestens ein Maskierer präsentiert wird, haben hingegen interaurale Zeitdifferenzen die größere Bedeutung für den Effekt der räumlichen Separation. Monaurale Lautstärke-Hinweise sind im Allgemeinen in interauralen Pegeldifferenzen enthalten und zeigen auch einen ähnlichen Zusammenhang mit dem Effekt der räumlichen Separation wie letztere: Nur wenn sich Distraktoren (aus Sicht der hörenden Person) nicht sowohl links als auch rechts

des Zielreizes befinden, kann eine wesentliche Mitwirkung von monauralen Hinweisen am Separationseffekt erwartet werden.

Während interaurale Zeit- und Pegeldifferenzen auf einer frühen Verarbeitungsstufe analysiert werden, finden die informatorische Maskierung und die Ausrichtung von Aufmerksamkeit wahrscheinlich auf einer höheren Ebene neuronaler Verarbeitung statt. Eine räumlich separierte Anordnung von Schallquellen reduziert neben energetischer Maskierung auch informatorische Maskierung – wahrscheinlich sogar in noch höherem Maße. Etwas schwierig erscheint es, informatorische Maskierung von Aufmerksamkeitsprozessen abzugrenzen. Da die räumliche Ausrichtung auditorischer Aufmerksamkeit zu einer schnelleren Bearbeitung von Reaktionsaufgaben führen kann, ist jedenfalls auch eine Beteiligung von Aufmerksamkeitsprozessen an dem Vorteil räumlicher Separation zu erwarten.

## 5. Akustische Suche in alltagsnäheren Situationen

Der Effekt räumlicher Separation wurde zumeist unter stark vereinfachten Bedingungen untersucht – mit oft nur wenigen Distraktoren, Präsentation des Zielreizes von einem festen Ort und Verwendung von Sprache als Zielreiz sowie Sprache oder Rauschen als Distraktoren. Beispielsweise lassen sich mehrere Studien identifizieren, in denen als Zielreiz ein sprachähnliches Signal von vorne und als Distraktor ein sprachähnliches Signal oder Rauschen von vorne in der nicht separierten Bedingung und von einer anderen Position in der separierten Bedingung präsentiert wurde, wobei unter Umständen nur bestimmte räumliche Hinweise implementiert wurden (Arbogast et al., 2002; Arbogast, Mason, & Kidd, 2005; Bronkhorst & Plomp, 1988, 1990; Garadat, Litovsky, Yu, & Zeng, 2009). Gelegentlich weichen Studien in einem der zentralen Aspekte von dieser Standard-situation ab, selten jedoch in mehreren oder gar allen Aspekten gleichzeitig.

Diese Beschränkung bisheriger Studien zum Effekt räumlich separierter Schallquellen auf derart artifizielle Bedingungen unterscheidet sich von realistischen Situationen in mindestens dreierlei Hinsicht: Im Alltag sind wir oft von zahlreichen akustischen Reizen gleichzeitig umgeben, von denen nicht alle für uns relevant sind. Zusätzlich handelt es sich bei den Schallereignissen oft nicht um homogene Reize einer einzelnen Kategorie wie etwa Sprache, und schließlich sind die Richtungen, aus denen der Schall kommt nicht immer vorhersehbar und auch nicht unbedingt über die Zeit gleichbleibend. Bereits ein kurzer Spaziergang durch einen Park offenbart, dass es viele verschiedene Schallquellen gibt (z.B. Sprache, Gelächter, Vogelgezwitscher, Hundebellen, Blätterrascheln, Schrittgeräusche, Fahrradklingeln etc.), die den verschiedensten unerwarteten und auch wechselnden Richtungen entspringen können. Eine Beschäftigung mit solchen alltagsrelevanten Situationen ist nicht nur sinnvoll im Hinblick auf einen eventuellen Anwendungsbezug, sondern auch notwendig, um die Bedeutung des Effektes räumlicher Separation für solche Bedingungen beurteilen zu können. In den folgenden Abschnitten 5.1 bis 5.3 wird auf die Bedeutung der einzelnen Aspekte – Anzahl der akustischen Reize, Sprache versus Umweltgeräusche und Vorhersagbarkeit räumlicher Positionen von Schallquellen – näher eingegangen, bevor in

Abschnitt 5.4 und 5.5 der bestehende Klärungsbedarf und die einzige dazu bisher vorliegende Untersuchung (Eramudugolla, McAnally, Martin, Irvine, & Mattingley, 2008) behandelt werden.

### 5.1. Anzahl der akustischen Reize

Unter den zahlreichen Untersuchungen zum Effekt der räumlichen Separation (für Übersichtsartikel siehe Bronkhorst, 2000; Ebata, 2003) lassen sich 17 Untersuchungen identifizieren, in denen vier oder mehr Reize gleichzeitig präsentiert wurden (Abouchacra et al., 2001; Bronkhorst & Plomp, 1992; Culling et al., 2004; Drullman & Bronkhorst, 2000; Ericson & McKinley, 1997; Hawley et al., 2004; MacDonald et al., 2002; McAnally & Martin, 2007; Nelson, Bolia, Ericson, & McKinley, 1998a, 1998b, 1999; Peissig & Kollmeier, 1997; siehe Tabelle 1). Die bisherige Literatur zeichnet jedoch ein nur uneindeutiges Bild, was den Einfluss der Anzahl auf den Effekt räumlich separierter Schallquellen angeht.

Hawley et al. (2004) konnten zeigen, dass sich Einflussfaktoren auf den Effekt räumlicher Separation in Abhängigkeit von der Anzahl der Reize unterschiedlich auswirken können. Zweiunddreissig Teilnehmer sollten Sprache erkennen, die gleichzeitig mit ein, zwei oder drei Distraktoren dargeboten wurde, wobei die Präsentation der Reize für die eine Hälfte der Teilnehmer binaural und für die andere monaural erfolgte. Durch den Vergleich der Leistung in den Gruppen mit binauraler und monauraler Reizpräsentation sollte die Bedeutung binauraler Hinweise im Vergleich zur Bedeutung monauraler Hinweise auf den Separierungseffekt analysiert werden.

Auch wenn Hawley et al. (2004) Additivität von monauralen und binauralen Effekten annehmen und den Vergleich verschiedener Anzahlen an Reizen nicht statistisch testeten, deuten die Ergebnisse darauf hin, dass es auch von der Anzahl der verwendeten Reize abhängt, wie groß der Einfluss monauraler und binauraler Effekte bei Sprachverständlichkeitsaufgaben mit räumlicher Separation der Reize ausfällt: In Situationen mit mehreren Distraktoren fiel der Effekt binauraler Hinweise – abhängig vom Reizmaterial – größer und der Effekt monauraler Hinweise kleiner aus als bei nur einem Distraktor. Zu ähnlichen Schlussfolgerungen kommen auch Yost, Dye und Sheft (1996).

Tabelle 1

Untersuchungen mit vier oder mehr gleichzeitig präsentierten akustischen Reizen und Variation der räumlichen Separation der Schallquellen.

Autoren	Jahr	Anzahl simultaner Reize	Zielreiz	Distraktor	Vorhersagbarkeit der Ziellokation	Zielreiz hat durchgehend Merkmale, die sich von Distraktoren unterscheiden	Anzahl Teilnehmer	Bemerkungen
Bronkhorst et al.	1992	2–7	Sprache	Moduliertes Rauschen	ja	ja	17	Kopfhörer, zusätzlich Untersuchung bei Schwerhörigkeit
Peissig et al.	1997	2–4	Sprache	Sprache oder Rauschen	ja	ja (zum Teil)	8	Kopfhörer, keine Separation außer für einen Distraktor
Ericson et al., Exp. 1	1997	4	Sprache	Sprache	nein (?)	ja (2 Hinweise)	4	Kopfhörer
Ericson et al., Exp. 2	1997	4	Sprache	Sprache	ja und nein (je nach Bedingung)	?	?	Kopfhörer
Nelson et al.	1998	1–8	Sprache	Sprache	nein (?)	nein (?)	8	
Nelson et al.	1998	1–8	Sprache	Sprache	nein (?)	nein (aber gleicher Hinweis)	8	Kopfhörer
Drullmann et al. Exp. 1	1999	2–5	Sprache	Sprache	blockweise ja	ja	12	Kopfhörer, Reize 4kHz-Tiefpass-gefiltert
Drullmann et al. Exp. 2	1999	2–5	Sprache	Sprache	nein	ja	10	Kopfhörer, Reize 4kHz-Tiefpass-gefiltert
Nelson et al.	1999	1–8	Sprache	Sprache	nein (?)	nein (aber gleiches <i>callsign</i> )	8	Kopfhörer oder Freifeld (intraindividuell)
Abouchacra et al.	2001	1–5	Sprache	3 x Sprache und Fahrgeräusche	nein (bei diotischer Präsentation ja)	nein (aber gleiches <i>callsign</i> )	8	Kopfhörer, Reihenfolgeeffekt in Anzahl der Reize
MacDonald et al.	2002	4	Sprache	Sprache	nein (?)	nein	30	Kopfhörer, keine Auswertung von falschen Alarmen oder Reaktionszeiten, visuelle Zusatzaufgabe
Culling et al.	2004	4	Sprache	Sprache, Rauschen mit sprachähnlichem Spektrum	ja	nein	36	Kopfhörer
Hawley et al.	2004	2–4	Sprache	Sprache rückwärts, Rauschen mit Sprachspektrum oder sprachähnliches Rauschen	ja	nein	32 (16 monaural und 16 binaural)	Kopfhörer, Distraktortext war bekannt
Best et al.	2005	2–4 (bei 4 kamen 3 von einer Position)	Gesang von Zebrafinken	Stimmengemisch von Zebrafinken, Rauschen oder moduliertes Rauschen	ja	ja	5	Kopfhörer, Identifikations-Trainingsphase

Autoren	Jahr	Anzahl simultaner Reize	Zielreiz	Distraktor	Vorhersagbarkeit der Ziellokation	Zielreiz hat durchgehend Merkmale, die sich von Distraktoren unterscheiden	Anzahl Teilnehmer	Bemerkungen
McAnally et al.	2007	0–5 (zufällige Abfolge, im Durchschnitt 2)	Sprache	Sprache	nein	nein (aber gleicher Hinweis)	6	Kopfhörer, seltene Zielreize
Eramudugolla et al. Exp. 1	2008	4, 6 oder 8	Umweltgeräusche	Umweltgeräusche	nein	nein	28	Kopfhörer, Geschw.-Genauigk.-Austausch
Eramudugolla et al. Exp. 2	2008	4, 6 oder 8	Umweltgeräusche	Umweltgeräusche	nein	nein	17	Kopfhörer, keine Reaktionszeiten

Nelson, Bolia, Ericson und Kinley (1998b) realisierten mit ein bis acht Reizen gleichzeitig eine weit gespannte Manipulation der Anzahl an Reizen und fanden dennoch keine Interaktion zwischen räumlicher Separation und Anzahl an Reizen. Sie setzten eine modifizierte Version des *Coordinate Response Measure* ein, innerhalb dessen alle Sätze die Form „Ready [Hinweiswort] go to [Farbe] [Zahl] now.“ besaßen. Die Teilnehmer sollten Nummern und Farben in Zielsätzen mit einem bestimmten Hinweiswort erkennen, während diese gleichzeitig mit bis zu sieben Distraktorsätzen (mit anderen Hinweiswörtern und irrelevanten Nummern und Farben) präsentiert wurden. Abgesehen davon, dass auch hier ausschließlich Sprache als Reizmaterial zum Einsatz kam, wurde bei dieser Untersuchung wie bei vielen anderen mit acht Teilnehmern nur eine geringe Teststärke erreicht. Unter Annahme einer Populationskorrelation von  $\rho = .5$  und  $\alpha = .05$  konnten dann mittelgroße Interaktionseffekte nach Cohen (1988) von  $f = 0.25$  mit der Variable räumliche Separation lediglich mit einer Teststärke von weniger als 50% nachgewiesen werden<sup>5</sup>. In solchen Situationen ist es schwierig zu beurteilen, ob die Anzahl der verwendeten Reize tatsächlich keine Rolle spielt, oder ob die mangelnde Teststärke hier für das Ausbleiben von Interaktionen verantwortlich ist.

Daher scheint also bisher nicht eindeutig geklärt zu sein, ob der Effekt räumlich separierter Schallquellen bei hohen Anzahlen an gleichzeitig präsentierten Reizen gleichermaßen

<sup>5</sup> Alle Teststärkeanalysen wurden mit G\*Power berechnet (Faul, Erdfelder, Lang, & Buchner, 2007).

zu beobachten ist oder ob es möglicherweise Abhängigkeiten von weiteren Variablen gibt – und ebenso, ob die gleichen Mechanismen zu einem eventuellen Vorteil räumlicher Separation bei einer größeren Anzahl an Reizen beitragen. Hierzu sind Untersuchungen nötig, die nicht nur eine hohe Anzahl an Reizen verwirklichen, sondern auch die nötige Teststärke besitzen, um eventuelle Interaktionen beispielsweise zwischen Anzahl und Separation aufdecken zu können.

## 5.2. Sprache versus Umweltgeräusche

Wie Tabelle 1 zeigt, gibt es etwa 17 Untersuchungen, in denen vier oder mehr Reize gleichzeitig präsentiert wurden. Davon verwendeten 14 Untersuchungen Sprache als Zielreiz (Abouchacra et al., 2001; Bronkhorst & Plomp, 1992; Culling et al., 2004; Drullman & Bronkhorst, 2000; Ericson & McKinley, 1997; Hawley et al., 2004; MacDonald et al., 2002; McAnally & Martin, 2007; Nelson et al., 1998a, 1998b, 1999; Peissig & Kollmeier, 1997).

Laut Byrne et al. (1994) zeichnet sich Sprache durch bestimmte Eigenschaften aus, die über verschiedene Sprachen hinweg weitgehend gleich bleiben. Berücksichtigt wurden dabei zwölf Sprachen: Dänisch, Deutsch, Englisch, Französisch, Japanisch, Kantonesisch, Mandarin, Russisch, Schwedisch, Singhalesisch, Vietnamesisch und Walisisch. Das Durchschnittsspektrum aller Sprachaufnahmen zeigte den höchsten Pegel im Bereich von 200 bis 600 Hz (siehe auch Abbildung 1). Bei den Frauen nahm der Pegel darunter stark ab, während dies bei den Männern erst unterhalb von 100 Hz der Fall war. Zu hohen Frequenzen hin fiel der Pegel in beiden Fällen kontinuierlich, aber weniger steil ab. Obwohl Byrne et al. (1994) Unterschiede zwischen verschiedenen Sprechern fanden, blieben die Standardabweichungen für beide Geschlechter unter 7 dB in allen Frequenzbereichen, wobei jeder Sprecher auf 70 dB Gesamtpegel normalisiert wurde. Sprache scheint also eine recht homogene Klasse von Geräuschen zu sein, innerhalb der nur wenig von ihrem typischen Durchschnittsspektrum abgewichen wird.

Neben dem typischen Durchschnittsspektrum besitzt Sprache etwa auch die Eigenheit, dass sich kein bestimmter Frequenzbereich identifizieren lässt, der für das Verstehen notwendig wäre (Moore, 2004). Beispielsweise sind Silben, die nur Frequenzen ober- oder un-

terhalb von 1800 Hz enthalten, zu 67% verständlich – normale Unterhaltungen sogar vollständig. Für Sätze genügt ein Frequenzband zwischen 1000 und 2000 Hz, um 90% Verständlichkeit zu erreichen.

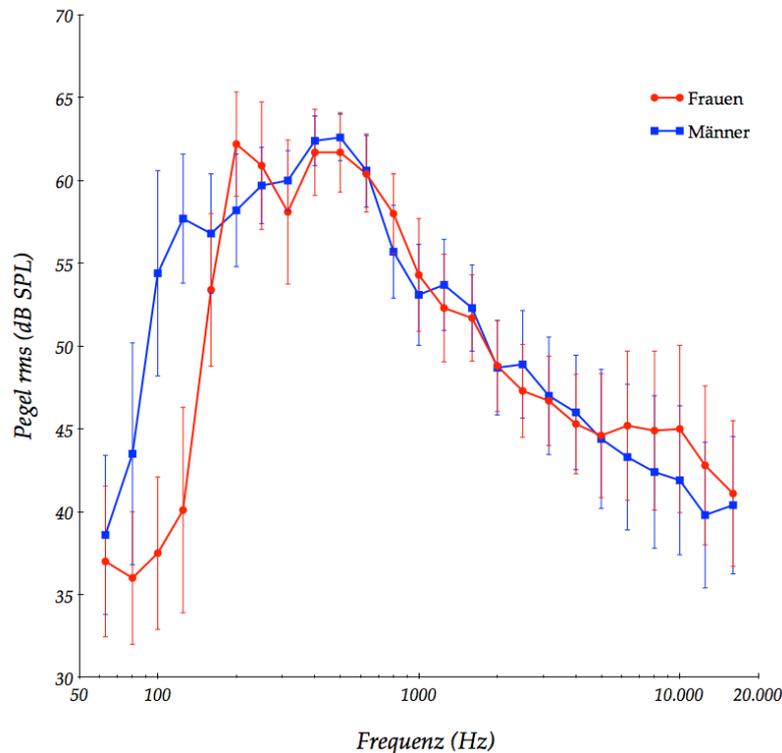


Abbildung 1: Langzeitspektrum von weiblicher und männlicher Sprache. Durchschnitt aus 17 Aufnahmen verschiedener Sprachen, jeweils auf 70 dB SPL normalisiert (nach Byrne et al., 1994). Die Fehlerbalken zeigen Standardabweichungen.

Gygi, Kidd und Watson (2004) weisen auf eine größere Variabilität zwischen zufällig ausgewählten Umweltgeräuschen hin als zwischen Sprachaufnahmen, da auch die Objekte oder Ereignisse, die die Umweltgeräusche verursachen, stärker variieren. In ihrer Studie, in der Personen Umweltgeräusche erkennen sollten, beobachteten Gygi et al., dass bei Umweltgeräuschen und Sprache zum Großteil zwar gleiche Frequenzbereiche für das Erkennen wichtig sind, höhere Frequenzbereiche bis 8000 Hz bei Umweltgeräuschen jedoch mehr Information zu besitzen scheinen. Chen und Spence (2011) konnten außerdem zeigen, dass Sprache und naturalistische Geräusche in semantischen Primingaufgaben unterschiedlich verarbeitet werden. Die Teilnehmer zeigten eine höhere Sensitivität bei visuel-

len Detektionsaufgaben, wenn kurz zuvor naturalistische Geräusche präsentiert wurden, nicht jedoch wenn dies gesprochene Wörter waren.

Schließlich lassen sich auch Unterschiede in der neuronalen Verarbeitung von Sprache im Vergleich zu nicht-sprachlichen Geräuschen wie Sinustönen oder Rauschen finden (Wolfe et al., 2009). Sowohl in der linken als in der rechten Hemisphäre zeigt sich mehr Aktivität im oberen Temporallappen, wenn Sprache statt anderer Geräusche gehört wird. Außerdem konnte bei Sprachreizen eine weiter anterior und ventral gelegene Aktivierung beobachtet werden als bei ähnlich komplexen, aber nicht verständlichen Vergleichsreizen.

Da bei hoher Reizähnlichkeit nicht nur energetische Maskierung, sondern auch Phänomene wie Aufmerksamkeitsprozesse oder informatorische Maskierung eine Rolle spielen und diese einen Einfluss auf den Separationseffekt haben können (Arbogast et al., 2002; Best et al., 2005), müssen Ergebnisse aus Untersuchungen mit Sprachreizen nicht unbedingt gleichermaßen für Situationen mit Umgebungsgeräuschen gelten. Möglicherweise werden räumliche Hinweise vor allem dann genutzt, wenn andere (spektrotemporale) Segmentierungshinweise schwach ausfallen. Dann wäre ein geringerer Effekt räumlicher Separation bei naturalistischen Geräuschen zu erwarten, da diese ja in ihren spektrotemporalen Eigenschaften stärker variieren und daher mehr alternative Segmentierungshinweise bieten als Sprachreize. Unterschiedliche Ergebnisse in Priming-Untersuchungen mit Umweltgeräuschen und Sprache (Chen & Spence, 2011) sowie eine Verarbeitung an unterschiedlichen Orten im zentralen Nervensystem (Wolfe et al., 2009) weisen zudem darauf hin, dass die Mechanismen der Verarbeitung von räumlichen Hinweisen bei diesen beiden akustischen Kategorien unterschiedlich ausfallen können und möglicherweise auch zu anderen Ergebnissen in akustischen Suchaufgaben führen.

### **5.3. Vorhersagbarkeit der Position einer Schallquelle**

Eine von drei verbleibenden Untersuchungen, die keine Sprache als Zielreiz verwendeten, ist die Untersuchung von Best, Ozmeral, Gallun, Sen und Shinn-Cunningham (2005), in der Personen Gesang von Zebrafinken wiedererkennen sollten. Da die Zielreize jedoch stets von vorne präsentiert wurden und auch durchgehend andere akustische Eigenschaf-

ten als die Distraktoren besaßen, lag auch hier eine von vielen natürlichen Bedingungen abweichende Situation vor, in denen die Positionen von Geräuschen weder fest noch vorab bekannt sein müssen und in denen dieselben Geräusche in manchen Situationen Zielreize oder in anderen Situationen Distraktoren sein können.

