



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Der Einfluss der Ähnlichkeit zwischen Hinweis- und Zielreiz in einem auditiven Hinweisreizparadigma“

verfasst von / submitted by

Moritz Klingel, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2017

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 840

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Psychologie UG2002

Betreut von / Supervisor:

Doz. Dr. Bernhard Laback

Danksagung

Ich danke meinem Betreuer Doz. Dr. Bernhard Laback für die Offenheit meinen Ideen gegenüber und die konsequente Unterstützung von der Planung über die Durchführung bis zum Abschluss der Arbeit.

Weiter danke ich Prof. Ulrich Ansorge für seine hilfreichen Anregungen und sein detailliertes Feedback.

Daneben danke ich dem Institut für Schallforschung, welches mich bereitwillig aufgenommen und durch seine angenehme Arbeitsatmosphäre überzeugt hat. Besonderer Dank gilt hier Michael Mihocic, der durch seine Programmierkünste die Umsetzung des Experiments erst ermöglichte.

Außerdem danke ich meiner Freundin, die mich nicht nur durch Korrekturlesen und angeregte Diskussionen unterstützt, sondern auch meine Stimmungsschwankungen ertragen und immer wieder aufs Neue motiviert hat.

Inhaltsverzeichnis

Abstract	3
Zusammenfassung	4
Theoretischer Hintergrund	5
Auditive Reize im Alltag	5
Selektive Aufmerksamkeit und das Hinweisreizparadigma	5
Unterschiede zwischen visueller und auditiver Modalität	7
Das Hinweisparadigma in der auditiven Modalität	8
Die selektive Aufmerksamkeit in der auditiven Modalität	9
Experiment 1	10
Anforderungen an das neue experimentelle Design	10
Ablauf und erwartete Ergebnisse	11
Methode	11
Versuchspersonen	11
Apparatur und Reize	11
Aufgabe	12
Prozedur	13
Ergebnisse	14
Reaktionszeiten	15
Trefferquoten	17
Diskussion	17
Experiment 2	18
Beschreibung und erwartete Ergebnisse	18
Methode	19
Versuchspersonen	19
Apparatur und Reize	19
Aufgabe und Prozedur	20
Ergebnisse	21
Reaktionszeiten	21
Trefferquoten	22
Generelle Diskussion	24
Erklärungen innerhalb der auditiven Modalität	25
Vergleich zwischen auditiver und visueller Modalität	27
Limitationen und weiterführende Forschung	28

Schlussfolgerung	30
Literaturverzeichnis.....	31
Abbildungsverzeichnis	35
Tabellenverzeichnis.....	35
Abkürzungsverzeichnis	35
Anhang	36

Abstract

Many studies in the visual domain used the cueing paradigm to study mechanisms of selective attention. We developed a new experimental paradigm to show similar mechanisms of selective attention in the auditory domain. In Experiment 1, participants ($n = 9$) had to selectively focus on a target speaker with predefined pitch, while one distractor speaker with different pitch was presented simultaneously. Target and distractor were spatially separated along the lateral dimension and were preceded by a speech mixture cue either at the same (congruent) or at a different (incongruent) lateral position relative to the target. Surprisingly, reaction times were shorter in incongruent than in congruent trials, irrespective if they had to react to the higher (group 1) or to the lower (group 2) pitch speaker. Former studies in the visual or auditory domain show, that the similarity between cue and target influences the direction of the spatial cueing effect. Although our cue consisted of mixed segments of the higher voice, group 1 and 2 did not differ in their performance. Therefore, we supposed that cue and target differed too much resulting in a negative spatial cueing effect. In Experiment 2 ($n = 19$), we varied the similarity between cue and target by testing three different cue types. Consistent with our hypotheses, a high similarity between cue and target resulted in a positive and a low similarity in a negative spatial cueing effect. As in the visual domain, the results are interpreted in terms of location benefits and costs depending on the cue-target similarity.

Zusammenfassung

Das Hinweisreizparadigma wurde in einer Vielzahl von Studien in der visuellen Modalität dazu genutzt, die Mechanismen der selektiven Aufmerksamkeit zu untersuchen. In der vorliegenden Studie wurde ein neues experimentelles Paradigma entwickelt, mit welchem die in der visuellen Modalität untersuchten Effekte auch in der auditiven Modalität sichtbar gemacht werden können. In Experiment 1 mussten die Versuchspersonen (Vpn) auf eine von zwei gleichzeitig präsentierten Stimmen reagieren, welche sich in ihrer Tonhöhe unterschieden. Beide Stimmen waren räumlich voneinander getrennt und wurden nach einem Sprachgemisch-Hinweisreiz präsentiert, welcher entweder an der gleichen lateralen Position der Zielstimme (kongruent) oder an der der Distraktorstimme (inkongruent) erklang. Überraschenderweise reagierten die Vpn ($n = 9$) in inkongruenten Durchgängen schneller auf die Zielstimme als in kongruenten Durchgängen, unabhängig davon ob sie auf die höhere (Gruppe H) oder tiefere (Gruppe T) Stimme reagieren mussten. Aus früheren Studien in der visuellen und auditiven Modalität ist bekannt, dass die Ähnlichkeit zwischen Hinweis- und Zielreiz einen Einfluss auf die Art des räumlichen Kongruenzeffektes hat. Da der Hinweisreiz aus übereinander gelagerten Sprachsegmenten der höheren Stimme erzeugt worden war und sich die Leistung der Vpn in beiden Gruppen nicht unterschied, wurde vermutet, dass der Hinweisreiz dem Zielreiz zu unähnlich war und sich deshalb ein negativer räumlicher Kongruenzeffekt (NRK) zeigte. In Experiment 2 ($n = 19$) wurde die Hinweis-Zielreiz-Ähnlichkeit (HZÄ) mittels dreier Hinweisreiztypen systematisch variiert. Entsprechend der Hypothesen führte eine hohe HZÄ zu einem positiven und eine geringe HZÄ zu einem NRK. Ähnlich wie in der visuellen Modalität können die Ergebnisse als Erleichterungs- und Kosteneffekte, welche von der HZÄ abhängen, interpretiert werden.