Wie auch Kidd et al. (2005) zeigen konnten, kann die Vorhersagbarkeit des Zielreizortes Detektionsaufgaben vereinfachen. Wenn von drei gleichzeitig präsentierten aber räumlich separierten Sätzen einer nur durch ein Hinweiswort statt durch seine räumliche Position angekündigt wurde, konnten die am Ende des Satzes gesprochenen Wörter weniger gut wiedergegeben werden, als wenn die Position des Zielsatzes vor dem Durchgang bekannt war. Sach, Hill und Bailey (2000) ließen ihre Teilnehmer Diskriminationsaufgaben mit synthetischen Geräuschen durchführen und manipulierten dabei die räumliche Position dieser Geräusche allein mit Hilfe interauraler Zeitdifferenzen. Auch hier zeigte sich, dass die Leistungen der Teilnehmer besser waren, wenn die Zielgeräusche an erwarteten Orten auftraten. Zu einem ähnlichen Ergebnis kommen Arbogast und Kidd (2000), die Maskierer mit vor allem informatorischer Maskierung verwendeten. Ein Zielgeräusch konnte genauer und schneller identifiziert werden, wenn es von der erwarteten statt einer anderen Position gespielt wurde.

Folglich ist eine räumlich ausgerichtete Aufmerksamkeit ebenso in anderen Untersuchungen mit nicht oder nur gelegentlich variierenden Zielpositionen zu erwarten – und darüber hinaus möglicherweise auch ein Einfluss dieser Aufmerksamkeitsausrichtung auf den Effekt der räumlichen Separation. Wird der Zielort nämlich nicht vorher bekannt gegeben und deswegen in manchen Durchgängen ein Ort ohne Ziel beachtet, so muss bei separierten Schallquellen unter Umständen nicht nur die Aufmerksamkeit neu ausgerichtet werden, sondern unter Umständen auch ein Distraktor im Aufmerksamkeitsfokus ignoriert werden, während sich bei nicht räumlich separierten Schallquellen nie allein ein Distraktor im Aufmerksamkeitsfokus befinden kann. Bei vorher bekanntem Zielort gibt es diesen Störeinfluss des beachteten Distraktors in der separierten Bedingung im Vergleich zur nicht separierten Bedingung jedoch nicht, weshalb die Vorhersagbarkeit des Zielortes eher

zu einem größeren Effekt der räumlichen Separation führen sollte, als wenn der Zielort nicht vorher bekannt ist.<sup>6</sup>

#### 5.4. Die Untersuchungen von Eramudugolla et al. (2008)

Abgesehen von den im Folgenden geschilderten Untersuchungen von Eramudugolla et al. (2008) scheint es keine Untersuchungen zu geben, in denen die drei oben genannten Bedingungen berücksichtigt wurden, die in vielen realistischen Hörsituationen als wahrscheinlich angenommen werden können: Erstens sind Menschen häufig von vielen Geräuschen gleichzeitig umgeben, zweitens handelt es sich oft auch um nicht-sprachliche Reize und drittens kann die Position eines zu beachtenden Reizes vielmals nicht vorhergesagt werden. Diese drei Bedingungen sollten jedoch eher zu einer Reduktion des Separationseffektes führen. Im Weiteren werden nun diese Untersuchungen erläutert.

Eramudugolla et al. (2008) waren der Auffassung, dass sich Ergebnisse von Studien mit nur einem Distraktor nicht auf Hörsituationen mit mehreren Distraktoren verallgemeinern lassen. Dies liegt nach Meinung der Autoren vor allem an der Beanspruchung der Aufmerksamkeit und der Anfälligkeit für Ablenkung, die in vielen Studien nicht berücksichtigt werden. Fraglich ist auch, ob räumliche Hinweise natürlicherweise auch in komplexen Hörsituationen genutzt werden.

Die Autoren führten daher eine Untersuchung durch, in der 28 Teilnehmer eine Aufgabe zur akustischen Suche mit mehreren Distraktoren bearbeiteten. Vier, sechs oder acht Geräusche wurden virtuell über Kopfhörer entweder von einem gemeinsamen Ort oder von verschiedenen Orten gleichzeitig präsentiert. Die Geräusche konnten von 0 bis 350° in der horizontalen Ebene mit einem Mindestabstand von 20° auftreten und umfassten eine weibliche Nachrichtensprecherin in Hindi, einen männlichen Kommentator in Englisch, gregorianischen Chorgesang, ein Klaviersolo, ein Cellosolo, ein Trompetensolo, synthetisches Schlagzeug, Vogelgezwitscher, Hühnergegacker, eine Polizeisirene und eine Alarmglocke. In jedem Durchgang hatten die Teilnehmer die Aufgabe, die An- bzw. Abwesenheit

---

<sup>6</sup> Zu einer ähnlichen Vorhersage kommt man, wenn man eine gleichzeitige Beachtung mehrerer Orte bei unbekanntem Zielort annimmt. Auch dann erhalten die Distraktoren in der räumlich separierten Bedingung mehr Aufmerksamkeit, als dies in der gleichen Bedingung bei bekanntem Zielort der Fall ist.

eines zufällig ausgewählten Zielgeräusches per Tastendruck anzugeben. Zwar zeigte sich ein Separationseffekt – die Sensitivität für räumlich separierte Geräusche war höher als für nicht separierte – jedoch waren die Reaktionszeiten in der räumlich separierten Bedingung bei abwesendem Zielreiz gleichzeitig auch langsamer. Aufgrund dieses Geschwindigkeits-Genauigkeits-Austauschs konnte ein eindeutiger Verarbeitungsvorteil zugunsten der räumlichen Separation nicht belegt werden.

In einer Folgestudie wurden 17 Teilnehmern dieselben akustischen Szenarios präsentiert, jedoch in unterschiedlicher Länge (250, 500, 1000, 1500 oder 3000 ms). Im direkten Anschluss wurde für 2000 ms weißes Rauschen eingespielt. Dieses Rauschen sollte die Teilnehmer dabei stören, auf ihr echoisches Gedächtnis zuzugreifen. Außerdem wurden sie instruiert, so genau wie möglich und beim Einsetzen des Rauschens zu reagieren. Es wurden jedoch keine Reaktionszeiten erhoben.

Auch in dieser Untersuchung zeigte sich eine höhere Sensitivität bei räumlich separierten Geräuschen, jedoch war der Effekt räumlicher Separation nicht für alle Präsentationsdauern gleich stark ausgeprägt. Post-hoc-Tests lieferten einen Separationseffekt in der Bedingung mit 250 ms Präsentationsdauer ( $p = .01$ ), jedoch keinen Separationseffekt in den Bedingungen mit 500, 1000 und 1500 ms Präsentationsdauer (alle  $p > .10$ ), während das Signifikanzniveau bei 3000 ms nur knapp verfehlt wurde ( $p = .07$ ). Die Autoren vermuteten, dass der Vorteil der räumlichen Separation bei kurzen Präsentationsdauern auf eine verbesserte Stromtrennung zurückgeht. Ein serielles räumliches Absuchen der auditorischen Szene – im Sinne eines sich von Position zu Position verlagernden Aufmerksamkeitsfokus – sei bei nur 250 ms dauernden Stimuli nämlich unwahrscheinlich.

## 5.5. Offene Fragen und Motivation der eigenen Untersuchungen

Wie in jeder Untersuchung bleiben auch bei Eramudugolla et al. (2008) Fragen offen: Angesichts von nur 17 Teilnehmern und Separationseffekten in den Bedingungen 500, 1500 und 3000 ms zumindest auf deskriptiver Ebene kann ein Separationseffekt bei längeren Präsentationsdauern nicht generell ausgeschlossen werden. So erreicht z.B. bei der gegebenen Stichprobengröße von  $N = 17$  ein gerichteter  $t$ -Test für einen mittelgroßen Separat-

ionseffekt nach Cohen (1988) von  $d_z = 0.5$  und einem  $\alpha$  von .05 lediglich eine Teststärke von knapp .63. Mit einer derartig kleinen Stichprobe können lediglich große Effekte von  $d_z = 0.83$  (oder größer) mit einer Fehlerwahrscheinlichkeit 1. oder 2. Art von .05 gefunden werden, wie eine Sensitivitätsanalyse zeigt. Möglicherweise wäre also der Separationseffekt in der Studie von Eramudugolla et al. (2008) bei längeren Präsentationsdauern in einer Untersuchung mit größerer Teilnehmerzahl signifikant geworden.

Etwas unklar bleibt auch, weshalb ein bei kurzen Präsentationsdauern auftretender Separationsvorteil bei längeren Präsentationsdauern verloren gehen sollte. Eramudugolla et al. (2008) spekulieren zwar, dass es möglicherweise Prozesse gäbe, die bei kurzen Präsentationsdauern wahrscheinlich nicht auftreten (serielles räumliches Absuchen) – erklären jedoch nicht, inwiefern diese einem Separationseffekt bei längeren Präsentationsdauern entgegenwirken. Auch schweigen sie zu der Frage, ob es Prozesse gäbe, die umgekehrt nur bei kurzen Präsentationsdauern vorkommen. Wenn diese speziell von der räumlichen Separation profitierten, könnte man möglicherweise dadurch ein Ausbleiben des Separationseffekts bei langen Präsentationsdauern erklären. Es müsste aber auch erklärt werden, warum die Prozesse, die bei kurzen Präsentationsdauern zu einem Separationseffekt führen, am Anfang langer Präsentationsdauern nicht stattfinden bzw. wodurch ein eventueller Separationsvorteil am Anfang einer langen Präsentation im weiteren zeitlichen Verlauf verloren geht.

Die Unsicherheit über die Befunde von Eramudugolla et al. (2008) wird verstärkt durch eine Untersuchung von Arrabito (2006), in der keine Interaktion zwischen Präsentationsdauer und räumlicher Separation der Schallquellen gefunden werden konnte. Personen hörten über einen Kopfhörer eine Aufnahme eines Torpedos – und in einer anderen Sitzung einen Sinuston – zusammen mit diotischem rosa Rauschen (d.h. Rauschen, dessen Energiedichte mit steigenden Frequenzen um 3 dB pro Oktave abnimmt) als Distraktor, wobei die Zielreize entweder ebenso diotisch oder mit Hilfe von kopfbezogenen Übertragungsfunktionen aus einem Winkel von 30°, 60° oder 90° dargeboten wurden. Die Dauer der Präsentation konnte dabei die Werte 12.5, 25, 50, 100, 200, 500 oder 1000 ms annehmen. Die Detektionsleistung war für einen Sinuston als Zielreiz besser als für eine Aufnahme

eines Torpedos, für um 60 oder 90° separierte Darbietung des Zielreizes besser als für 30° separierte und für 1000 ms Präsentationsdauer besser als für 12.5 ms. Es konnte jedoch – wie bereits erwähnt – keine Interaktion zwischen Separation und Präsentationsdauer und auch keine Interaktion höheren Grades gefunden werden. Der Separationseffekt zeigte sich hier also im Unterschied zu Eramudugolla et al. bei allen Präsentationsdauern etwa gleichermaßen.

Für die Interpretation der Ergebnisse von Eramudugolla et al. (2008) wäre auch wichtig zu wissen, ob es auch in der zweiten Untersuchung einen Geschwindigkeits-Genauigkeitsaustausch gegeben hat. Ohne die Reaktionszeiten betrachtet zu haben, lässt sich nicht sagen, wie gut die Teilnehmer sich an die Instruktionen gehalten haben und tatsächlich immer gleichermaßen genau und ohne Eile reagiert haben. Um dies zu unterstützen, verwendeten die Autoren ein zweisekündiges weißes Rauschen, das das Gemisch aus Geräuschen rückwärts maskieren sollte. Obwohl hier der Begriff der rückwärtsgerichteten Maskierung (*backward masking*; bei anderen Autoren auch *premasking*) verwendet wurde, ist er vermutlich nicht im akustischen Sinne gemeint. Dann nämlich würde die Wahrnehmung eines Geräusches durch ein kurz darauf folgendes lauterer Geräusch (das weiße Rauschen) maskiert werden (also nicht gehört werden können), was jedoch nur – abhängig von der Lautstärke des Maskierers – für einen Zeitraum bis etwa 20 ms unmittelbar vor dem maskierendem Geräusch möglich ist (Fastl & Zwicker, 2007). Eramudugolla et al. meinen vermutlich eine Beeinträchtigung des akustischen Gedächtnisses durch die Präsentation von Rauschen. Dies ist jedoch in der Literatur nicht belegt. In einer Untersuchung zum Stimulus-Suffix-Effekt präsentierten Morton, Crowder und Prussin (1971) eine Reihe von Ziffern, die die Teilnehmer behalten sollten. Folgte auf die Reihe der Ziffern das Suffix *zero*, so konnten mehr Fehler (vor allem in den letzten Ziffern) bei der Wiedergabe der Ziffern als ohne ein Suffix beobachtet werden. Ein Suffix in Form eines weißen Rauschens – egal ob laut oder leise – beeinträchtigte das Gedächtnis hingegen kaum. Auf eine Betrachtung von Reaktionszeiten kann also auch mit weißem Rauschen als vermeintliche Rückwärtsmaskierung nicht verzichtet werden, wenn ein Geschwindigkeits-Genauigkeitsaustausch ausgeschlossen werden soll.

In der Literatur findet sich also bislang keine eindeutige Evidenz dafür, dass der Effekt räumlicher Separation auch in komplexen, alltagsnahen akustischen Situationen auftreten kann, die sich durch mehrere Schallquellen gleichzeitig, das Auftreten von Umweltgeräuschen und Unsicherheit in räumlichen Positionen sowie spektrotemporalen Eigenschaften auszeichnen. Um diese Lücke zu schließen, wird eine Replikation der Untersuchung von Eramudugolla et al. (2008) angestrebt, bei der jedoch ein eventueller Geschwindigkeits-Genauigkeitsaustausch durch Erheben von Reaktionszeiten kontrolliert und eine höhere Teststärke durch höhere Teilnehmerzahl erreicht werden soll. Ferner sollen Einflussfaktoren auf den Separationseffekt in Situationen ohne Relevanz der Lokation näher untersucht werden. Dies sollte Aufschluss geben über akustische Suchprozesse, die – im Gegensatz zu Suchprozessen in der visuellen Modalität (siehe auch Eckstein, 2011) – bisher nur wenig untersucht worden sind.

## 6. Eigene Untersuchungen

In diesem Kapitel werden vier eigene Untersuchungen zum Effekt der räumlichen Separation vorgestellt. In der ersten Untersuchung wurde eine Suchaufgabe in einem Raum mit Reflexionen durchgeführt, während die zweite Untersuchung in einem reflexionsarmen Raum stattfand. Anschließend wurde der Zielreiz in der dritten Untersuchung durch einen akustischen statt visuellen Hinweis gekennzeichnet. In der vierten Untersuchung wurde der Einfluss von Ähnlichkeit zwischen Ziel- und Distraktorreiz auf den Separationseffekt untersucht.

### 6.1. Untersuchung 1 - Separationseffekt im Raum mit Reflexionen

In einer ersten Untersuchung wird geprüft, ob sich die Ergebnisse von Eramudugolla et al. (2008) replizieren lassen. Gleichzeitig wird dabei kontrolliert, ob ein Geschwindigkeits-Genauigkeitsaustausch einen gegebenenfalls auftretenden Separationsvorteil erklären kann. Daher wurde nicht nur die Sensitivität, sondern auch Reaktionszeiten als abhängige Variablen erfasst. Um die Hörsituation möglichst realistisch zu gestalten, wurden alle in der Untersuchung dargebotenen Reize mit einem Messmikrophon in einem reflexionsfreien Raum aufgenommen, so dass Frequenzen zwischen 100 Hz und 20 kHz näherungsweise mit unverfälschter Amplitude<sup>7</sup> und ohne Reflexionen des Aufnahmeraums wiedergegeben werden konnten. Die Untersuchung selbst fand in einem akustisch nicht behandelten Raum statt, und die räumliche Separation der Schallquellen wurde nicht virtuell simuliert, sondern mit Hilfe von mehreren Lautsprechern realisiert. Außerdem wurde auf Basis einer A-priori-Teststärkeanalyse eine angemessen hohe Zahl an Teilnehmern untersucht, um bei gegebenenfalls ausbleibenden Separationseffekten (vor allem bei längeren Präsentationsdauern) nicht mit dem Problem mangelnder Teststärke konfrontiert zu sein.

Für die Bedingungen mit räumlich separierten Geräuschen werden eine höhere Sensitivität oder kürzere Reaktionszeiten oder beides erwartet – im Vergleich zu den Bedingungen mit räumlich nicht separierten Geräuschen. Es wird – konträr zu Eramudugolla et al.

---

<sup>7</sup> Im Gegensatz dazu verwendeten Eramudugolla et al. (2008) Aufnahmen, die teilweise nur Frequenzen bis etwa 6 oder 11 kHz berücksichtigten.

(2008) – mangels stichhaltiger Begründung keine Interaktion zwischen der räumlichen Separation und der Präsentationsdauer vermutet, sondern etwa gleich große Separationsvorteile für alle Präsentationsdauern.

### 6.1.1. Methode

#### 6.1.1.1. Stichprobe

Von den 109 Personen, die an der Untersuchung teilnahmen, wurden 17 von der Auswertung ausgeschlossen. Bei acht Personen trat während der Durchführung unvermeidbarer Lärm außerhalb des Laborraums auf, bei einer Person stürzte die Experimentalsoftware vor dem Ende der Untersuchung ab, eine Person veränderte während der Untersuchung ihre Sitzposition, bei zwei Personen wurden Lautstärkeschwankungen durch Wackelkontakte der Stecker festgestellt und fünf zeigten in mindestens 30 Durchgängen Reaktionszeiten, die als für eine normale Reizverarbeitung und Reaktion zu schnell (weniger als 200 ms) oder zu langsam (mehr als 5000 ms) definiert wurden<sup>8</sup>. Bei diesen zu frühen oder zu späten Reaktionen handelt es sich wahrscheinlich um ein Verhalten, das nicht als ein instruktionsgemäßes Bearbeiten der Aufgabe interpretiert werden kann. Die hier beschriebenen Ergebnisse beziehen sich auf die Daten der verbleibenden 92 Personen, von denen 62 weiblich waren. Das Alter lag zwischen 19 und 39 Jahren mit einem Durchschnitt von etwa 26 Jahren. Die Teilnehmer wurden wahlweise mit Teilnehmerstunden oder zehn Euro entlohnt.

#### 6.1.1.2. Material

Als Stimuli dienten zufällige Kombinationen von vier, sechs oder acht aus elf Geräuschen. Eine Sprecherin (persische Sprache), ein Sprecher (deutsche Sprache), ein Sänger (französische Sprache), eine Klarinette, eine Mundharmonika, lautes Pfeifen, eine Triangel, eine Trommel mit Snareteppich, eine Fahrradhupe, eine Rassel und ein bellender Hund wurden mit einem Brüel & Kjær Schallanalysator Modell 2250 in einem reflexionsarmen Raum aufgezeichnet, wobei Frequenzen bis mindestens 20 kHz berücksichtigt wurden. Für jedes

---

<sup>8</sup> Das gleiche Kriterium hinsichtlich Anzahl an Antizipationen oder verspäteten Reaktionen wurde in den folgenden zwei Untersuchungen (mit ebenfalls 960 Gesamtdurchgängen) verwendet.

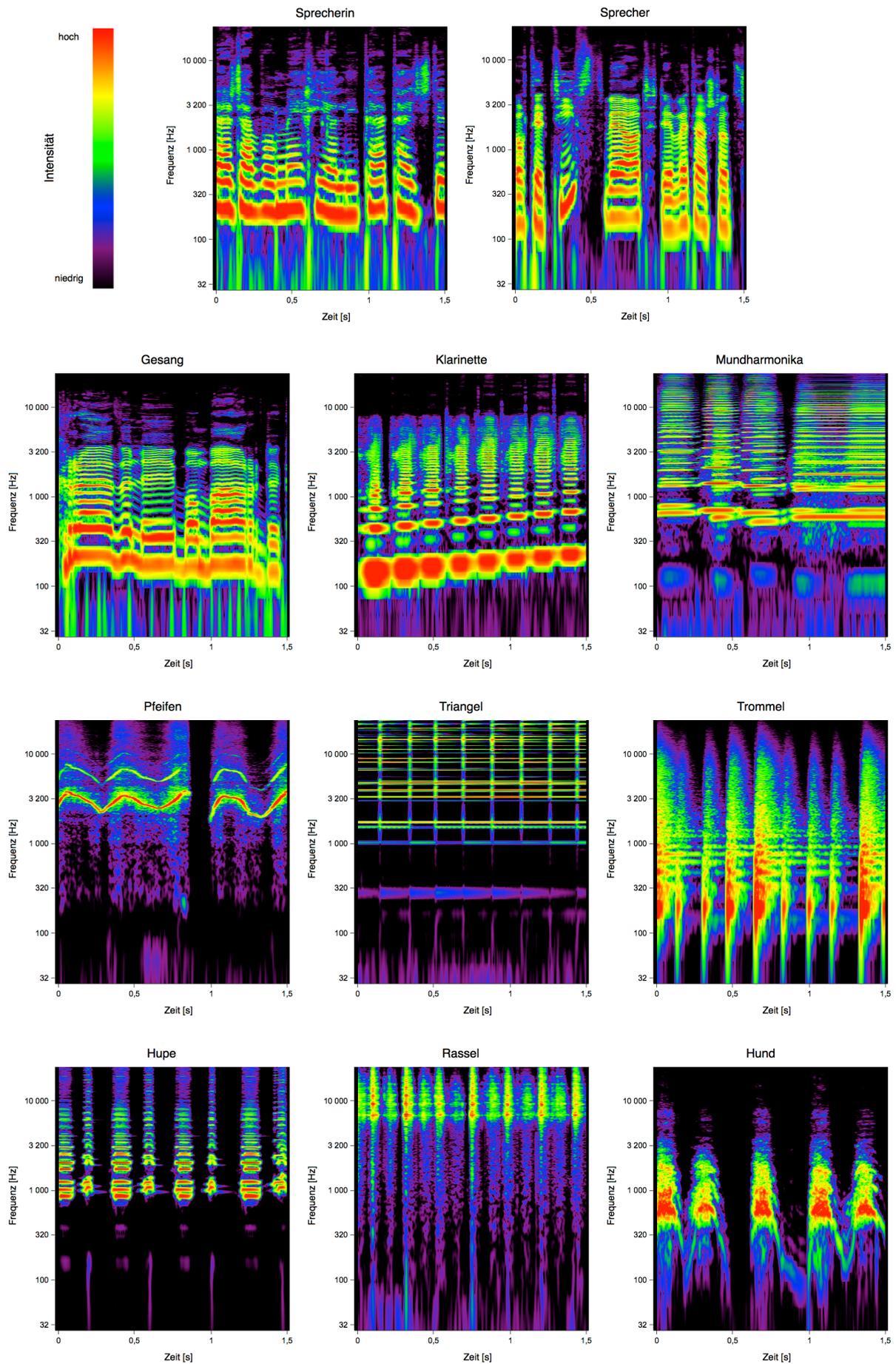


Abbildung 2: Spektrogramme der elf Geräusche. Die x- und y-Achsen zeigen die Zeit und die Frequenz, für die die Intensität farblich kodiert ist.

Geräusch wurde ein 1500 ms langer Abschnitt erstellt, von dem je nach Bedingung des Faktors „Dauer“ die ersten 250, 500, 1000 ms oder der gesamte Abschnitt von 1500 ms präsentiert wurden. Alle Geräusche wurden mit der Software Amadeus Pro v1.4 anhand der maximalen RMS-Leistung<sup>9</sup> als Referenzwert normalisiert. In der Bedingung „zusammen“ wurden alle präsentierten Geräusche aus einem zufällig ausgewählten Lautsprecher gespielt, während in der Bedingung „getrennt“ jedes präsentierte Geräusch einem anderen zufällig ausgewählten Lautsprecher zugeordnet wurde. Vier gleichzeitig präsentierte Geräusche erreichten eine Gesamtlautstärke von etwa 57 dB (A), sechs Geräusche etwa 58,5 dB (A) und acht etwa 60 dB (A). Für diese Messung wurde das Messmikrofon in der Mitte des Lautsprecherkreises nach oben gerichtet positioniert, wobei sich Mikrofon und Lautsprecher etwa in durchschnittlicher Ohrhöhe der sitzenden Teilnehmer befanden. Für die Lautstärkemessung machte es keinen Unterschied, ob die Geräusche alle über einen oder alle über verschiedene Lautsprecher präsentiert wurden. Abbildung 2 zeigt Spektrogramme der Geräusche in 1500 ms Länge.

Die Untersuchung fand in einem 3,2 x 2 m großem Raum mit 3 m Höhe statt, der zusätzlich noch eine Nische von 1,5 x 1 m aufwies. Die Nachhallzeit  $T_{60}$  des Raumes betrug im Frequenzbereich ab 1 kHz etwa 400 ms. Für niedrige Frequenzen nahm die Nachhallzeit zu und lag für 100 Hz etwa bei 700 ms. Vor den Teilnehmern, in einem Abstand von einem Meter, war ein Bildschirm positioniert, auf dem alle visuelle Information präsentiert wurde. Acht aktive Lautsprecher von Typ Yamaha MSP3 wurden im Kreis um den Sitzplatz der Teilnehmer aufgebaut, so dass zwei benachbarte Lautsprecher stets einen Winkel von 45° einschlossen. Um Vorne-hinten-Konfusionen unwahrscheinlicher zu machen, wurde die Anordnung im Vergleich zu einer symmetrischen Anordnung und zum Bildschirm um 11,25° nach links (von oben betrachtet) gedreht, so dass jeder Lautsprecher zu eventuellen Phantom-Lautsprechern (die nach vorne oder hinten gespiegelt sind) eine Winkelweite von 22,5° besaß. Der Abstand der Lautsprecher zum Mittelpunkt betrug 85 cm, und die Höhe ihres akustischen Zentrums (d.h. die Mitte zwischen den Mittelpunkten von Hoch- und Tieftöner) lag bei 120 cm. Abbildung 3 zeigt eine Aufsicht der Versuchsanordnung.

---

<sup>9</sup> Für die Option „Maximale RMS-Leistung“ berechnet Amadeus durchschnittliche RMS-Pegel für 100ms-Abschnitte und wählt den größten als Referenz.

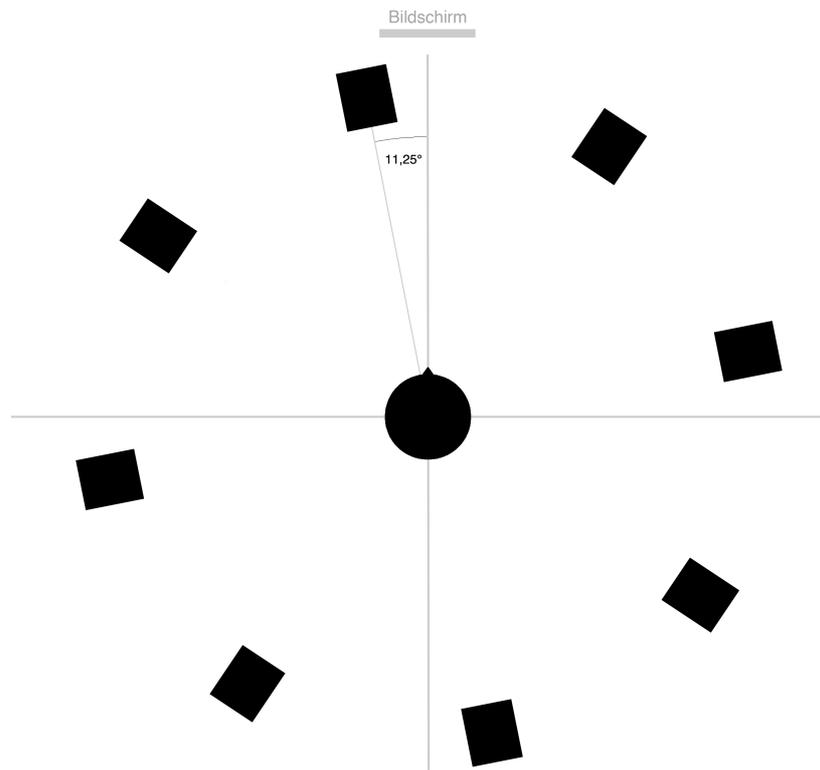


Abbildung 3: Teilnehmerposition und Anordnung der Lautsprecher und des Bildschirms.

Alle visuellen und akustischen Stimuli wurden mit Hilfe eines Rechners mit Intel-Core-i7-860-Prozessor, der Experimentalsoftware Neurobehavioral Systems Presentation Version 14.3, einer 7:1-Soundkarte vom Typ Creative Sound Blaster X-Fi XtremeGamer und einem Monitor EIZO FlexScan S1961 dargeboten. Reaktionszeiten wurden mittels einer USB-Full-Speed-Tastatur von Empirisoft mit 10 Tasten gemessen, die Reaktionen in einer Auflösung von 1 ms erfasst.

#### 6.1.1.3. Versuchsdurchführung

Während der Instruktion hörten sich die Teilnehmer alle in der Untersuchung vorkommenden Geräusche in allen Darbietungslängen mindestens ein Mal an und wurden gebeten, in der Untersuchung vor allem so korrekt, aber auch so schnell wie möglich zu reagieren. Sie absolvierten 50 Übungsdurchgänge und 960 Experimentaldurchgänge im Einzelversuch.

Am Beginn eines Durchgangs wurde für 1 s ein Fixationskreuz und anschließend der Name des zu suchenden Geräuschs (Frau, Sprecher, Gesang, Klarinette, Mundharmonika,

Pfeifen, Triangel, Trommel, Hupe, Rassel oder Hund) in der Mitte des Monitors präsentiert. Die Geräusche wurden 750 ms nach Einblenden des Namens dargeboten. Nach der Präsentationsdauer, die 250 ms, 500 ms, 1000 ms oder 1500 ms lang sein konnte, wurde 3000 ms lang unkorreliertes weisses Rauschen aus allen acht Lautsprechern gespielt. Richtige, falsche oder zu späte Reaktionen wurden 1000 ms lang visuell als solche rückgemeldet. Zu frühe Reaktionen wurden als falsch rückgemeldet. Richtige Reaktionen waren 200 ms bis 5000 ms nach dem Beginn der Geräusche möglich.

Für jeden Durchgang wurde per Zufall ein Zielreiz und drei, fünf oder sieben Distraktoren ausgewählt, wobei jedes Geräusch nur einmal auftreten konnte. Außerdem wurden die Lautsprecherpositionen zufällig ausgewählt, wobei jedes Geräusch – außer in den Bedingungen ohne Separation – von einem anderen Lautsprecher präsentiert wurde. Die Separation (d.h. die räumlich gemeinsame oder getrennte Präsentation der Geräusche) und die Anzahl der Geräusche (vier, sechs oder acht Geräusche) wurde blockweise variiert, so dass sich sechs Experimentalblöcke ergaben, deren Reihenfolge randomisiert wurde. Jede Präsentationsdauer trat in einem Block gleich häufig auf. Das zu suchende Geräusch war in der Hälfte der Durchgänge anwesend und in der anderen Hälfte abwesend. Innerhalb der Blöcke wurde die Reihenfolge der Durchgänge vollständig randomisiert. Bei Anwesenheit des Geräusches sollte jeder zweite Teilnehmer die rechte Taste und bei Abwesenheit die linke Taste der Tastatur drücken; bei den verbleibenden Teilnehmern war die Tastenzuordnung umgekehrt. Die Tastatur lag auf dem Schoß der Teilnehmer, die ihre Finger über den Tasten halten sollten. Nach jeweils 20 Durchgängen konnten die Personen eine Pause machen und die Untersuchung selbstständig fortsetzen. Pro Person dauerte die Untersuchung etwa 90 Minuten.

#### 6.1.1.4. *Versuchsplan*

Der Untersuchung lag ein  $2 \times 2 \times 3 \times 4$ -faktorieller Versuchsplan mit den messwiederholten Faktoren Separation, Anwesenheit (des zu suchenden Geräusches), Anzahl und Dauer zugrunde. Der Faktor Separation besaß die Ausprägungen „zusammen“ (alle Geräusche wurden aus einem Lautsprecher gespielt) und „getrennt“ (alle Geräusche wurden aus verschiedenen Lautsprechern gespielt), der Faktor Anwesenheit die Ausprägungen „anwe-

send“ und „abwesend“ (bezogen auf das zu suchende Geräusch), der Faktor Anzahl (der gleichzeitig präsentierten Geräusche) die Ausprägungen „4 Geräusche“, „6 Geräusche“ und „8 Geräusche“ und der Faktor Dauer (der Geräusche) die Ausprägungen „250 ms“, „500 ms“, „1000 ms“ und „1500 ms“.

Die Sensitivität  $P_r$  (Snodgrass & Corwin, 1988) und die Reaktionszeit bei korrekten Antworten waren die primären abhängige Variablen.  $P_r$  ist definiert als die Differenz aus Treffer-Rate und der Rate falscher Alarme ( $P_r = Hits - FA$ ).

Um die Teststärke für einen Interaktionseffekt zwischen Separation und Dauer zu berechnen, wurden – in dieser wie auch in den folgenden beiden Untersuchungen – Differenzen zwischen den Ausprägungen des Faktors Separation (d.h. zusammen versus getrennt) für alle Stufen des Faktors Präsentationsdauer gebildet und die Teststärkenberechnung für eine einfaktorielle Varianzanalyse mit Messwiederholung bezüglich dieser Differenzvariablen durchgeführt. Eine A-priori-Analyse mit einer angenommenen Populationskorrelation von  $\rho = .5$  zeigte, dass eine Stichprobengröße von 55 Personen erforderlich ist, um einen Effekt (d.h. einen Unterschied in der Größe des Separierungseffekts für die verschiedenen Präsentationsdauern) von  $f = 0.2$  mit  $\alpha = \beta = .05$  nachweisen zu können. Eine Sensitivitätsanalyse mit der tatsächlich verfügbaren Stichprobengröße von 92 Teilnehmern und ansonsten gleichen Kennwerten ergab, dass ein Effekt von  $f = 0.154$  oder größer gefunden werden konnte. Nach Cohen (1988) kann ein Effekt von  $f = 0.25$  als mittelgroß und ein Effekt von  $f = 0.1$  als klein bezeichnet werden.

Für alle in dieser Arbeit berichteten Untersuchungen wurde ein multivariater Ansatz zur Auswertung messwiederholter Versuchspläne gewählt (MANOVA). Als Effektstärkemaß wird das partielle  $\eta^2$  berichtet. Post-Hoc-Tests auf einen Separationseffekt wurden gerichtet und unter Verwendung der Bonferoni-Holm-Prozedur für multiples Testen gerechnet (Holm, 1979).

### 6.1.2. Ergebnisse

In den Ergebnisteilen dieser Arbeit werden ausschließlich der Haupteffekt Separation, alle signifikanten Haupteffekte und alle signifikanten Interaktionen erwähnt. Alle Kennwerte

der jeweiligen statistischen Analyse werden jedoch in einer tabellarischen Übersicht dargestellt.

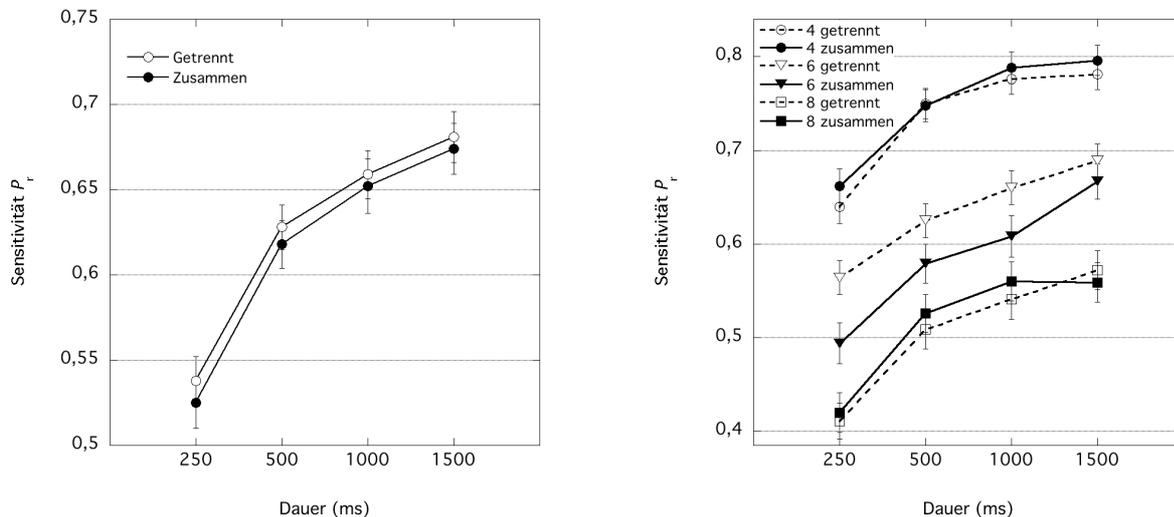


Abbildung 4: Die Sensitivität  $P_r$  für Untersuchung 1 in Abhängigkeit von der Separation und Präsentationsdauer (links) sowie in Abhängigkeit von der Separation, der Anzahl und der Präsentationsdauer der gleichzeitig präsentierten Geräusche (rechts). Die Fehlerbalken zeigen Standardfehler.

Die abhängige Variable Sensitivität nahm deskriptiv mit der Präsentationsdauer zu und war bei räumlich getrennter Darbietung der Geräusche tendenziell größer als bei Darbietung der Geräusche zusammen von einem Ort (siehe Abbildung 4, links). Außerdem war die Sensitivität deskriptiv größer, je geringer die Anzahl der gleichzeitig präsentierten Geräusche war, wobei ein Vorteil räumlich separierter Schallquellen nur in der Bedingung mit sechs Geräuschen gleichzeitig zu erkennen ist (siehe Abbildung 4, rechts). Eine  $2 \times 3 \times 4$ -MANOVA mit den messwiederholten Faktoren Separation, Anzahl und Dauer zeigte keinen Haupteffekt der Separation, aber einen Haupteffekt der Anzahl und einen Haupteffekt der Dauer. Die einzige signifikante Interaktion war die zwischen Separation und Anzahl. Die statistischen Kennwerte sind in Tabelle 2 aufgeführt.

Gerichtete Post-hoc-Tests auf adjustiertem Alphaniveau ( $\alpha = .05/3 \approx .017$ ) lieferten nur in der Bedingung mit sechs Geräuschen einen signifikanten Separationsvorteil,  $F(1,91) = 8.138$ ,  $p = .003$ ,  $\eta^2 = .082$ , und nicht in den Bedingungen mit vier,  $F(1,91) = 0.930$ ,  $p = .169$ ,  $\eta^2 = .010$ , oder acht Geräuschen,  $F(1,91) = 0.248$ ,  $p = .310$ ,  $\eta^2 = .003$ . In den letztgenannten

Tabelle 2

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Sensitivität für Untersuchung 1. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		$F$	$p$	$\eta^2$
	Zähler	Nenner			
Separation	1	91	1.105	.296	.012
<b>Anzahl</b>	<b>2</b>	<b>90</b>	<b>224.882</b>	<b>&lt; .001</b>	<b>.833</b>
<b>Dauer</b>	<b>3</b>	<b>89</b>	<b>107.839</b>	<b>&lt; .001</b>	<b>.784</b>
<b>Separation x Anzahl</b>	<b>2</b>	<b>90</b>	<b>4.760</b>	<b>.011</b>	<b>.096</b>
Separation x Dauer	3	89	0.080	.970	.003
Anzahl x Dauer	6	86	1.216	.306	.078
Separation x Anzahl x Dauer	6	86	1.032	.410	.067

Bedingungen trat je nach Präsentationsdauer oft sogar ein tendenzieller Separationsnachteil auf.

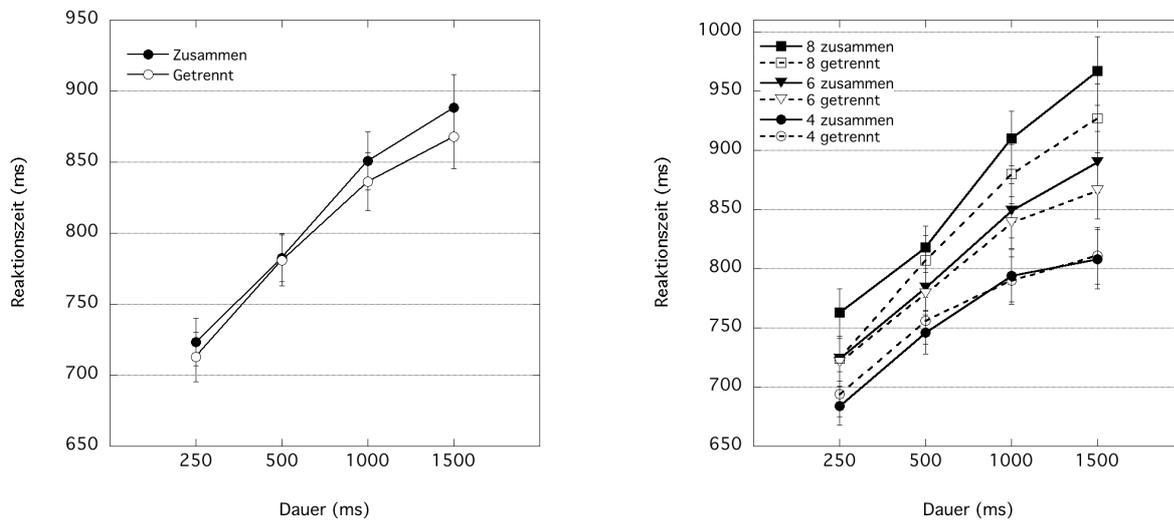


Abbildung 5: Die Reaktionszeit für Untersuchung 1 in Abhängigkeit von der Separation und der Präsentationsdauer (links) sowie in Abhängigkeit der Separation, der Anzahl und der Präsentationsdauer der gleichzeitig präsentierten Geräusche (rechts). Die Fehlerbalken zeigen Standardfehler.

Die abhängige Variable Reaktionszeit nahm deskriptiv mit der Präsentationsdauer zu und war bei räumlich getrennter Darbietung der Geräusche tendenziell kleiner als bei Darbietung der Geräusche zusammen von einem Ort (siehe Abbildung 5, links). Außerdem waren die Reaktionszeiten kürzer, je geringer die Anzahl der gleichzeitig präsentierten Geräusche war (siehe Abbildung 5, rechts).

Tabelle 3

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Reaktionszeit für Untersuchung 1. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		<i>F</i>	<i>p</i>	$\eta^2$
	Zähler	Nenner			
Separation	1	91	2.149	.146	.023
<b>Anzahl</b>	<b>2</b>	<b>90</b>	<b>41.193</b>	<b>&lt; .001</b>	<b>.478</b>
<b>Dauer</b>	<b>3</b>	<b>89</b>	<b>75.346</b>	<b>&lt; .001</b>	<b>.717</b>
<b>Anwesenheit</b>	<b>1</b>	<b>91</b>	<b>334.301</b>	<b>&lt; .001</b>	<b>.786</b>
Separation*Anzahl	2	90	1.222	.299	.026
Separation*Dauer	3	89	2.014	.118	.064
Separation*Anwesenheit	1	91	0.740	.392	.008
<b>Anzahl*Dauer</b>	<b>6</b>	<b>86</b>	<b>9.454</b>	<b>&lt; .001</b>	<b>.397</b>
Anzahl*Anwesenheit	2	90	2.938	.058	.061
<b>Dauer*Anwesenheit</b>	<b>3</b>	<b>89</b>	<b>36.408</b>	<b>&lt; .001</b>	<b>.551</b>
Separation*Anzahl*Dauer	6	86	0.781	.587	.052
Separation*Anzahl*Anwesenheit	2	90	0.280	.756	.006
Separation*Dauer*Anwesenheit	3	89	1.959	.126	.062
Anzahl*Dauer*Anwesenheit	6	86	1.546	.173	.097
Separation*Anzahl*Dauer*Anwesenheit	6	86	1.753	.118	.109

Eine 2 x 2 x 3 x 4-MANOVA mit den messwiederholten Faktoren Anwesenheit, Separation, Anzahl und Dauer zeigte einen Haupteffekt der Anwesenheit, keinen Haupteffekt der Separation, einen Haupteffekt der Anzahl und einen Haupteffekt der Dauer. Außerdem

wurde eine Interaktion zwischen Anzahl und Dauer gefunden. Gemessen an den Tendenzen der Effekte der zugehörigen einzelnen Bedingungen wird die Interpretation der zugehörigen Haupteffekte jedoch nicht eingeschränkt<sup>10</sup> (für die konkrete Ausprägung dieser nur nachrangig interessierenden Interaktion sei lediglich auf Abbildung 5 rechts verwiesen).

Weiterhin ergab sich eine Interaktion zwischen Dauer und Anwesenheit, die gemessen an den Tendenzen der Effekte der zugehörigen einzelnen Bedingungen ebenfalls keinen Einfluss auf die Interpretation der zugehörigen Haupteffekt hat. Die statistischen Kennwerte zeigt Tabelle 3.

### 6.1.3. Diskussion

In Untersuchung 1 konnte eine erhöhte Sensitivität bei räumlich separierten Geräuschen im Vergleich zu nicht separierten Geräuschen in der Bedingung mit sechs gleichzeitig präsentierten Geräuschen gefunden werden, nicht jedoch bei vier oder acht Geräuschen. In den Reaktionszeiten zeigten sich keine mit der Variable Separation in Zusammenhang stehenden Effekte, sondern lediglich tendenziell schnellere Reaktionszeiten bei separierten Geräuschen. Der Separationseffekt in der Sensitivität war also eher geringer als in der Untersuchung von Eramudugolla et al. (2008) ausgeprägt. Außerdem zeigte sich kein Hinweis auf einen Einfluss der Präsentationsdauer auf den Separationseffekt, während Eramudugolla et al. (2008) einen größeren Effekt bei kurzen Präsentationsdauern berichteten.

Einen insgesamt nicht stark ausgeprägten Separationsunterschied könnte man alleine deswegen erwarten, da interaurale Pegeldifferenzen im Wesentlichen nur dann zu einem Vorteil führen, wenn der Zielreiz sich maximal links oder rechts des Teilnehmers befinden. Da dies bei vier, sechs oder acht zufällig positionierten Schallquellen meist nicht der Fall ist, tragen interaurale Pegeldifferenzen also nur wenig zu einem eventuellen Separationsvorteil bei. Außerdem sind Effekte der Aufmerksamkeitsausrichtung wahrscheinlich nur

---

<sup>10</sup> Interaktionen, an denen die Variable Separation nicht beteiligt ist, werden hier und im Folgenden als nicht einschränkend betrachtet, wenn die Effekte der zugehörigen einzelnen Bedingungen sich zumindest tendenziell und auf mit den Haupteffekten konsistente Weise unterscheiden. Hier nehmen beispielsweise die Reaktionszeiten für jede Anzahl mit längerer Dauer zumindest tendenziell zu und ebenso für jede Dauer mit zunehmender Anzahl. Die Interpretation der Haupteffekte wird daher als nicht eingeschränkt angenommen. Aus Übersichtsgründen wird bei Interaktionen, an denen die Variable Separation nicht beteiligt ist, auf die Darstellung von Post-Hoc-Tests verzichtet.

begrenzt relevant, da die Position des Zielreizes in jedem Durchgang zufällig ausgewählt wurde und somit nicht vorhersagbar war. Ein eventueller Separationsvorteil in dieser Untersuchung geht daher vermutlich hauptsächlich auf interaurale Zeitdifferenzen zurück und sollte deshalb auch nicht so groß ausfallen wie in Situationen, in denen auch eine Nutzung von interauralen Pegeldifferenzen oder einer Aufmerksamkeitsausrichtung möglich ist.

Ein weiterer möglicher Grund für den geringen Separationseffekt ist die Darbietung der Reize über Lautsprecher in einem realen Raum mit Reflexionen. Die Anwesenheit von Reflexionen (oder Nachhall, wobei Nachhall als ein Gemisch aus vielen Reflexionen betrachtet werden kann) kann mehrere Konsequenzen für die Identifikation von akustischen Reizen haben. Beispielsweise muss Sprache, um die gleiche Verständlichkeit zu erreichen, in einer Umgebung mit Nachhall lauter präsentiert werden als in einer Umgebung ohne Nachhall oder auch mit leiserem Nachhall (Bronkhorst & Plomp, 1990). Die Erkennbarkeit der Geräusche war durch den Nachhall des Raumes in Untersuchung 1 also herabgesetzt und die Aufgabenschwierigkeit damit höher, als dies in einer Umgebung ohne Nachhall der Fall gewesen wäre. Ebenso fanden Ruggles und Shinn-Cunningham (2011), dass das Wiedergeben einer Zahlenfolge, die von zwei räumlich um  $15^\circ$  getrennten Distraktor-Zahlenfolgen eingeschlossen war, bei einer längeren Nachhallzeit weniger gut gelang und häufiger Zahlen aus den Distraktorfolgen genannt wurden als bei kurzer Nachhallzeit oder gar keinem Nachhall. Die räumliche Ausrichtung von Aufmerksamkeit scheint also bei Nachhall auch weniger gut zu gelingen.

Nachhall kann auch monaurale Hinweise wie das Signal-Rausch-Verhältnis beeinflussen (Bronkhorst & Plomp, 1990). Während durch Kopfschatten bedingte Pegelunterschiede zwischen den Ohren in hallfreier Umgebung 20 dB betragen können, findet man keine so hohen Pegelunterschiede in Räumen mit Nachhall, da der Hall ja im Allgemeinen aus vielen Richtungen kommt. In dem von Bronkhorst und Plomp (1990) verwendeten Raum war der indirekte Schall nur 9 dB leiser als der Direktschall. Als Folge konnte ein Separationsvorteil für die Bedingungen mit frontalem Zielreiz und links oder rechts platziertem Distraktor (im Vergleich zu ebenfalls frontalem Distraktor) von nur etwa 6 dB gefunden

werden, während dieser Vorteil unter hallfreien Bedingungen in einer Untersuchung von 1988 (Bronkhorst & Plomp) noch 4 dB höher gewesen war.

Ähnlich wie das Zielreiz-Distraktor-Lautstärkeverhältnis kann auch das Hören in den Lücken durch Reflexionen im Raum beeinträchtigt werden. Bronkhorst und Plomp (1990) verwendeten kontinuierliches und moduliertes Rauschen als Maskierer, wobei bei letzterem die Amplituden der Frequenzbereiche ober- und unterhalb von 1 kHz so moduliert wurden, wie dies bei Sprache in den Amplituden der gleichen Bereichen vorgefunden wurde. Der Vorteil dieses modulierten Maskierers gegenüber einem kontinuierlichen wurde durch die Reflexionen auch hier verringert. Da die Reflexionen auch in die Lücken des modulierten Rauschens fallen können, verschlechtert sich in den Lücken folglich das Lautstärkeverhältnis zwischen Zielreiz und Maskierer, was eine nahe liegende Erklärung für den kleineren Vorteil im Vergleich zum Freifeld ist.

Konsistent dazu konnten Culling, Hodder und Toh (2003) eine Reduktion des Separationsvorteils in einer Untersuchung mit normal oder monoton intonierten Sprachsignalen bei Nachhall in einem virtuellen Raum beobachten: Solange eine schalltote virtuelle Umgebung verwendet wurde, waren beide Signale räumlich getrennt von interferierender Sprache besser verständlich, als wenn alle Signale von einem Ort präsentiert wurden. Wurde jedoch ein virtueller Raum mit einer Nachhallzeit von 400 ms simuliert, so ging der Vorteil durch räumliche Separation verloren. Dieses Ergebnis konnte in einer weiteren Untersuchung repliziert werden, wobei sich die Beseitigung des Separationseffektes durch Raumhall nicht nur bei natürlicher, sondern auch bei invertiert intonierter Sprache zeigte.

Weitere Untersuchungen kommen ebenfalls zu dem Ergebnis, dass Raumhall die Wirksamkeit räumlicher Hinweise reduziert (Koehnke & Besing, 1996; Lee & Shinn-Cunningham, 2008). Ein nur zum Teil ähnliches Bild zeigt die Untersuchung von Kidd, Mason, Brughera und Hartmann (2005): Die Autoren boten verschiedenartige Sprachsignale dar, die zu unterschiedlich starker energetischer oder informatorischer Maskierung führen sollten. Insgesamt zeigte sich, dass der Nachhall eines Raumes den Separationsvorteil vor allem bei Distraktoren mit hoher energetischer Maskierung des Zielreizes stark beeinträchtigen. Bei Distraktoren mit hoher informatorischer Maskierung des Zielreizes wurde hin-

gegen keine Reduktion des Separationsvorteils durch Nachhall gefunden. Auch Freyman, Helfer, McCall und Clifton (1999) kamen zu dem Schluss, dass der Vorteil räumlicher Separation bei informatorischer Maskierung weniger stark durch Nachhall beeinträchtigt wird als bei energetischer Maskierung.

Eine nachteilige Wirkung von Nachhall ist ebenso für Lokalisationsaufgaben bekannt. Giguere und Abel (1993) verwendeten Rauschbänder mit einer Breite von einer Dritteloctave als Zielreize und erhielten als Ergebnis, dass eine Verlängerung der Nachhallzeit eines Raumes von, je nach Frequenz, 0,15 bis 0,2 s auf 0,6 bis 1 s zu schlechteren Lokalisationsleistungen führte. Zu ähnlichen Ergebnissen kommt auch Begault (1992) in einer Untersuchung mit über Kopfhörer dargebotenen Reizen. In der Bedingung mit simuliertem Nachhall verschlechterte sich die Genauigkeit der horizontalen und vertikalen Lokalisation im Vergleich zu einer Bedingung ohne Nachhall.

Jedoch gibt es eine Untersuchung, aus deren Ergebnis ein Vorteil durch bestimmte frühe Reflexionen geschlussfolgert wurde: Hartmann (1983) ging der Frage nach, ob die Genauigkeit, mit der Schallquellen lokalisiert werden können, von der Akustik des Raumes abhängt. Dazu wurde ein Raum verwendet, dessen Deckenhöhe variiert werden konnte, wobei Reflexionen von der Decke (mit der gleichen horizontalen Richtung wie die Quelle) den Hörer bei niedriger Decke früher erreichten als bei hoher Deckenhöhe. Im Ergebnis zeigte sich, dass die Lokalisationsleistung bei niedriger Decke besser war als bei hoher. Der Autor interpretiert zwar, dass frühe Reflexionen aus der Richtung der Quelle die Richtung der (horizontalen) Lokalisation der Quelle verstärken, jedoch existierte keine Kontrollbedingung ohne Reflexionen. In der Bedingung mit hoher Deckenhöhe (und schlechterer Lokalisationsleistung) trafen die Reflexionen von der Decke (mit der gleichen horizontalen Richtung wie die Quelle) später beim Hörer ein als die lateralen Reflexionen (mit anderen horizontalen Richtungen). Da die Reflexionen in der Bedingung mit der niedrigen Decke (und besserer Lokalisationsleistung) schon vor den lateralen Reflexionen eintrafen, seien sie nach Meinung des Autors für die Lokalisation der Quelle unterstützend wirksam. Allein aus den Ergebnissen dieser Studie, d.h. ohne eine reflexionsfreie Kontrollbedingung, kann der absolute Effekt der Reflexionen jedoch nicht beurteilt werden.

Eine spätere Untersuchung von Rakerd und Hartmann (1985) mit vertikalen und lateralen oder ohne Reflexionen kommt zu einem anderen Ergebnis als die vorherige Studie: Die Genauigkeit der horizontalen Lokalisation unterschied sich für Bedingungen mit vertikalen und ohne Reflexionen bei Verwendung von perkussiven Reizen nicht – langsam einschwingende Reize führten mit vertikalen Reflexionen tendenziell zu schlechteren Lokalisationen im Vergleich zur Bedingung ohne Reflexionen. Laterale Reflexionen hatten jedenfalls wie erwartet eine ungenauere (horizontale) Lokalisation von Schallquellen zur Folge, was in der zuvor erwähnten Untersuchung zumindest tendenziell auch der Fall war. Somit erscheint es unplausibel, dass räumliche Hinweise durch einzelne Reflexionen unterstützt werden.

Der Nachhall eines Raumes scheint also die Lokalisation von Schallquellen zu beeinträchtigen und den Vorteil räumlich separierter Schallquellen bei energetischer Maskierung erheblich zu reduzieren. Vor dem Hintergrund, dass der Mensch sich nicht schon immer in kubischen Räumen aufhielt, ist plausibel, dass das menschliche Gehör evolutionär besser an die akustischen Bedingungen im Freien angepasst ist. In der freien Umgebung, wo weniger Reflexionen als in einem modernen Raum anzutreffen sind, sollte räumliches Hören erwartungsgemäß besser funktionieren. Eine Rolle mag außerdem spielen, dass Situationen mit vielen Reflexionen komplexer sind als Situationen mit nur wenigen oder ohne Reflexionen, womit auch höherer perzeptueller Verarbeitungsaufwand einhergehen könnte.

Da Eramudugolla et al. (2008) im Unterschied zu der vorliegenden Untersuchung einen virtuellen Raum ohne Nachhall verwendeten, kann der dort gefundene deutliche Separationseffekt in der Sensitivität möglicherweise durch günstigere akustische Bedingungen ohne störenden Nachhall zustande gekommen sein. Wenn dies der Fall ist, sollte bei Durchführung der beschriebenen Untersuchung in einer Umgebung ohne Nachhall ebenso ein deutlicher Separationseffekt zu finden sein. In Untersuchung 2 wird diese Hypothese empirisch geprüft.

## 6.2. Untersuchung 2 - Separationseffekt im Raum ohne Reflexionen

Eine weitere Untersuchung sollte die Frage klären, ob die begrenzte Ausprägung des Effektes der räumlichen Separation in Untersuchung 1 auf den Nachhall des Raumes zurückgeführt werden kann und dieser Effekt bei Reduktion des Nachhalls auch für alle Anzahlen gleichzeitig präsentierter Geräusche nachgewiesen werden kann. Untersuchung 2 wurde daher in einem reflexionsarmen Raum mit sehr geringem Nachhall durchgeführt. Somit besteht hier eine noch höhere Vergleichbarkeit zur Studie von Eramudugolla et al. (2008), die einen virtuellen, reflexionsfreien Raum verwendeten.

Für die Bedingungen mit räumlich separierten Geräuschen werden wieder eine höhere Sensitivität oder kürzere Reaktionszeiten oder beides erwartet – im Vergleich zu den Bedingungen mit räumlich nicht separierten Geräuschen. Es wird – konträr zu Eramudugolla et al. (2008) – keine Interaktion zwischen der räumlichen Separation und der Präsentationsdauer vermutet, sondern etwa gleich große Separationsvorteile für alle Präsentationsdauern. Der Separationseffekt sollte größer ausfallen als in Untersuchung 1. Da die Methode von Untersuchung 2 in weiten Teilen identisch mit der von Untersuchung 1 ist, werden hier nur abweichende Teile dargestellt.

### 6.2.1. Methode

#### 6.2.1.1. Stichprobe

Von den 88 Personen, die an der Untersuchung teilnahmen, wurden neun von der Auswertung ausgeschlossen. Bei fünf Personen stürzte die Experimentalsoftware vor dem Ende der Untersuchung ab, und vier zeigten zu viele Antizipationen. Die hier beschriebenen Ergebnisse beziehen sich auf die Daten der verbleibenden 79 Personen, von denen 58 weiblich waren. Das Alter lag zwischen 19 und 46 Jahren mit einem Durchschnitt von 25 Jahren.

### 6.2.1.2. Material

Die Untersuchung wurde in einem reflexionsarmen Raum durchgeführt, dessen Decke, Boden und Seitenwände mit porösen Absorbieren (nämlich Glasfaserkeilen) ausgestattet waren. Die Nachhallzeit betrug höchstens etwa 100 ms für Frequenzen größer als 100 Hz.

### 6.2.1.3. Versuchsplan

Eine A-priori-Analyse mit einer angenommenen Populationskorrelation von  $\rho = .5$  zeigte, dass eine Stichprobengröße von 55 Personen erforderlich ist, um einen Effekt (d.h. einen Unterschied in der Größe des Separierungseffekts für die verschiedenen Präsentationsdauern) von  $f = 0.2$  mit  $\alpha = \beta = .05$  nachweisen zu können. Eine Sensitivitätsanalyse mit einer Stichprobengröße von 79 Teilnehmern und ansonsten gleichen Kennwerten ergab, dass ein Effekt von  $f = 0.166$  oder größer gefunden werden konnte.

### 6.2.2. Ergebnisse

Dem deskriptiven Erscheinungsbild nach nahm die Sensitivität mit der Präsentationsdauer zu und war bei räumlich getrennter Darbietung größer als bei Darbietung zusammen von einem Ort (siehe Abbildung 6, links). Auch kann bei wenigen gleichzeitig präsentier-

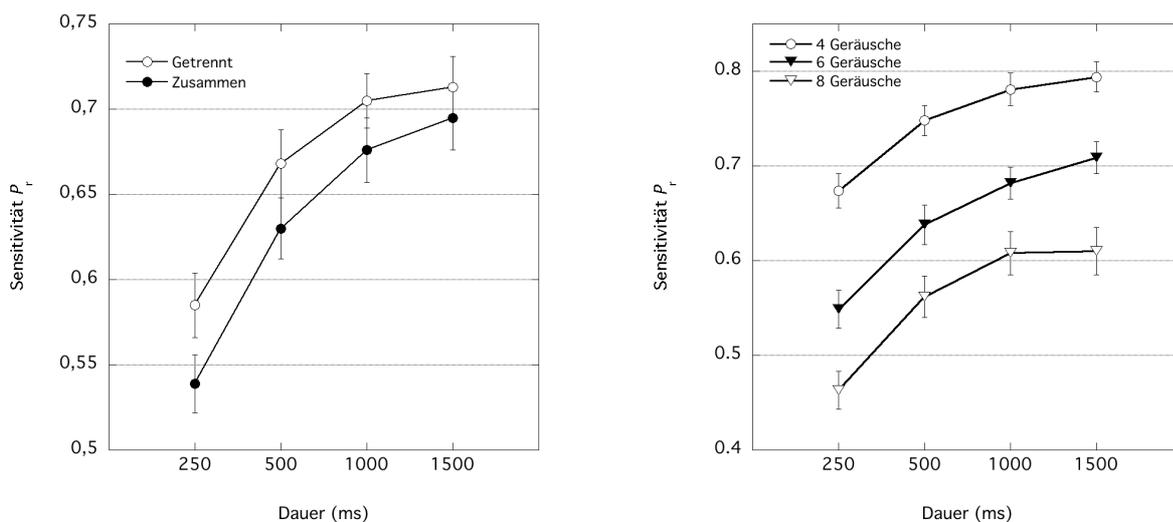


Abbildung 6: Die Sensitivität  $P_r$  für Untersuchung 2 in Abhängigkeit von der Separation und der Präsentationsdauer (links) sowie in Abhängigkeit von der Anzahl und der Präsentationsdauer der gleichzeitig präsentierten Geräusche (rechts). Die Fehlerbalken zeigen Standardfehler.

ten Geräuschen deskriptiv eine höhere Sensitivität als bei vielen beobachtet werden (siehe Abbildung 6, rechts).

Tabelle 4

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Sensitivität für Untersuchung 2. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		$F$	$p$	$\eta^2$
	Zähler	Nenner			
<b>Separation</b>	<b>1</b>	<b>78</b>	<b>17.584</b>	<b>&lt; .001</b>	<b>.184</b>
<b>Anzahl</b>	<b>2</b>	<b>77</b>	<b>147.918</b>	<b>&lt; .001</b>	<b>.793</b>
<b>Dauer</b>	<b>3</b>	<b>76</b>	<b>72.738</b>	<b>&lt; .001</b>	<b>.742</b>
Separation x Anzahl	2	77	0.819	.445	.021
Separation x Dauer	3	76	1.487	.225	.055
Anzahl x Dauer	6	73	1.139	.349	.086
Separation x Anzahl x Dauer	6	73	1.229	.301	.092

Für die abhängige Variable Sensitivität zeigte eine 2 x 3 x 4-MANOVA mit den messwiederholten Faktoren Separation, Anzahl und Dauer einen Haupteffekt der Separation mit

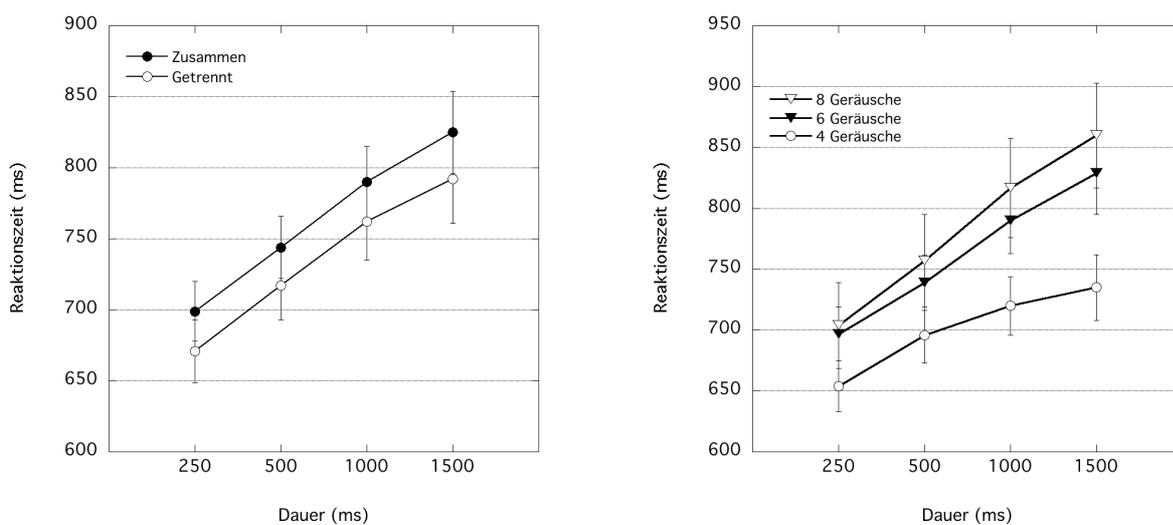


Abbildung 7: Die Reaktionszeit für Untersuchung 2 in Abhängigkeit von der Separation und der Präsentationsdauer (links) sowie in Abhängigkeit von der Anzahl und der Präsentationsdauer der gleichzeitig präsentierten Geräusche (rechts). Die Fehlerbalken zeigen Standardfehler.

höherer Sensitivität in den Bedingungen mit getrennten Geräuschen. Zusätzlich konnte ein Haupteffekt der Anzahl beobachtet werden mit augenscheinlich höherer Sensitivität bei geringer Anzahl an Geräuschen und ein Haupteffekt der Dauer, wobei die Sensitivität dem visuellen Eindruck nach mit längerer Dauer zunahm. Die statistischen Kennwerte sind in Tabelle 4 aufgeführt.

Tabelle 5

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Reaktionszeit für Untersuchung 2. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		<i>F</i>	<i>p</i>	$\eta^2$
	Zähler	Nenner			
<b>Separation</b>	<b>1</b>	<b>78</b>	<b>11.690</b>	<b>.001</b>	<b>.130</b>
<b>Anzahl</b>	<b>2</b>	<b>77</b>	<b>38.065</b>	<b>&lt; .001</b>	<b>.497</b>
<b>Dauer</b>	<b>3</b>	<b>76</b>	<b>40.543</b>	<b>&lt; .001</b>	<b>.615</b>
<b>Anwesenheit</b>	<b>1</b>	<b>78</b>	<b>319.218</b>	<b>&lt; .001</b>	<b>.804</b>
Separation*Anzahl	2	77	1.535	.222	.038
Separation*Dauer	3	76	0.315	.815	.012
Separation*Anwesenheit	1	78	1.582	.212	.020
<b>Anzahl*Dauer</b>	<b>6</b>	<b>73</b>	<b>10.465</b>	<b>&lt; .001</b>	<b>.462</b>
Anzahl*Anwesenheit	2	77	0.307	.736	.008
<b>Dauer*Anwesenheit</b>	<b>3</b>	<b>76</b>	<b>14.198</b>	<b>&lt; .001</b>	<b>.359</b>
Separation*Anzahl*Dauer	6	73	1.568	.169	.114
Separation*Anzahl*Anwesenheit	2	77	0.927	.400	.024
Separation*Dauer*Anwesenheit	3	76	0.757	.522	.029
Anzahl*Dauer*Anwesenheit	6	73	1.237	.298	.092
Separation*Anzahl*Dauer*Anwesenheit	6	73	0.648	.691	.051

Deskriptiv nahmen die Reaktionszeiten mit der Präsentationsdauer zu und waren bei getrennter Darbietung der Geräusche kürzer als bei Darbietung zusammen von einem Ort

(siehe Abbildung 7, links). Außerdem zeigten sich bei wenigen Geräuschen deskriptiv kürzere Reaktionszeiten als bei vielen (siehe Abbildung 7, rechts).

Für die abhängige Variable Reaktionszeit zeigte eine  $2 \times 2 \times 3 \times 4$ -MANOVA mit den messwiederholten Faktoren Anwesenheit, Separation, Anzahl und Dauer einen Haupteffekt der Anwesenheit mit kürzeren Reaktionszeiten bei anwesendem Zielreiz und einen Haupteffekt der Separation mit kürzeren Reaktionszeiten bei getrennter Präsentation. Zusätzlich konnte ein Haupteffekt der Anzahl beobachtet werden mit augenscheinlich kürzeren Reaktionszeiten bei geringer Anzahl an Geräuschen und ein Haupteffekt der Dauer mit deskriptiv kürzeren Reaktionszeiten bei kürzerer Präsentation. Außerdem wurde eine Interaktion zwischen Anzahl und Dauer (siehe Abbildung 7, rechts) sowie eine zwischen Dauer und Anwesenheit gefunden. Gemessen an den Tendenzen der Effekte der zugehörigen einzelnen Bedingungen schränken beide Interaktionen die Interpretation der Haupteffekte jedoch nicht ein. Tabelle 5 zeigt eine Übersicht über alle statistischen Kennwerte der Effekte für die Variable Reaktionszeit.

### 6.2.3. Diskussion

Unter reflexionsarmen Bedingungen zeigte sich sowohl in der Sensitivität als auch in den Reaktionszeiten ein Vorteil räumlich separierter Geräusche im Vergleich zu von einem Ort präsentierten Geräuschen. Dieser Effekt der räumlichen Separation hatte Bestand bei gleichzeitiger Präsentation von vier, sechs und acht Geräuschen. Der Haupteffekt durch räumliche Separation in der Untersuchung von Eramudugolla et al. (2008) konnte also repliziert und gleichzeitig ein Geschwindigkeits-Genauigkeitsaustausch ausgeschlossen werden. Untersuchung 2 liefert somit bisher fehlende Evidenz für die Nutzung räumlicher Hinweise in einer komplexen akustischen Suchaufgabe, in der die Position des Ziel nicht nur aufgabenirrelevant, sondern auch unvorhersagbar ist. Im Gegensatz zu den Ergebnissen von Eramudugolla et al. (2008) erreichte die Interaktion zwischen Separation und Dauer jedoch weder in der Sensitivität noch in den Reaktionszeiten das Signifikanzniveau ( $p = .225$  und  $p = .815$ ).

Ein Grund für das Auftreten der Interaktion zwischen der räumlicher Separation und der Präsentationsdauer in der zweiten Untersuchung von Eramudugolla et al. (2008) könnte eventuell auch in der geringen Teilnehmerzahl in Kombination mit möglicherweise nicht-normalverteilten Daten zu suchen sein, was theoretisch zu einem falsch positiven Ergebnis der Varianzanalyse geführt haben könnte. Nach dem zentralen Grenzwerttheorem geht die Verteilung von Mittelwerten mit wachsendem Stichprobenumfang zwar in eine Normalverteilung über, jedoch wird die dafür nötige Größe der Stichprobe bei etwa 30 verortet (Bortz & Schuster, 2010). Diese Anzahl wurde in der Untersuchung 2 von Eramudugolla et al. (2008) mit 17 Teilnehmern jedoch unterschritten, so dass die Wahrscheinlichkeit, eine Verzerrung der Ergebnisse durch diesen Einfluss zu erhalten, gegeben war. Im Gegensatz dazu gingen in Untersuchung 2 dieser Arbeit 79 Personen in die Auswertung ein, so dass hier die Wahrscheinlichkeit falsch positiver Ergebnisse geringer war.

Die Lokation der Geräusche ist für die Bearbeitung der Suchaufgabe nicht relevant, da die Teilnehmer vorab keine Information darüber erhalten, von welchem Ort der Zielreiz oder die Distraktoren eingespielt werden werden und auch keine Angaben über die Lokationen machen sollen. Wie die Sensitivität in den Bedingungen ohne räumliche Separation zeigt, kann die Aufgabe auch ohne für das Zielgeräusch spezifische räumliche Hinweise ausreichend genau bearbeitet werden. Hinzu kommt, dass räumliche Hinweise in der auditiven Modalität häufig als weniger wirkungsvoll erachtet werden – etwa im Vergleich zu Hinweisen bezüglich Harmonizität oder des Ein- und Ausschwingverhaltens (Darwin, 2006). Dennoch zeigt sich ein Vorteil durch räumliche Separation sowohl für leichtere (lange Präsentationsdauer und wenige Geräusche gleichzeitig) als auch für schwerere (kurze Präsentationsdauer und viele Geräusche gleichzeitig) akustische Suchaufgaben.

Auch wenn ein Vergleich über verschiedene Untersuchungen methodisch weniger günstig ist als ein Vergleich zwischen Bedingungen innerhalb einer Untersuchung, konnte die Nachhallzeit als Einflussfaktor auf den Separationseffekt bestätigt werden. Neben der Nachhallzeit haben wahrscheinlich noch zahlreiche weitere Parameter einen Einfluss auf den Effekt räumlich separierter Schallquellen. Aus Untersuchungen zur akustischen Aufmerksamkeit etwa ist bekannt, dass sowohl Frequenzhinweise als auch Hinweise mit Rauminformation die Aufmerksamkeit lenken und in der Folge die Leistung bei Detek-

tions- oder Diskriminationsaufgaben beeinflussen können (Mondor, Breau, & Milliken, 1998). In der folgenden Untersuchung 3 wird der Frage nachgegangen, ob die Modalität des Hinweisreizes vor jedem Durchgang – und damit die Darbietung von Frequenzhinweisen – ebenso eine modulierende Wirkung auf die Nutzung räumlicher Hinweise hat. Im nächsten Abschnitt wird erläutert, aus welchen Gründen eine modulierende Wirkung erwartet werden kann.

### 6.3. Untersuchung 3 - Separationseffekt mit akustischem Hinweis

In Untersuchung 2 zeigte sich ein Separationseffekt, obwohl räumliche Hinweise in akustischen Such- oder Identifikationsaufgaben nach Meinung mehrerer Autoren vor allem dann in Betracht gezogen werden, wenn es keine anderen Unterschiede zwischen akustischen Ereignissen gibt (Darwin, 2006; Scharf, 1998; Shinn-Cunningham, 2005). Da größtmögliche spektrotemporale Unterschiedlichkeit ein Kriterium bei der Auswahl der verwendeten Reize war und dennoch ein Separationsvorteil beobachtet werden konnte, wurden spektrotemporale Hinweise möglicherweise wegen begrenzter Repräsentation im Gedächtnis nicht in vollem Umfang genutzt. In Studien, in denen die Aufmerksamkeit durch einen Hinweis auf einen bestimmten Frequenzbereich gelenkt wurde, zeigte sich, dass auch von der Frequenz unabhängige Eigenschaften eines Ziels schneller und genauer erkannt werden können, wenn sich auch das Ziel in der Nähe dieses Frequenzbereiches befindet (Mondor & Bregman, 1994). Spektrotemporale Eigenschaften des Zielgeräusches sollten also dann besser genutzt werden können, wenn die Aufmerksamkeit auf passende Frequenzbereiche gerichtet ist. In den Untersuchungen 1 und 2 mussten die Teilnehmer bei Erscheinen des Hinweiswortes am Anfang eines Durchgangs auf ihr Langzeitgedächtnis zurückgreifen und spektrotemporale Eigenschaften des Zielgeräusches rekonstruieren, um einen Abgleich mit den gehörten Ereignissen vollziehen zu können. Dieser Abgleich sollte leichter fallen, wenn das Zielgeräusch nicht durch ein Hinweiswort, sondern akustische Präsentation des Zielgeräusches vor jedem Durchgang gekennzeichnet wird. Dann läge – unabhängig von der Repräsentation im Langzeitgedächtnis – die spektrotemporale Information im echoischen Gedächtnis vor, was auf der anderen Seite die Nutzung räumlicher Hinweise unwahrscheinlicher machen sollte (Darwin, 2006; Scharf, 1998). War also

eine begrenzte Repräsentation spektrottemporaler Information im Gedächtnis an dem Vorteil durch räumliche Hinweise beteiligt, so sollte bei akustischem Hinweis für das Zielgeräusch (der mit dem Zielgeräusch identisch ist) vor jedem Durchgang ein geringerer als bei visuellem Zielhinweis oder gar kein Separationseffekt mehr zu finden sein.

Für die Bedingungen mit räumlich separierten Geräuschen im Vergleich zu nicht separierten Geräuschen werden eine höhere Sensitivität oder kürzere Reaktionszeiten oder beides erwartet – jedoch in geringerem Ausmaß als in Untersuchung 2. Es wird keine Interaktion zwischen der räumlichen Separation und der Präsentationsdauer vermutet, sondern etwa gleich große Separationsvorteile für alle Präsentationsdauern. Die Methode von Untersuchung 3 ist in weiten Teilen identisch mit der von Untersuchung 2, weshalb hier nur die Unterschiede dargestellt sind.

### 6.3.1. Methode

#### 6.3.1.1. Stichprobe

Von den 81 Teilnehmern der zweiten Untersuchung wurden vier wegen zu vieler Antizipationen oder verspäteter Reaktionen von der Auswertung ausgeschlossen. In die hier beschriebenen Ergebnisse gingen die Daten von 77 Personen ein, von denen 54 weiblich waren. Das Alter der Teilnehmer lag zwischen 18 und 40 Jahren, mit einem Durchschnitt von 25 Jahren.

#### 6.3.1.2. Versuchsdurchführung

Am Beginn eines Durchgangs wurde das zu suchende Geräusch als Hinweis akustisch dargeboten – und zwar aus allen Lautsprechern gleichzeitig. In der Bedingung mit 250 ms Präsentationsdauer der Geräusche wurde der Hinweis 250 ms lang präsentiert, danach 500 ms Stille und schließlich das Geräuschgemisch. In allen anderen Bedingungen war der Hinweis 500 ms lang und die darauf folgende Pause 250 ms. Auf dem Monitor wurde währenddessen weiterhin das Fixationskreuz angezeigt.

### 6.3.1.3. Versuchsplan

Eine A-priori-Analyse mit einer angenommenen Populationskorrelation von  $\rho = .5$  zeigte, dass eine Stichprobengröße von 55 Personen erforderlich ist, um einen Effekt (Unterschied in der Größe des Separierungseffekts für die verschiedenen Präsentationsdauern) von  $f = 0.2$  mit  $\alpha = \beta = .05$  nachweisen zu können. Eine Sensitivitätsanalyse mit einer Stichprobengröße von 77 Teilnehmern und ansonsten gleichen Kennwerten ergab, dass ein Effekt von  $f = 0.168$  oder größer gefunden werden konnte.

### 6.3.2. Ergebnisse

Dem deskriptiven Erscheinungsbild nach nahm die Sensitivität mit der Präsentationsdauer zu und war bei räumlich getrennter Darbietung ähnlich groß wie bei Darbietung zusammen von einem Ort (siehe Abbildung 8, links). Wie bereits in den Experimenten zuvor gefunden, kann bei wenigen gleichzeitig präsentierten Geräuschen deskriptiv eine höhere Sensitivität als bei vielen beobachtet werden (siehe Abbildung 8, rechts).

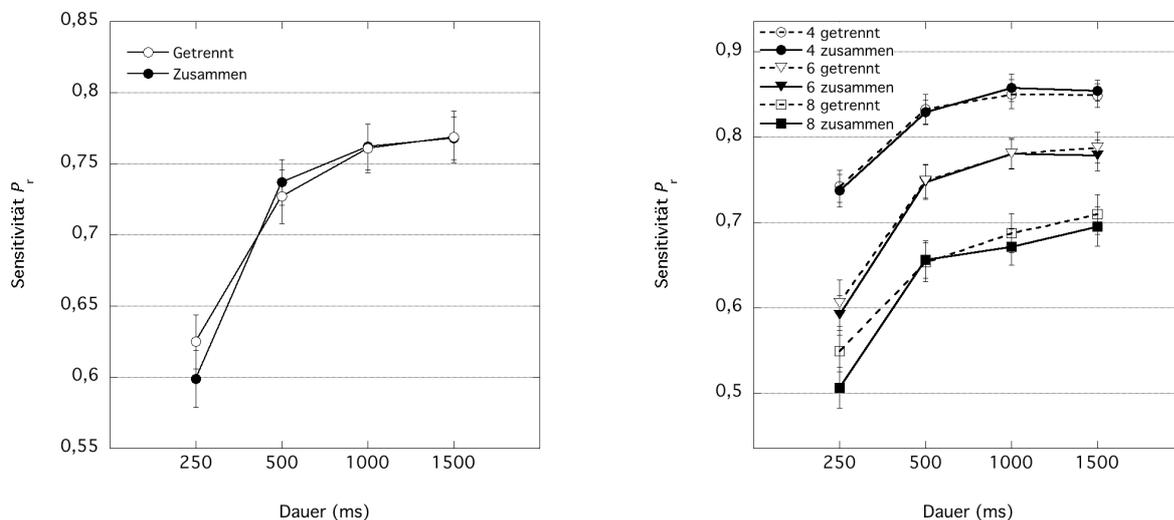


Abbildung 8: Die Sensitivität  $P_r$  für Untersuchung 3 in Abhängigkeit von der Separation und der Präsentationsdauer (links) sowie in Abhängigkeit von der Separation, der Anzahl und der Präsentationsdauer (rechts) der gleichzeitig präsentierten Geräusche. Die Fehlerbalken zeigen Standardfehler.

Für die abhängige Variable Sensitivität  $P_r$  zeigte eine  $2 \times 3 \times 4$ -MANOVA mit den messwiederholten Faktoren Separation, Anzahl und Dauer keinen Haupteffekt der Separation (siehe Abbildung 8, links), aber einen Haupteffekt der Anzahl (siehe Abbildung 8, rechts),

und einen Haupteffekt der Dauer. Es trat eine Interaktion zwischen Anzahl und Dauer auf. Abbildung 8 (rechts) zeigt, dass die Sensitivität bei vier Geräuschen und 1500 ms Dauer zumindest tendenziell kleiner war als bei 1000 ms. Die statistischen Kennwerte sind in Tabelle 6 aufgeführt.

Tabelle 6

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Sensitivität für Untersuchung 3. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		<i>F</i>	<i>p</i>	$\eta^2$
	Zähler	Nenner			
Separation	1	76	1.059	.307	.014
<b>Anzahl</b>	<b>2</b>	<b>75</b>	<b>113.887</b>	<b>&lt; .001</b>	<b>.752</b>
<b>Dauer</b>	<b>3</b>	<b>74</b>	<b>96.771</b>	<b>&lt; .001</b>	<b>.797</b>
Separation x Anzahl	2	75	0.632	.534	.017
Separation x Dauer	3	74	0.590	.623	.023
<b>Anzahl x Dauer</b>	<b>6</b>	<b>71</b>	<b>3.575</b>	<b>.004</b>	<b>.232</b>
Separation x Anzahl x Dauer	6	71	0.330	.919	.027

Deskriptiv nahmen die Reaktionszeiten mit der Präsentationsdauer zu und waren bei getrennter Darbietung der Geräusche tendenziell kürzer als bei Darbietung zusammen von einem Ort (siehe Abbildung 9, links). Außerdem zeigten sich bei wenigen Geräuschen deskriptiv kürzere Reaktionszeiten als bei vielen (siehe Abbildung 9, rechts).

Für die abhängige Variable Reaktionszeit zeigte eine 2 x 2 x 3 x 4-MANOVA mit den messwiederholten Faktoren Separation, Anzahl, Dauer und Anwesenheit keinen Haupteffekt der Separation, aber einen Haupteffekt der Anzahl, einen Haupteffekt der Dauer und einen Haupteffekt der Anwesenheit mit längeren Reaktionszeiten bei abwesendem Zielreiz.

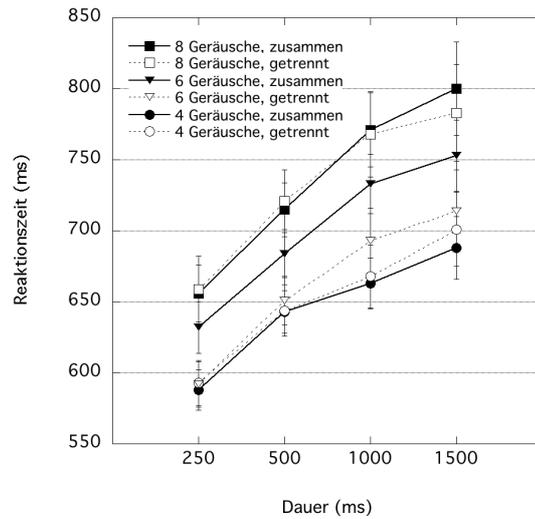
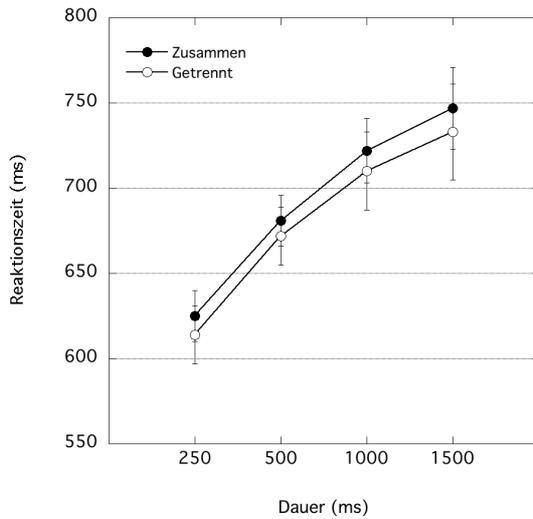


Abbildung 9: Die Reaktionszeit für Untersuchung 3 in Abhängigkeit von der Separation und der Präsentationsdauer (links) sowie der Separation, der Anzahl und der Präsentationsdauer (rechts) der gleichzeitig präsentierten Geräusche. Die Fehlerbalken zeigen Standardfehler.

Es trat eine Interaktion zwischen Separation und Anzahl auf (Abbildung 9, rechts). Gerichtete Post-hoc-Tests lieferten einen Separationsvorteil auf adjustiertem Alphaniveau ( $\alpha = .05/3 \approx .017$ ) für die Bedingung mit sechs gleichzeitig präsentierten Geräuschen,  $F(1,76) =$

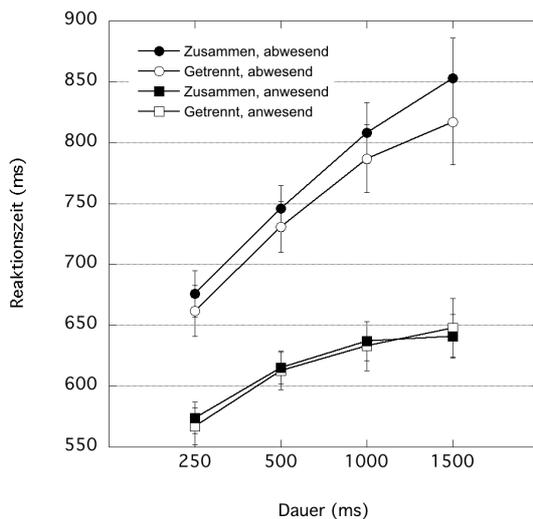


Abbildung 10: Die Reaktionszeit für Untersuchung 3 in Abhängigkeit von der Anwesenheit des Zielgeräusches, der Separation und der Präsentationsdauer der gleichzeitig präsentierten Geräusche. Die Fehlerbalken zeigen Standardfehler.

6.404,  $p = .007$ ,  $\eta^2 = .078$ , nicht jedoch für die Bedingungen mit vier,  $F(1,76) = 0.173$ ,  $p = .339$ ,  $\eta^2 = .002$ , oder acht Geräuschen,  $F(1,76) = 0.038$ ,  $p = .423$ ,  $\eta^2 < .001$ . Es wurde auch eine Interaktion zwischen Separation und Anwesenheit (Abbildung 10) sowie zwischen Separation, Dauer und Anwesenheit gefunden (ebenfalls Abbildung 10).

Tabelle 7

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Reaktionszeit für Untersuchung 3. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		F	p	$\eta^2$
	Zähler	Nenner			
Separation	1	76	1.556	.216	.020
<b>Anzahl</b>	<b>2</b>	<b>75</b>	<b>17.932</b>	<b>&lt; .001</b>	<b>.324</b>
<b>Dauer</b>	<b>3</b>	<b>74</b>	<b>65.072</b>	<b>&lt; .001</b>	<b>.725</b>
<b>Anwesenheit</b>	<b>1</b>	<b>76</b>	<b>106.761</b>	<b>&lt; .001</b>	<b>.584</b>
<b>Separation x Anzahl</b>	<b>2</b>	<b>75</b>	<b>3.709</b>	<b>.029</b>	<b>.090</b>
Separation x Dauer	3	74	0.243	.866	.010
<b>Separation x Anwesenheit</b>	<b>1</b>	<b>76</b>	<b>5.902</b>	<b>.017</b>	<b>.072</b>
<b>Anzahl x Dauer</b>	<b>6</b>	<b>71</b>	<b>2.410</b>	<b>.035</b>	<b>.169</b>
<b>Anzahl x Anwesenheit</b>	<b>2</b>	<b>75</b>	<b>4.005</b>	<b>.022</b>	<b>.096</b>
<b>Dauer x Anwesenheit</b>	<b>3</b>	<b>74</b>	<b>15.705</b>	<b>&lt; .001</b>	<b>.389</b>
Separation x Anzahl x Dauer	6	71	0.994	.436	.077
Separation x Anzahl x Anwesenheit	2	75	3.030	.054	.075
<b>Separation x Dauer x Anwesenheit</b>	<b>3</b>	<b>74</b>	<b>3.109</b>	<b>.031</b>	<b>.112</b>
Anzahl x Dauer x Anwesenheit	6	71	1.109	.366	.086
Separation x Anzahl x Dauer x Anwesenheit	6	71	1.407	.224	.106

Nach Adjustierung des Alphaniveaus und bei gerichteter Testung ( $\alpha = .05/2 = .025$ ) kann ein Separationseffekt nicht für die Bedingung mit anwesendem Zielreiz ( $F(1,76) = 0.027$ ,  $p = .435$ ,  $\eta^2 < .001$ ), aber für die mit abwesendem Zielreiz ( $F(1,76) = 4.277$ ,  $p = .021$ ,

$\eta^2 = .053$ ) gefunden werden. Eine weiterführende Untersuchung der Dreifachinteraktion mittels gerichteter Post-hoc-Tests liefert auf adjustiertem Alphaniveau ( $\alpha = .05/8 \approx .006$ ) keine Separationseffekte für die Bedingungen mit anwesendem Zielreiz und den Präsentationsdauern 250ms ( $F(1,76) = 0.798, p = .187, \eta^2 = .010$ ), 500ms ( $F(1,76) = 0.070, p = .396, \eta^2 = .001$ ), 1000ms ( $F(1,76) = 0.102, p = .375, \eta^2 = .001$ ) und 1500ms ( $F(1,76) = 0.288, p = .296, \eta^2 = .004$ ) und auch nicht für die Bedingungen mit abwesendem Zielreiz und den Präsentationsdauern 250ms ( $F(1,76) = 2.430, p = .062, \eta^2 = .031$ ), 500ms ( $F(1,76) = 2.546, p = .057, \eta^2 = .032$ ), 1000ms ( $F(1,76) = 2.621, p = .055, \eta^2 = .033$ ) und 1500ms ( $F(1,76) = 5.284, p = .012, \eta^2 = .065$ ).

Weiterhin zeigte sich eine Interaktion zwischen Anzahl und Dauer sowie zwischen Anzahl und Anwesenheit, die – gemessen an den Tendenzen der Effekte der zugehörigen einzelnen Bedingungen – die Interpretation der Haupteffekte jedoch nicht einschränkt. Die statistischen Kennwerte aller Effekte der messwiederholten Varianzanalyse mit der abhängigen Variable Reaktionszeit zeigt Tabelle 7.

### 6.3.3. Diskussion

Mit akustischem Zielhinweis in Untersuchung 3 konnte kein Haupteffekt der Separation gefunden werden. Lediglich für eine Anzahl von sechs gleichzeitig präsentierten Reizen zeigte sich noch ein Sensitivitätsvorteil für räumlich separierte Geräusche. Außerdem konnte ein Separationsvorteil für abwesende Zielreize in den Reaktionszeiten beobachtet werden, nicht jedoch für anwesende Zielreize. Die Post-hoc-Tests anlässlich der Dreifachinteraktion zwischen Separation, Dauer und Anwesenheit erreichten das adjustierte Alphaniveau zwar nicht, zeigten aber eine ähnliche Tendenz: Bei abwesendem Zielreiz lagen die  $p$ -Werte (gerichteter) Tests auf Separationseffekte in den vier Bedingungen unterschiedlicher Präsentationsdauer stets unterhalb von .063 – bei anwesendem Zielreiz jedoch stets oberhalb von .186. Gleichzeitig schränken die Tendenzen der Effekte der zugehörigen einzelnen Bedingungen nicht die Interpretation des Separationseffektes für abwesende Zielreize ein.

Der fehlende Haupteffekt der Separation und sein selektives Auftreten in Durchgängen mit abwesendem Zielreiz sprechen dafür, dass spektrotemporale Hinweise vermehrt ge-

nutzt werden, wenn sie – bedingt durch einen akustischen Hinweis – besser im Gedächtnis repräsentiert sind, nämlich in Durchgängen mit anwesendem Zielreiz. In Durchgängen mit abwesendem Zielreiz, in denen spektrotemporale Information nicht erfolgreich abgeglichen werden kann, nutzen Personen wahrscheinlich in stärkerem Maße auch räumliche Hinweise, um die Suchaufgabe zu bewältigen. Dies legen zumindest die Ergebnisse von Untersuchung 3 nahe, die passenderweise vor allem bei abwesendem Zielreiz Evidenz für einen Separationsvorteil aufweisen. Außerdem reagierten die Teilnehmer bei akustischem Hinweis in Untersuchung 3 genauer und schneller als bei visuellem Hinweis in Untersuchung 2, was darauf hindeutet, dass die Nutzung spektrotemporaler Hinweise hier effizienter ist als die Nutzung räumlicher Hinweise. Somit scheinen räumliche Hinweise in diesem Paradigma in einer alternativen Strategie zur Lösung der Aufgabe genutzt zu werden – anscheinend aber vor allem dann, wenn andere Strategien (z.B. Nutzung spektrotemporaler Hinweise) nicht so leicht zum Ziel führen.

Dass die Nutzung räumlicher Information abhängig von den verwendeten Hinweisen auf das zu identifizierende Ziel ist, steht im Einklang mit einer Untersuchung von Ihlefeld und Shinn-Cunningham (2008). Sie präsentierten über Kopfhörer zwei modifizierte Sätze, die aus amplitudenmodulierten Sinustönen bestanden, also synthetisch klangen, aber noch verständlich waren und zu möglichst wenig gegenseitiger energetischer Maskierung führen sollten. Der Zielsatz wurde durch ein Hinweiswort gekennzeichnet, das auf den gleichen Sinustönen wie das Ziel basierte (also eine ähnliche Klangfarbe besaß) oder die gleiche simulierte Lokation belegte oder beides. Der Prozentsatz richtig wiedergegebener Wörter des Zielsatzes war in der Bedingung mit gleicher Klangfarbe des Hinweises unabhängig von der räumlichen Separation zwischen Ziel- und Distraktorsatz. In den Bedingungen mit gleicher Lokation sowie mit gleicher Lokation und gleicher Klangfarbe zeigte sich jedoch ein Effekt der räumlichen Separation. In letzteren beiden Bedingungen war vorab ausgerichtete Aufmerksamkeit auf das Ziel für die Bearbeitung der Aufgabe jedoch zumindest möglich (beim Hinweis mit nur gleicher Lokation sogar nötig), während die Lokation des Ziels in den Untersuchungen der vorliegenden Arbeit nicht vorhersagbar war. In Untersuchung 3 kann außerdem bei separierten Schallquellen – im Gegensatz zu Ihlefeld und Shinn-Cunningham – trotz Hinweises ohne Lokationsinformation eine höhe-

re Sensitivität zumindest für die Bedingung mit sechs gleichzeitig präsentierten Geräuschen gefunden werden. Zusätzlich existierten in Untersuchung 3 (im Gegensatz zu der Untersuchung von Ihlefeld und Shinn-Cunningham) auch Durchgänge mit abwesendem Ziel, in denen eine Separation zu schnelleren Reaktionszeiten führte. Im Vergleich zu Untersuchung 2 scheint – insofern teilkonsistent mit Ihlefeld und Shinn-Cunningham – der Separationseffekt in Untersuchung 3 mit akustischem und Klangfarbeninformation lieferndem Hinweis jedoch insgesamt eher reduziert zu sein.

#### 6.4. Untersuchung 4 - Separationseffekt und Ähnlichkeit von Reizen

Die bisherigen Untersuchungen lassen die Interpretation zu, dass räumliche Hinweise vor allem dann genutzt werden, wenn spektrotemporale Information als Unterscheidungsmerkmal weniger gut geeignet ist. In Untersuchung 4 soll diese Erklärung durch Manipulation der Ähnlichkeit zwischen Ziel- und Distraktorreizen überprüft werden. Eine mögliche Hypothese wäre, dass die spektrotemporale Information sich schwieriger abgleichen lässt, wenn sich Ziel und Distraktor ähneln, da die Gefahr der Verwechslung dann eher gegeben ist. Zumindest konnte eine geringere Sprachverständlichkeit gefunden werden bei höherer Ähnlichkeit zwischen Zielsprache und Distraktor – hinsichtlich des Stimulustyps (Sprache versus Sprache im Vergleich zu Sprache versus Rauschen; Hawley et al., 2004) und bei Sprache hinsichtlich des Geschlechts der Sprecher (Ericson & McKinley, 1997). In einer Bedingung mit hoher Ähnlichkeit zwischen Ziel und Distraktor könnte daher ein größerer Effekt der räumlichen Separation erwartet werden als in einer Bedingung mit geringerer Ähnlichkeit.

Dieses Ergebnis würde auch mit einer Untersuchung von Noble und Perrett (2002) übereinstimmen, die von Sprechern weiblichen oder männlichen Geschlechts vorgetragene Erzählungen mithilfe eines Lautsprechers oder dreier Lautsprecher präsentierten. Dadurch ergaben sich Bedingungen mit ähnlichen Stimmen (Ziel- und Distraktorsprecher des gleichen Geschlechts) und mit weniger ähnlichen Stimmen (Ziel- und Distraktorsprecher verschiedenen Geschlechts). Die zu beachtende Erzählung wurde stets von vorne präsentiert, die beiden irrelevanten Erzählungen entweder auch von vorne oder jeweils von Positionen 30° links bzw. rechts. Die Aufgabe der Teilnehmer bestand darin, die Lautstärke der

Zielerzählung so einzustellen, dass fast vollständige subjektive Verständlichkeit (das sogenannte 90%-Kriterium) erreicht wurde. Der Vorteil durch räumliche Separation erwies sich in der Bedingung mit weiblichem Ziel- und Distraktorsprecher als größer als mit weiblichem Ziel- und männlichem Distraktorsprecher, während sich die beiden Bedingungen mit männlichem Zielsprecher nur tendenziell unterschieden. Ein größerer Separationsvorteil konnte also eher bei akustisch ähnlichen Reizen gefunden werden. Dieses Ergebnis spricht also dafür, dass der Separationseffekt eher größer ausfällt, wenn sich Ziel- und Distraktorsprache hinsichtlich Grundfrequenz und Klangfarbe ähneln.

Auf ähnliche Weise lassen sich auch die Ergebnisse von Hawley et al. (2004) interpretieren. In ihrer Untersuchung sollten die Teilnehmer Sprache verstehen, die gleichzeitig mit ein bis drei Distraktoren präsentiert wurde. Die Distraktoren konnten entweder aus Sprache bestehen, aus rückwärts gespielter Sprache, aus Rauschen mit dem gleichen Spektrum wie Sprache oder aus mit der Lautstärkehüllkurve von Sprache moduliertem Rauschen, das ebenso das gleiche Spektrum wie Sprache besaß. In den Bedingungen mit zwei oder drei Distraktoren war der Vorteil durch räumliche Separation größer, wenn Sprache oder rückwärts gespielte Sprache verwendet wurde anstatt einer der Rauschdistraktortypen. Bei Ähnlichkeit in der Klangfarbe zwischen Ziel und Distraktor war also auch hier der Separationseffekt ausgeprägter.

Zu einem anderen Ergebnis kommen Shackleton, Meddis und Hewitt (1994), die die Rolle binauraler Hinweise und Unterschiede in der Grundfrequenz untersuchten und dazu zwei gleichzeitig präsentierte Vokale identifizieren ließen. Als binaurale Hinweise kamen interaurale Zeitdifferenzen, interaurale Pegeldifferenzen oder beide Hinweise gleichzeitig zum Einsatz, wobei alle binauralen Hinweise zu vergleichbaren Effekten führten. Außerdem besaßen die Vokale entweder die gleiche Grundfrequenz, oder die Grundfrequenzen unterschieden sich um einen Halbton. Wurden zwei verschiedene Vokale dargeboten, so war der Vorteil räumlich separierter Präsentation für Vokale mit verschiedenen Grundfrequenzen ausgeprägter als für Vokale der gleichen Grundfrequenz. Hier scheint eine räumliche Separation vor allem dann einen Vorteil zu bieten, wenn eine Objektbildung anhand

anderer Hinweise (Grundfrequenz) hinreichend gut erreicht wurde.<sup>11</sup> Die Nutzung räumlicher Hinweise scheint bei verschiedenen Vokalen also möglicherweise eine Konsequenz aus dem Vorhandensein von Hinweisen zur Grundfrequenz zu sein und findet nicht vor allem dann statt, wenn die Notwendigkeit dazu besteht, sondern wenn Frequenzhinweise die Trennung von Reizen und damit die Zuordnung ihrer räumlichen Hinweise ermöglichen.

Somit weisen zwei Untersuchungen auf einen größeren Separationseffekt bei ähnlicheren Reizen hin (Hawley et al., 2004; Noble & Perrett, 2002) und eine auf einen kleineren Separationseffekt (Shackleton et al., 1994). Allerdings erforderte die Aufgabe bei letzterer geteilte Aufmerksamkeit im Gegensatz zu den anderen Studien, deren Aufgabe mit Hilfe von selektiver Aufmerksamkeit gelöst werden sollte. Zudem weisen zumindest die beiden erstgenannten Untersuchungen etwa mit stets von vorne präsentem Ziel, anderen, teils subjektiven Leistungskriterien oder ausschließlicher Verwendung von Sprachreizen Unterschiede zu dem in dieser Arbeit verwendeten Paradigma der akustischen Suche auf, so dass sich nur mit vergleichsweise großer Unsicherheit eine Vorhersage hinsichtlich des Einflusses von Ähnlichkeit zwischen Ziel und Distraktor im Paradigma der akustischen Suche ableiten lässt. In der folgenden Untersuchung soll dieser empirischen Frage nachgegangen werden. Dazu wurde eine ähnliche akustische Suchaufgabe wie in Untersuchung 1, 2 und 3 verwendet, jedoch mit verändertem Reizmaterial. Die Ähnlichkeit der Reize wurde manipuliert durch die Verwendung verschiedener Musikinstrumentenklänge, die mit gleicher oder verschiedener Tonhöhe gespielt wurden. Um die höhere Aufgabenschwierigkeit bei ähnlichen Ziel- und Distraktorreizen zu kompensieren, wurden immer nur zwei Reize gleichzeitig dargeboten. Die Methode von Untersuchung 4 ist in Teilen identisch mit der von Untersuchung 2, weshalb hier nur abweichende Punkte beschrieben sind.

---

<sup>11</sup> Bei identischem Vokal wurde kein Vorteil räumlicher Separation gefunden und – ebenso im Gegensatz zu Durchgängen mit verschiedenen Vokalen – ein Nachteil verschiedener Grundfrequenz, was darauf hindeutet, dass die Aufgabe in diesem Falle auf eine andere Weise gelöst wird. Möglicherweise werden zwei identische Vokale mit gleicher Frequenz als nur *ein* Vokal wahrgenommen und dann – mit dem Vorwissen, dass immer zwei Vokale präsentiert wurden – in der Antwort als zwei gleiche Vokale bezeichnet.

### 6.4.1. Methode

#### 6.4.1.1. Stichprobe

108 Personen, davon 73 weiblich, gingen in die Auswertung ein. Das Alter lag zwischen 19 und 39 Jahren mit einem Mittelwert von 24 Jahren. Nicht ausgewertet wurden die Daten von fünf Personen<sup>12</sup>, die den Vortest nicht bestanden hatten, zwei Personen, die während der Instruktion abgebrochen hatten, einer Person, die ein negatives  $P_r$  in drei von vier Bedingungen zeigte (und daher die Aufgabe möglicherweise falsch verstanden hatte) und einer Person, die einarmig war und deswegen nicht die Finger über beide Reaktionstasten halten konnte. Die Teilnehmer wurden wahlweise mit einer halben Teilnehmerstunde oder drei Euro entlohnt.

#### 6.4.1.2. Material

Als akustische Reize diente eine Kombination von zwei aus acht Einzelklängen<sup>13</sup>. Dazu wurden eine gezupfte Bratsche, eine gestrichene Bratsche, eine Mundharmonika und eine Querflöte in den Tonhöhen E4 (Grundton: ca. 330 Hz) und Ais4 (Grundton ca. 466 Hz) mit einer Abtastrate von 48 kHz in einem reflexionsarmen Raum aufgezeichnet, auf 500 ms Länge beschnitten und in den letzten etwa 10 ms ausgeblendet. Die Klänge wurden auf das gleiche quadratische Mittel des RMS-Pegels normalisiert. Die Auswahl der Aufnahmen für jeden Durchgang erfolgte zufällig, jedoch mit der Einschränkung, dass in der Bedingung „gleich hoch“ zwei Aufnahmen mit gleicher Tonhöhe ausgewählt wurden und in der Bedingung „verschieden hoch“ zwei Aufnahmen mit verschiedener Tonhöhe. Die Paare bestanden immer aus Klängen von zwei verschiedenen Instrumenten.

---

<sup>12</sup> Da die Dokumentation der Anzahl der Personen, die den Vortest nicht bestanden haben, möglicherweise nicht immer instruktionsgemäß erfolgte, kann eine höhere Zahl abgewiesener Personen nicht ausgeschlossen werden.

<sup>13</sup> Um die höhere Aufgabenschwierigkeit bei ähnlichen Ziel- und Distraktorreizen zu kompensieren, wurden immer nur zwei Reize gleichzeitig präsentiert. Eine Vorstudie hatte ergeben, dass viele Teilnehmer die Suche in akustischen Szenarios mit mehr als zwei gleichzeitig gespielten Einzelnoten von Musikinstrumenten nicht mehr bewältigen können.

### 6.4.1.3. Versuchsdurchführung

Durch einen Vortest sollte sichergestellt werden, dass die Teilnehmer die in der Untersuchung vorkommenden Klänge erkennen konnten. Zunächst konnten sich alle Teilnehmer die acht Klänge beliebig oft anhören. Danach sollten die in zufälliger Reihenfolge präsentierten Klänge den Namen der Musikinstrumente (die in der Untersuchung als visuelle Hinweise verwendet wurden) per Mausklick zugeordnet werden. Der Vortest galt als bestanden, wenn in den letzten 24 von 80 Durchgängen mindestens 23 richtige Antworten erfolgten, wobei bei Nichtbestehen ein weiterer Versuch erlaubt war. Personen, die nicht 23 richtige Antworten erreichten, nahmen nicht an der eigentlichen Untersuchung teil. Wie in den vorangegangenen Untersuchungen wurde den Teilnehmern Geräusch-Szenarios (hier zwei Musikinstrumentenklänge) vorgespielt, und sie sollten bei Anwesenheit eines zuvor visuell angekündigten Instrumentes die rechte Taste und bei Abwesenheit die linke Taste einer Tastatur drücken. Richtige, falsche, zu frühe oder zu späte Reaktionen wurden 1 s lang visuell als solche rückgemeldet. Die Separation und die Tonhöhe der Klänge wurde blockweise variiert, mit zufälliger Reihenfolge der Blöcke, während die Präsentationsdauer stets 500 ms betrug. Die Bedingungen „gleich hoch“ und „verschieden hoch“ waren hinsichtlich des Instrumentes, der Tonhöhe des Zieles und der Position im Raum parallel. Die Untersuchung umfasste 200 Durchgänge und dauerte pro Person etwa 15 Minuten.

### 6.4.1.4. Versuchsplan

Für die Untersuchung der Sensitivität und der Reaktionszeiten lag ein  $2 \times 2 \times 2$ -faktorieller Versuchsplan mit den messwiederholten Faktoren Separation, Anwesenheit (des zu suchenden Geräusches) und Tonhöhe vor. Der Faktor Separation besaß die Ausprägungen „zusammen“ (alle Geräusche wurden aus einem Lautsprecher gespielt) und „getrennt“ (alle Geräusche wurden aus verschiedenen Lautsprechern gespielt), der Faktor Anwesenheit die Ausprägungen „anwesend“ und „abwesend“ und der Faktor Tonhöhe die Ausprägungen „gleich hoch“ und „verschieden hoch“.

Um die Teststärke für einen Interaktionseffekt zwischen Separation und Tonhöhe zu berechnen, wurden Differenzen zwischen den Ausprägungen des Faktors Separation (d.h.

zusammen und getrennt) für beide Stufen des Faktors Tonhöhe gebildet und die Teststärkenberechnung für eine einfaktorielles Varianzanalyse mit Messwiederholung durchgeführt. Eine A-priori-Analyse mit einer angenommenen Populationskorrelation von  $\rho = .5$  zeigte, dass eine Stichprobengröße von 84 Personen erforderlich ist, um einen Effekt (Unterschied in der Größe des Separierungseffekts für die Tonhöhen) von  $f = 0.2$  mit  $\alpha = \beta = .05$  nachweisen zu können. Eine Sensitivitätsanalyse mit einer Stichprobengröße von 108 Teilnehmern und gleichen anderen Kennwerten ergab, dass ein Effekt von  $f = 0.175$  oder größer gefunden werden konnte.

#### 6.4.2. Ergebnisse

Dem deskriptiven Erscheinungsbild nach war die Sensitivität bei räumlich getrennter Darbietung größer als bei Darbietung zusammen von einem Ort (siehe Abbildung 11, links). Auch scheint die Sensitivität bei verschiedenen hohen Klängen größer als bei gleich hohen Klängen zu sein. Dem Anschein nach war der Separationsvorteil in der Bedingung mit gleicher Tonhöhe größer als bei verschiedener Tonhöhe.

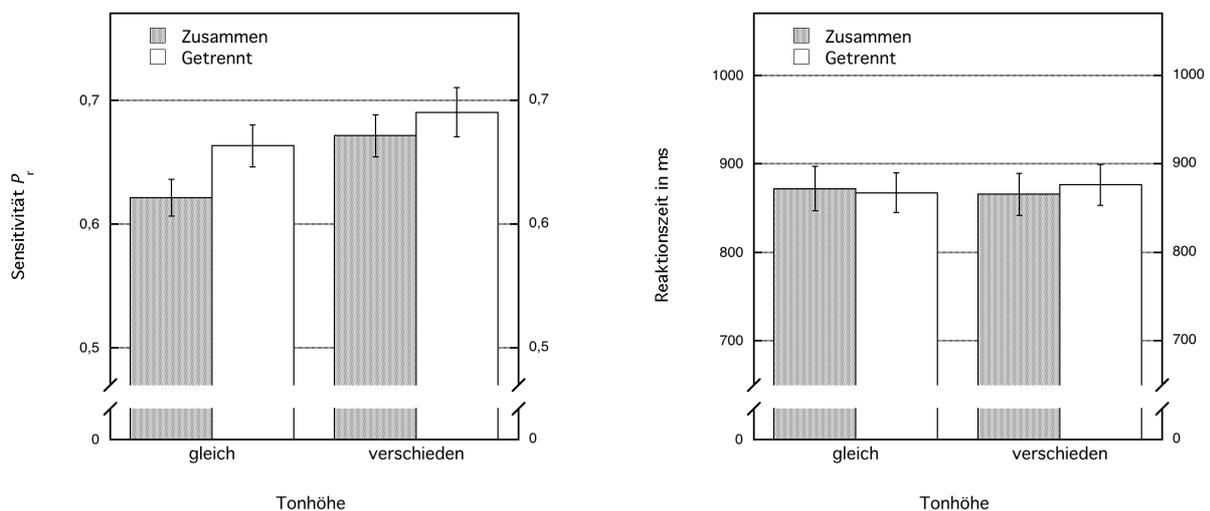


Abbildung 11: Die Sensitivität  $P_r$  (links) und die Reaktionszeiten (rechts) für Untersuchung 4 in Abhängigkeit von der Separation und der Tonhöhe der gleichzeitig präsentierten Klänge. Die Fehlerbalken zeigen Standardfehler.

Für die abhängige Variable Sensitivität zeigte eine 2 x 2-MANOVA mit den messwiederholten Faktoren Separation und Tonhöhe einen Haupteffekt der Separation,  $F(1,107) = 5.686$ ,  $p = .019$ ,  $\eta^2 = .050$  und einen Haupteffekt der Tonhöhe,  $F(1,107) = 10.184$ ,  $p = .002$ ,  $\eta^2 = .087$ . Eine Interaktion zwischen Separation und Tonhöhe konnte nicht gefunden werden ( $F(1,107) = 0.965$ ,  $p = .328$ ,  $\eta^2 = .009$ ). Um den sich deskriptiv andeutenden Unterschied in der Größe des Separationseffekts für Geräuschkombinationen gleicher und unterschiedlicher Tonhöhe zu überprüfen, wurden gerichtete Post-hoc-Tests (mit adjustiertem Alpha-niveau:  $\alpha = .05/2 = .025$ ) durchgeführt. Sie lieferten nur in der Bedingung mit gleicher Tonhöhe einen Separationseffekt,  $F(1,107) = 6.619$ ,  $p = .006$ ,  $\eta^2 = .058$ , während in der Bedingung mit verschiedener Tonhöhe kein Separationseffekt nachgewiesen werden konnte,  $F(1,107) = 1.135$ ,  $p = .145$ ,  $\eta^2 = .010$ .

Tabelle 8

Statistische Kennwerte der 2 x 3 x 4-MANOVA der abhängigen Variable Reaktionszeit für Untersuchung 4. Signifikante Effekte ( $p \leq .05$ ) sind fettgedruckt.

Effekt	Freiheitsgrade		<i>F</i>	<i>p</i>	$\eta^2$
	Zähler	Nenner			
Separation	1	107	0.080	.778	.001
Tonhöhe	1	107	0.021	.885	<.001
<b>Anwesenheit</b>	<b>1</b>	<b>107</b>	<b>390.003</b>	<b>&lt; .001</b>	<b>.785</b>
Separation x Tonhöhe	1	107	0.774	.381	.007
Separation x Anwesenheit	1	107	0.111	.739	.001
<b>Tonhöhe x Anwesenheit</b>	<b>1</b>	<b>107</b>	<b>7.105</b>	<b>.009</b>	<b>.062</b>
Separation x Tonhöhe x Anwesenheit	1	107	0.687	.409	.006

Deskriptiv waren die Reaktionszeiten in allen Bedingungen ungefähr ähnlich lang (siehe Abbildung 11, rechts). Eine 2 x 2 x 2-MANOVA mit den messwiederholten Faktoren Separation, Anzahl, Dauer und Anwesenheit zeigte keinen Haupteffekt der Separation, keinen Haupteffekt der Tonhöhe, aber einen Haupteffekt der Anwesenheit mit längeren Reak-

tionszeiten bei abwesendem Zielreiz. Außerdem trat eine Interaktion zwischen Tonhöhe und Anwesenheit auf. Zumindest tendenziell waren die Reaktionszeiten bei anwesendem Ziel in der Bedingung mit verschiedener Tonhöhe kürzer – bei abwesendem Ziel jedoch in der Bedingung mit gleicher Tonhöhe. Die statistischen Kennwerte aller Effekte der Varianzanalyse mit der abhängigen Variable Reaktionszeit zeigt Tabelle 8.

### 6.4.3. Diskussion

In der Sensitivität konnten Haupteffekte der Separation und der Tonhöhe gefunden werden, ohne dass ein Geschwindigkeits-Genauigkeitsaustausch diese relativieren würde. Obwohl die Interaktion zwischen den beiden Haupteffekten nicht signifikant wurde, konnte nur in der Bedingung mit gleicher Tonhöhe ein Separationseffekt nachgewiesen werden, während dieser Effekt in der Bedingung mit verschiedener Tonhöhe statistisch nicht belegt werden konnte. Gleichzeitig finden sich in den Reaktionszeiten schwache Tendenzen zu Gunsten der getrennten Darbietung bei gleicher Tonhöhe und zu Gunsten der Darbietung von einem Ort bei verschiedener Tonhöhe. Dass die Interaktion in der Sensitivitätsanalyse das Signifikanzniveau nicht erreicht hat, lag wahrscheinlich nicht an einem eventuellen Deckeneffekt, da ein durchschnittliches  $P_r$  von weniger als .7 in der besten von vier Bedingungen eigentlich noch Raum für Verbesserungen in der Sensitivität zulässt. Möglicherweise ist der Interaktionseffekt so klein, dass die Teststärke der vorliegenden Untersuchung (die ausgerichtet war auf eine Effektgröße von  $f = 0.175$  bei  $\beta = .95$ ) für einen Nachweis nicht ausreichte. Insofern liefert der Vergleich des Separationseffektes bei ähnlichen und unähnlichen Reizen nicht ganz eindeutig zu interpretierende Ergebnisse. Allerdings legen die Ergebnisse eher den Schluss nahe, dass räumliche Hinweise vor allem dann genutzt werden, wenn eine Aufgabe mit Hilfe von frequenzbezogenen Hinweisen nur schwer bewältigt werden kann. Im Einklang damit ist der Haupteffekt der Tonhöhe in der Sensitivität auch tendenziell größer als der Haupteffekt der Separation. Bei den hier verwendeten Ausprägungen für Separation und Tonhöhe scheint also letztere eher die größere Hilfe bei der Suche von Musikinstrumentenklängen zu sein. Das Befundmuster eines Separationsvorteils bei ähnlichen Reizen lässt sich somit eher mit den Untersuchun-

gen von Noble und Perrett (2002) und Hawley et al. (2004) vereinbaren als mit der Studie von Shackleton et al. (1994).

## 7. Allgemeine Diskussion

In diesem Abschnitt werden zunächst die Ergebnisse der eigenen Untersuchungen kurz zusammengefasst und mit der bestehenden Literatur verglichen. Daraufhin werden Interpretationen beschrieben, Bezüge zum theoretischen Hintergrund hergestellt und abschließend Schlussfolgerungen aus den Untersuchungen dargestellt.

In Untersuchung 1 wurde die akustische Suche mit visuellem Zielhinweis in einem Raum mit Reflexionen durchgeführt. Ausschließlich für sechs gleichzeitig präsentierte Geräusche und lediglich in der Sensitivität fand sich ein Vorteil räumlich separierter Schallquellen – bei vergleichbar langen Reaktionszeiten. Untersuchung 2 wurde in einem reflexionsarmen Raum durchgeführt und lieferte einen generellen Vorteil der Separation sowohl in der Sensitivität als auch in den Reaktionszeiten. Auch Untersuchung 3 wurde in einem reflexionsarmen Raum durchgeführt; jedoch wurde der Zielhinweis akustisch statt visuell dargeboten. Ein Vorteil räumlicher Separation zeigte sich nur in den Reaktionszeiten und auch nur bei sechs gleichzeitig präsentierten Geräuschen oder bei abwesendem Ziel, während die dazugehörigen Sensitivitäten sich nicht unterschieden. In Untersuchung 4 wurde schließlich die Ähnlichkeit in der Tonhöhe zwischen den Reizen variiert und (wie in den Untersuchungen 1 und 2) ein visueller Zielhinweis verwendet. Obwohl keine Interaktion zwischen Separation und Tonhöhe gefunden werden konnte, lieferten Post-hoc-Tests nur für gleiche und nicht für verschiedene Tonhöhen einen Vorteil räumlich separierter Schallquellen.

Dass Untersuchung 1 nicht für alle Anzahlen gleichzeitig präsentierter Geräusche einen Vorteil räumlich separierter Schallquellen bei der akustischen Suche liefert, widerspricht zunächst den Ergebnissen von Eramudugolla et al. (2008), die jedoch einen virtuellen Raum ohne Reflexionen oder Nachhall verwendeten. Da mehrere Untersuchungen in der Literatur eine Minderung des Separationseffektes bei Anwesenheit von Reflexionen oder komplexem Raumhall offenbarten (z.B. Culling et al., 2003; Kidd, Mason, et al., 2005; Koehnke & Besing, 1996), geht der in vielen Bedingungen von Untersuchung 1 nicht erkennbare Separationsvorteil wahrscheinlich auf die Reflexionen des Versuchsraumes zu-

rück. Ähnlich wie in der Untersuchung von Koehnke und Besing (1996) führten der Raumeinfluss in Untersuchung 1 dieser Arbeit jedoch nicht zum völligen Fehlen des Separationsvorteils; er konnte allerdings nur in der Bedingung mit sechs gleichzeitig präsentierten Geräuschen beobachtet werden.

Untersuchung 2 wurde in einem reflexionsarmen Raum durchgeführt und zeigte einen Vorteil räumlich separierter Schallquellen für alle Anzahlen gleichzeitig präsentierter Geräusche und sowohl in der Sensitivität als auch in den Reaktionszeiten. Dieses Teilergebnis ist konsistent zu Eramudugolla et al. (2008), wobei in der Untersuchung der vorliegenden Arbeit zusätzlich ein Geschwindigkeits-Genauigkeitsaustausch ausgeschlossen werden kann. Zwar lassen sich natürliche Bedingungen finden, die mehr Reflexionen als der verwendete reflexionsarme Raum aufweisen, jedoch können im Freien oft wiederum erheblich weniger Reflexionen als in dem akustisch unbehandelten Raum aus Untersuchung 1 wahrgenommen werden. Da beide Untersuchungen zumindest in einem Teil der Bedingungen Separationsvorteile nachweisen können, ist ein ähnlicher Effekt auch für viele natürliche Situationen zu erwarten.

Eramudugolla et al. (2008) berichten jedoch auch von einem größeren Separationsvorteil bei einer Präsentationsdauer der Reize von 250 ms als bei 500, 1000, 1500 oder 3000 ms. Dieses Befundmuster zeigte Untersuchung 2 nicht. Weder in der Sensitivität noch in den Reaktionszeiten konnte eine Interaktion zwischen Separation und Dauer oder eine höhere Interaktion beobachtet werden. Ebenso lieferten auch die Untersuchungen 1 und 3 keinen Hinweis auf eine Abhängigkeit des Separationseffektes von der Präsentationsdauer der Geräusche. Da Arrabito (2006) ebenso keinen Einfluss der Präsentationsdauer findet, ist eine Abhängigkeit des Separationseffektes von der Präsentationsdauer unwahrscheinlich.

Die Ausprägung des Separationseffekt im Paradigma der akustischen Suche scheint jedoch von der Art des Zielhinweises abzuhängen. Untersuchung 3 lieferte nur in der Bedingung mit sechs gleichzeitig präsentierten Geräuschen sowie bei abwesendem Ziel einen Separationsvorteil in den Reaktionszeiten. Ihlefeld und Shinn-Cunningham (2008) konnten in einer ähnlichen Bedingung ihrer Studie jedoch keinen Effekt räumlich separierter Schallquellen nachweisen, führten aber auch keine Durchgänge mit abwesendem Ziel

durch. Bei akustischen Suchaufgaben wird frequenzbezogene Information im Vergleich zu räumlicher Information anscheinend zumindest tendenziell bevorzugt genutzt.

In Untersuchung 4 wurde der Separationseffekt für ähnliche und unähnliche Reize verglichen. Auch wenn die Interaktion zwischen Separation und Tonhöhe nicht signifikant wurde, scheint ein Separationseffekt vor allem bei ähnlichen Reizen vorzufinden zu sein. Damit sind die Ergebnisse eher konsistent mit Noble und Perrett (2002) sowie Hawley et al. (2004) und stehen im Widerspruch zu Shackleton et al. (1994), deren Aufgabe jedoch geteilte Aufmerksamkeit notwendig machte. Für Suchaufgaben, die mit den in dieser Arbeit berichteten Untersuchungen vergleichbar sind, ist wahrscheinlich, dass ein Separationsvorteil bei ähnlicheren Reizen größer ausfällt.

Grundsätzlich ist eine Manipulation von Variablen zwischen verschiedenen Untersuchungen methodisch weniger günstig als eine Manipulation von Variablen innerhalb einer Untersuchung, da sich die Rahmenbedingungen für verschiedene Untersuchungen leicht unterscheiden können. Die Interpretation wird jedoch im Falle der Untersuchungen 1 bis 3 durch hohe Gleichheit in Stichprobe, Material, Durchführung und Versuchsplan zumindest erleichtert. So wäre eine mögliche Interpretation der vorliegenden Untersuchungsergebnisse, dass die Nutzung räumlicher Hinweise in Suchaufgaben mit akustischen Reizen vornehmlich dann stattfindet, wenn die Verarbeitung räumlicher Hinweise für die Bewältigung der Aufgabe besonders hilfreich ist. Reflexionen eines Raumes etwa erschwerten in Untersuchung 1 ein Profitieren von räumlich separierten (statt räumlich vereinten) Schallquellen, so dass in der Bedingung mit räumlicher Separation aller akustischen Reize nicht generell bessere Detektionsleistungen beobachtet werden konnten. Dieses Ergebnis steht im Einklang mit Befunden in der Literatur, die eine beeinträchtigte Lokalisation bei Anwesenheit von Reflexionen im Raum aufzeigen (Begault, 1992; Giguere & Abel, 1993; Rakerd & Hartmann, 1985).

Untersuchung 2 ohne Reflexionen im Raum zeigt jedoch, dass es auch Situationen gibt, in denen räumliche Hinweise die Leistung beim Detektieren von akustischen Reizen verbessern, obwohl die Aufgabe auch ohne diese Hinweise bewältigt werden kann. In komplexen Hörsituationen mit bis zu acht gleichzeitig präsentierten naturalistischen Geräuschen

zeigten sich sowohl eine bessere Sensitivität als auch kürzere Reaktionszeiten, wenn die Geräusche mit räumlicher Separation dargeboten wurden. Dies ist zumindest der Fall, wenn ein visueller Hinweis den Zielreiz kennzeichnete und die spektrotemporale Information des Zielreizes aus dem Langzeitgedächtnis abgerufen werden musste. Anscheinend sind dann räumliche Hinweise hilfreich genug, um die relativ schwierige Suche nach spektrotemporalen Hinweisen zu unterstützen.

Wurde der Zielreiz in Untersuchung 3 vor jedem Durchgang jedoch akustisch dargeboten, war die spektrotemporale Information des Zielreizes leichter verfügbar, so dass die Suche der gleichen Information in dem Gemisch der Geräusche auch leichter durchgeführt werden konnte. Dies wird auch dadurch bestätigt, dass der Separationseffekt nur in den Bedingungen mit abwesendem Zielreiz zu finden war. Anscheinend wird nur in Situationen, in denen ein Abgleich zwischen der spektrotemporalen Information des Zielgeräusches und der des Gemisches aus verschiedenen Geräuschen schwierig oder unmöglich ist, verstärkt auf die Nutzung räumlicher Hinweise zurückgegriffen.

Diese Regel scheint auch in Hörsituationen mit nur zwei gleichzeitig präsentierten Reizen zu gelten. Die Musikinstrumente in Untersuchung 4 wurden in der gleichen oder verschiedenen Tonhöhen gespielt, wobei die spektrotemporale Information im letzteren Falle wegen geringerer Überlappung leichter zugänglich gewesen sein sollte. Auch wenn das Ergebnis auch noch deutlicher hätte ausfallen können, zeigte sich ein signifikanter Vorteil räumlicher Separation nur in der Bedingung mit gleicher Tonhöhe – also der Bedingung, in der die spektrotemporale Information schwieriger zu extrahieren war. Zusammenfassend gesagt nehmen räumliche Hinweise verglichen mit spektrotemporalen Hinweisen also oft eine nachrangige Rolle bei derartigen akustischen Suchaufgaben ein. Jedoch können Menschen von räumlichen Hinweisen profitieren, sobald spektrotemporale Information alleine schwerlich zur Lösung führt – selbst wenn die Aufgabe eigentlich keine Verarbeitung von räumlichen Hinweisen erfordert.

Unterschiede zwischen Sprache und Umweltgeräuschen, wie von Gygi et al. (2004), Chen und Spence (2011) oder Wolfe et al. (2009) beschrieben, scheinen sich auf den Effekt räumlich separierter Schallquellen nicht wesentlich auszuwirken. Der für Sprache gut belegte

Vorteil räumlich getrennter Darbietung zeigte sich auch für naturalistische Geräusche in den vorliegenden Untersuchungen – und zwar unter ähnlichen Bedingungen, wie sie auch für Sprache gelten: In Situationen ohne Nachhall und ohne Klangfarbeninformation enthaltenden Hinweis ist der Separationseffekt für Sprache wie auch für naturalistische Geräusche deutlich ausgeprägt.

In Situationen mit unbekanntem Zielort ist grundsätzlich eher ein geringerer Separationsvorteil zu erwarten, da es dann auch Durchgänge gibt, an deren Anfang ein Distraktor beachtet und das Ziel noch ignoriert wird, weil die Aufmerksamkeit (infolge irrtümlichen Ratens) auf einen falschen Ort ausgerichtet wurde. Solche Durchgänge gibt es bei bekanntem Zielort (und instruktionsgemäßem Verhalten) nicht, da die Orte mit Distraktoren in der separierten Bedingung stets ignoriert werden. Nach Vorliegen der Untersuchungen dieser Arbeit lässt sich aber sagen, dass der Effekt räumlicher Separation auch in komplexen Szenarien ohne vorab auf das Ziel ausgerichtete Aufmerksamkeit gefunden werden kann. Selbst wenn acht Geräusche nur 250 ms lang präsentiert wurden (und somit eine Neuausrichtung von Aufmerksamkeit in dieser kurzen Zeit unwahrscheinlich war), ließ sich in Untersuchung 2 ein Vorteil räumlich separierter Schallquellen nachweisen.

In den Untersuchungen 1, 2, 3 und 4 konnten eher kleine Effektgrößen (partielle  $\eta^2$ ) der Separation in der Sensitivität gefunden werden: .012 (in der Bedingung mit sechs Geräuschen: .082), .184, .014 und .050 (bei gleicher Tonhöhe: .058) – ähnlich in den Reaktionszeiten: .023, .130, .020 (in der Bedingung mit sechs Geräuschen: .078 und bei abwesendem Ziel: .053) und .001. Es gehen also nur etwa bis zu 20% der Varianz in den Messwerten auf die Manipulation der räumlichen Separation zurück. Ein Grund dafür könnte darin zu suchen sein, dass nicht alle Hinweise, die zu einem Vorteil räumlich separierter Schallquellen führen können, in der hier verwendeten Konstellation genutzt wurden oder werden konnten. Hilfreiche Hinweise sind interaurale Zeit- und Pegeldifferenzen, wobei letztere nur wirksam sind, wenn der Zielreiz am weitesten lateral gelegen ist (Bronkhorst & Plomp, 1992; Hawley et al., 2004), so dass an einem Ohr ein günstigeres Zielreiz-Distraktorlautstärkeverhältnis anliegt. Ein beträchtlicher Vorteil räumlicher Separation lässt sich auch bei informatorischer Maskierung beobachten (Arbogast et al., 2002; Best et al., 2005). Diese

kann bei dem in den Untersuchungen 1 bis 3 verwendeten Reizmaterial jedoch als weniger ausgeprägt angenommen werden, da einander möglichst unähnliche Reize für diese Untersuchungen ausgewählt wurden. Hörer profitieren auch davon, wenn ihnen der Ort des Zielreizes im Vorhinein bekannt ist und sie ihre Aufmerksamkeit entsprechend ausrichten können (Arbogast & Kidd, 2000; Kidd et al., 2005; Sach et al. 2000). Da der Ort des Zielgeräusches in allen Untersuchungen dieser Arbeit jedoch bei jedem Durchgang zufällig ausgewählt wurde, war der Ort des Ziels unbekannt, so dass eine vorbereitende Ausrichtung der Aufmerksamkeit nicht möglich war. Insgesamt scheint der Effekt der räumlichen Separation in den Untersuchungen 1 bis 3 im Wesentlichen also auf interaurale Zeitdifferenzen zurückzugehen.

Diese Erklärung für den relativ kleinen Separationseffekt könnte man in einer weiteren Untersuchung überprüfen, z.B. indem man eine ähnliche Prozedur wie in Untersuchung 1 oder 2 verwendet, die Reize jedoch von virtuellen Positionen per Kopfhörer präsentiert, um interaurale Zeit- und Pegeldifferenzen unabhängig voneinander manipulieren zu können. In einer Bedingung mit Zeitdifferenzen sollte sich – wenn die oben beschriebenen Annahmen zutreffen – ein Separationseffekt finden lassen, während eine Bedingung nur mit Pegeldifferenzen keinen oder einen kleineren Separationseffekt zeigen sollte. Dies sollte zumindest dann der Fall sein, wenn alle oder die meisten Durchgänge Distraktoren als die am weitesten links und rechts präsentierten Reize enthalten und der Zielreiz somit von diesen meist räumlich eingeschlossen ist.

Eine weitere Untersuchungs idee wäre, die Hypothese zu prüfen, dass der verringerte Separationseffekt in Untersuchung 3 allein auf die akustische Modalität und nicht auf die in den verwendeten Hinweisen enthaltene spektrotemporale Information zurückgeht. Dazu könnte man gesprochene Wörter als Hinweise einsetzen, die das zu suchende Ziel benennen, aber den konkreten Klang des Ziels nicht vermitteln.

Die vorliegenden Untersuchungen liefern insgesamt Evidenz dafür, dass ein Vorteil räumlich separierter Schallquellen auch bei bis zu acht gleichzeitig präsentierten Geräuschen zu finden und nicht durch einen Geschwindigkeits-Genauigkeitsaustausch zu erklären ist. Anders als teilweise in der Literatur berichtet, ist dieser Vorteil nicht von der Präsen-

tionsdauer der Geräusche abhängig. Erfolgt die Deklaration des Ziels durch Hinweise mit Frequenzinformation, so scheinen vor allem diese spektralen Hinweise und weniger räumliche Hinweise genutzt zu werden. Der Effekt der räumlichen Separation fällt daher bei akustischem Hinweis geringer aus als bei visuellem und zeigt sich am ehesten bei abwesendem Ziel. Auch bei im Vergleich zum Ziel unähnlichen Distraktoren wird Frequenzinformation bevorzugt, solange eine Suchaufgabe dadurch leicht lösbar ist. Ähnliche Distraktoren, d.h. weniger hilfreiche Frequenzinformation, führen zumindest tendenziell zu stärkerer Verwendung von räumlichen Hinweisen und somit auch zu einem größeren Separationseffekt.

## 8. Literaturverzeichnis

- Abouchacra, K. S., Breitenbach, J., Mermagen, T., & Letowski, T. (2001). Binaural helmet: Improving speech recognition in noise with spatialized sound. *Human Factors*, 43(4), 584-594. doi: 10.1518/001872001775870368
- Arbogast, T. L., & Kidd, G., Jr. (2000). Evidence for spatial tuning in informational masking using the probe-signal method. *Journal of the Acoustical Society of America*, 108(4), 1803-1810. doi: 10.1121/1.1289366
- Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*, 112(5), 2086-2098. doi: 10.1121/1.1510141
- Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 117(4), 2169-2180. doi: 10.1121/1.1861598
- Arrabito, G. R. (2006). Three-dimensional auditory display for enhancing detection of passive sonar signals. *Human Factors*, 48(3), 465-473. doi: 10.1518/001872006778606769
- Baldis, J. J. (2001). *Effects of Spatial Audio on Memory, Comprehension, and Preference during Desktop Conferences*. Paper presented at the CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems, New York.
- Begault, D. R. (1992). Perceptual Effects of Synthetic Reverberation on 3-Dimensional Audio Systems. *Journal of the Audio Engineering Society*, 40(11), 895-904.
- Begault, D. R. (1999). Virtual acoustic displays for teleconferencing: Intelligibility advantage for "telephone-grade" audio. *Journal of the Audio Engineering Society*, 47(10), 824-828.
- Best, V., Ozmeral, E., Gallun, F. J., Sen, K., & Shinn-Cunningham, B. G. (2005). Spatial unmasking of birdsong in human listeners: energetic and informational factors. *The Journal of the Acoustical Society of America*, 118(6), 3766-3773. doi: 10.1121/1.2130949
- Blauert, J. (1997). *Spatial Hearing*. Cambridge, Massachusetts: The MIT Press.
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin: Springer.

- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. 1990. Auditory scene analysis: The perceptual organization of sound xiii, 773 pp Cambridge, MA, US: The MIT Press; US.
- Bregman, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 10-36). New York: Oxford University Press.
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acustica*, 86(1), 117-128.
- Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America*, 83(4), 1508-1516. doi: 10.1121/1.395906
- Bronkhorst, A. W., & Plomp, R. (1990). A clinical test for the assessment of binaural speech perception in noise. *Audiology*, 29(5), 275-285. doi: 10.3109/00206099009072858
- Bronkhorst, A. W., & Plomp, R. (1992). Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *Journal of the Acoustical Society of America*, 92(6), 3132-3139. doi: 10.1121/1.404209
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., . . . Ludvigsen, C. (1994). An International Comparison of Long-Term Average Speech Spectra. *Journal of the Acoustical Society of America*, 96(4), 2108-2120. doi: 10.1121/1.410152
- Chen, Y.-C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1554-1568. doi: 10.1037/a0024329
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates Inc.
- Colburn, H. S., Shinn-Cunningham, B. G., Kidd, G., & Durlach, N. (2006). The perceptual consequences of binaural hearing. *International Journal of Audiology*, 45(Supplement 1), 34-44. doi: 10.1080/14992020600782642
- Culling, J. F., & Akeroyd, M. A. (2010). Spatial Hearing. In D. R. Moore (Ed.), *The Oxford Handbook of Auditory Science: Hearing*. New York: Oxford University Press Inc.

- Culling, J. F., Hawley, M. L., & Litovsky, R. Y. (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *Journal of the Acoustical Society of America*, *116*(2), 1057-1065. doi: 10.1121/1.1925967
- Culling, J. F., Hodder, K. I., & Toh, C. Y. (2003). Effects of reverberation on perceptual segregation of competing voices. *Journal of the Acoustical Society of America*, *114*(5), 2871-2876. doi: 10.1121/1.1616922
- Darwin, C. J. (2006). Contributions of binaural information to the separation of different sound sources. *International Journal of Audiology*, *45*(Suppl1), S20-S24. doi: <http://dx.doi.org/10.1080/14992020600782592>
- Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *Journal of the Acoustical Society of America*, *114*(5), 2913-2922. doi: 10.1121/1.1616924
- Dirks, D. D., & Wilson, R. H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech and Hearing Disorders*, *12*(1), 5-38.
- Doll, T. J., & Hanna, T. E. (1995). Spatial and spectral release from masking in three-dimensional auditory displays. *Human Factors*, *37*(2), 341-355. doi: 10.1518/001872095779064573
- Drullman, R., & Bronkhorst, A. W. (2000). Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation. *Journal of the Acoustical Society of America*, *107*(4), 2224-2235. doi: 10.1121/1.428503
- Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, *113*(6), 2984-2987. doi: 10.1121/1.1570435
- Ebata, M. (2003). Spatial unmasking and attention related to the cocktail party problem. *Acoustical Science and Technology*, *24*(5), 208-219. doi: 10.1250/ast.24.208
- Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11*(5). doi: 10.1167/11.5.14

- Eramudugolla, R., McAnally, K. I., Martin, R. L., Irvine, D. R. F., & Mattingley, J. B. (2008). The role of spatial location in auditory search. *Hearing research*, 238(1-2), 139-146. doi: 10.1016/j.heares.2007.10.004
- Ericson, M. A., & McKinley, R. L. (1997). The intelligibility of multiple talkers separated spatially in noise. *Binaural and spatial hearing in real and virtual environments* (pp. 701-724). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc; England.
- Fastl, H., & Zwicker, E. (2007). *Psychoacoustics*. Berlin: Springer.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. doi: 10.3758/BF03193146
- Feddersen, W. E., Sandel, T. T., Teas, D. C., & Jeffress, L. A. (1957). Localization of high-frequency tones. *Journal of the Acoustical Society of America*, 29, 988-991. doi: 10.1121/1.1909356
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88(4), 1725-1736. doi: 10.1121/1.400247
- Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*, 106(6), 3578-3588. doi: 10.1121/1.428211
- Garadat, S. N., Litovsky, R. Y., Yu, G., & Zeng, F.-G. (2009). Role of binaural hearing in speech intelligibility and spatial release from masking using vocoded speech. *Journal of the Acoustical Society of America*, 126(5), 2522-2535. doi: 10.1121/1.3238242
- Giguere, C., & Abel, S. M. (1993). Sound Localization - Effects of Reverberation Time, Speaker Array, Stimulus Frequency, and Stimulus Rise Decay. *Journal of the Acoustical Society of America*, 94(2), 769-776. doi: 10.1121/1.408206
- Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America*, 115(3), 1252-1265. doi: 10.1121/1.1635840
- Hartmann, W. M. (1983). Localization of Sound in Rooms. *Journal of the Acoustical Society of America*, 74(5), 1380-1391. doi: 10.1121/1.390163

- Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *Journal of the Acoustical Society of America*, *115*(2), 833-843. doi: 10.1121/1.1639908
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65-70.
- Ihlefeld, A., & Shinn-Cunningham, B. G. (2008). Disentangling the effects of spatial cues on selection and formation of auditory objects. *The Journal of the Acoustical Society of America*, *124*(4), 2224-2235. doi: 10.1121/1.2973185
- Kidd, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *Journal of the Acoustical Society of America*, *118*(6), 3804-3815. doi: 10.1121/1.2109187
- Kidd, G., Mason, C. R., Brughera, A., & Hartmann, W. M. (2005). The role of reverberation in release from masking due to spatial separation of sources for speech identification. *Acta Acustica United with Acustica*, *91*(3), 526-536.
- Koehnke, J., & Besing, J. M. (1996). A procedure for testing speech intelligibility in a virtual listening environment. *Ear and Hearing*, *17*(3), 211-217. doi: 10.1097/00003446-199606000-00004
- Kulkarni, A., & Colburn, H. S. (2000). Variability in the characterization of the headphone transfer-function. *Journal of the Acoustical Society of America*, *107*(2), 1071-1074. doi: 10.1121/1.428571
- Lee, A. K. C., & Shinn-Cunningham, B. G. (2008). Effects of reverberant spatial cues on attention-dependent object formation. *JARO-Journal of the Association for Research in Otolaryngology*, *9*(1), 150-160. doi: 10.1007/S10162-007-0109-4
- MacDonald, J. A., Balakrishnan, J. D., Orosz, M. D., & Karplus, W. J. (2002). Intelligibility of speech in a virtual 3-D environment. *Human Factors*, *44*(2), 272-286. doi: 10.1518/0018720024497934
- Macpherson, E. A., & Middlebrooks, J. C. (2000). Localization of brief sounds: Effects of level and background noise. *Journal of the Acoustical Society of America*, *108*(4), 1834-1849. doi: 10.1121/1.1310196

- McAnally, K. I., & Martin, R. L. (2007). Spatial audio displays improve the detection of target messages in a continuous monitoring task. *Human Factors, 49*(4), 688-695. doi: 10.1518/001872007X215764
- Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America, 22*, 167-173. doi: 10.1121/1.1906584
- Mondor, T. A., Breau, L. M., & Milliken, B. (1998). Inhibitory processes in auditory selective attention: Evidence of location-based and frequency-based inhibition of return. *Perception & Psychophysics, 60*(2), 296-302. doi: 10.3758/BF03206038
- Mondor, T. A., & Bregman, A. S. (1994). Allocating attention to frequency regions. *Attention, Perception & Psychophysics, 56*(3), 268-276. doi: 10.3758/BF03209761
- Moore, B. C. J. (2004). *An Introduction to the Psychology of Hearing* (Fifth ed.): Academic Press.
- Morton, J., Crowder, R. G., & Prussin, H. A. (1971). Experiments with the stimulus suffix effect. *Journal of Experimental Psychology, 91*(1), 169-190.
- Nelson, W. T., Bolia, R. S., Ericson, M. A., & McKinley, R. L. (1998a). *Monitoring the Simultaneous Presentation of Multiple Spatialized Speech Signals in the Free Field*. Paper presented at the 16th International Congress on Acoustics and the 135th Meeting of the Acoustical Society of America.
- Nelson, W. T., Bolia, R. S., Ericson, M. A., & McKinley, R. L. (1998b). Monitoring the simultaneous presentation of spatialized speech signals in a virtual acoustic environment. *Proceedings of the 1998 IMAGE Conference, 159-166*.
- Nelson, W. T., Bolia, R. S., Ericson, M. A., & McKinley, R. L. (1999). *Spatial audio displays for speech communications: A comparison of free and virtual acoustic environments*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 43rd Meeting.
- Noble, W., & Perrett, S. (2002). Hearing speech against spatially separate competing speech versus competing noise. *Perception & Psychophysics, 64*(8), 1325-1336.
- Oxenham, A. J., Fligor, B. J., Mason, C. R., & Kidd, G., Jr. (2003). Informational masking and musical training. *Journal of the Acoustical Society of America, 114*(3), 1543-1549. doi: 10.1121/1.1598197

- Peissig, J., & Kollmeier, B. (1997). Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *Journal of the Acoustical Society of America*, 101(3), 1660-1670. doi: 10.1121/1.418150
- Rakerd, B., & Hartmann, W. M. (1985). Localization of sound in rooms, II: The effects of a single reflecting surface. *The Journal of the Acoustical Society of America*, 78(2), 524-533.
- Richards, V. M., & Kidd, G., Jr. (2009). Auditory masking with complex stimuli. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (Fourth ed., pp. 343-352). Cambridge, MA: Massachusetts Institute of Technology.
- Ruggles, D., & Shinn-Cunningham, B. G. (2011). Spatial selective auditory attention in the presence of reverberant energy: individual differences in normal-hearing listeners. *Jaro-Journal of the Association for Research in Otolaryngology*, 12(3), 395-405. doi: 10.1007/s10162-010-0254-z
- Sach, A. J., Hill, N. I., & Bailey, P. J. (2000). Auditory spatial attention using interaural time differences. *Journal of experimental psychology Human perception and performance*, 26(2), 717-729. doi: 10.1037//0096-1523.26.2.717
- Scharf, B. (1998). Auditory attention: The psychoacoustical approach. In H. Pashler, *Attention* (pp. 75-117). Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.
- Shackleton, T. M., Meddis, R., & Hewitt, M. J. (1994). The Role of Binaural and Fundamental-Frequency Difference Cues in the Identification of Concurrently Presented Vowels. *Quarterly Journal of Experimental Psychology - Human Experimental Psychology*, 47(3), 545-563. doi: 10.1080/14640749408401127
- Shinn-Cunningham, B. G. (2005). *Influences of spatial cues on grouping and understanding sound*. Paper presented at the Proceedings of Forum Acusticum, Budapest.
- Shinn-Cunningham, B. G., Santarelli, S., & Kopco, N. (2000). Tori of confusion: Binaural localization cues for sources within reach of a listener. *Journal of the Acoustical Society of America*, 107(3), 1627-1636. doi: 10.1121/1.428447
- Slattery, W. H., 3rd, & Middlebrooks, J. C. (1994). Monaural sound localization: acute versus chronic unilateral impairment. *Hearing Research*, 75(1-2), 38-46. doi: 10.1016/0378-5955(94)90053-1

- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of Measuring Recognition Memory - Applications to Dementia and Amnesia. *Journal of Experimental Psychology: General*, 117(1), 34-50.
- Spence, C. J., & Driver, J. (1994). Covert Spatial Orienting in Audition - Exogenous and Endogenous Mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, 20(3), 555-574. doi: 10.1037/0096-1523.20.3.555
- Stern, R. M., & Trahiotis, C. (1995). Models of binaural interaction. *Hearing* (pp. 347-386). San Diego, CA: Academic Press; US.
- Wightman, F. L., & Kistler, D. J. (1997). Monaural sound localization revisited. *Journal of the Acoustical Society of America*, 101(2), 1050-1063. doi: 10.1121/1.418029
- Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R. S., Lederman, S. J., & Merfeld, D. M. (2009). *Sensation and Perception*. Sunderland, Massachusetts U.S.A.: Sinauer Associates, Inc.
- Yost, W. A., Dye, R. H., Jr., & Sheft, S. (1996). A simulated "cocktail party" with up to three sound sources. *Attention, Perception, & Psychophysics*, 58(7), 1026-1036. doi: 10.3758/BF03206830
- Zhang, P. X., & Hartmann, W. M. (2010). On the ability of human listeners to distinguish between front and back. *Hearing Research*, 260(1-2), 30-46. doi: 10.1016/j.heares.2009.11.001

Hiermit erkläre ich, dass ich die vorliegende Arbeit eigenständig und ohne unerlaubte Hilfe angefertigt habe. Die Dissertation wurde in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen Institution eingereicht. Außerdem habe ich bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 02.10.2012

(Gunnar Regenbrecht)