Theoretischer Hintergrund

Auditive Reize im Alltag

In unserem Alltag sind wir einer Vielzahl von gleichzeitig auf uns eindringenden auditiven Reizen ausgesetzt. Dabei ist es in manchen Situationen wichtig, dass wir trotz laufendem Autoradio die Sirene eines heranfahrenden Krankenwagens bemerken, um ihm ausweichen zu können, und wiederum in anderen Situation von Vorteil, dass sich eine Orchestermusikerin auf ihr eigenes Instrument konzentrieren kann, obwohl gleichzeitig laute Paukenschläge erklingen. Ein ähnliches Beispiel bietet das von Cherry (1953) beschriebene Cocktail-Party-Problem. Stellen Sie sich folgende Situation vor: Sie befinden sich auf einer Party mit mehreren Gästen und Musik ertönt aus den Lautsprechern eines Raumes. Sie unterhalten sich mit einer Freundin und neben Ihnen unterhalten sich zwei Ihrer Freunde. Obwohl in diesem Moment mehrere Schallsignale Ihre Ohren erreichen, ist es Ihnen möglich, den Worten Ihrer Gesprächspartnerin zu folgen. Diese scheinbar banale Fähigkeit erfordert eine Reihe komplexer Mechanismen. Zunächst muss Ihnen klar sein, wie die Stimme Ihrer Freundin klingt. Sie müssen sich auf die Merkmale wie Tonhöhe, Klangfarbe und Sprachmelodie sowie die räumliche Position der Stimme einstellen. Mit diesem Wissen können Sie dann Ihre Aufmerksamkeit gezielt auf das Schallsignal Ihrer Gesprächspartnerin richten. Allerdings müssen Sie es dazu aus einer Flut von interferierenden Schallsignalen herausfiltern (für ein Review des Cocktail-Party-Problems siehe Haykin & Chen, 2005). Die folgende Forschungsarbeit leistet einen Beitrag, die hierbei involvierten Prozesse besser zu verstehen. Konkret geht es um die Frage, wie die räumliche Kongruenz eines kurz zuvor präsentierten akustischen Reizes die Aufmerksamkeit auf einen Zielreiz beeinflusst und welche Reizeigenschaften unsere Aufmerksamkeit unwillkürlich beeinflussen.

Selektive Aufmerksamkeit und das Hinweisreizparadigma

Selektive Aufmerksamkeit ist ein Kennzeichen jedes Wahrnehmungsprozesses. Bestimmte Objekte oder Merkmale von Objekten werden dabei fokussiert und andere ausgeblendet (Ansorge & Leder, 2017). Unter welchen Bedingungen die selektive Aufmerksamkeit erfolgreich auf ein interessierendes Objekt gelenkt werden kann, ist der Gegenstand einer breiten Forschungsdebatte. Im Mittelpunkt steht hier die Frage, inwieweit die Aufmerksamkeit eher reizgetrieben (engl. bottom-up) oder eher absichtsgetrieben (engl. top-down) gesteuert wird. Posner (1980), einer der Pioniere in diesem Gebiet, zeigte mithilfe des von ihm entwickelten Hinweisreizparadigmas, dass sich einerseits die räumliche Aufmerksamkeit unabhängig von der Blickrichtung verlagern lässt und andererseits diese räumliche Aufmerksamkeitsverlagerung unter bestimmten Umständen unwillkürlich stattfindet. Im Hinweisreizpara-

digma wird der unwillkürliche Einfluss eines präsentierten visuellen Hinweisreizes auf einen nachfolgenden visuellen Zielreiz untersucht (Posner, 1980). Entscheidend ist hierbei, welche Reaktion auf den Zielreiz gezeigt werden soll. Je nach den Eigenschaften der Reize oder der Art der Aufgabe lassen sich Rückschlüsse auf zugrundeliegende Mechanismen der Aufmerksamkeit ziehen. Beispielsweise zeigte sich in mehreren Studien in der visuellen Modalität, dass plötzlich auftretende periphere Hinweisreize (engl. *exogenous onset cues*), die zufällig entweder an der gleichen oder an einer anderen räumlichen Position wie der Zielreiz präsentiert wurden, die Reaktionszeiten der Versuchspersonen (Vpn) systematisch beeinflussten. Die Vpn berichteten dabei beispielsweise die Position oder den Inhalt des Zielreizes (für ein Review siehe Carrasco, 2011). Stimmt die Positionen von Hinweis- und Zielreiz überein, konnten die Vpn die Aufgabe schneller lösen. Dies wurde so interpretiert, dass die räumliche Aufmerksamkeit durch den Hinweisreiz bereits auf die Position gerichtet wird, an welcher in weiterer Folge der Zielreiz erscheint, und somit die Verarbeitung des Zielreizes erleichtert wird (z. B. Posner & Cohen, 1984). Stimmt die Positionen von Hinweis- und Zielreiz jedoch nicht überein, führte dies entsprechend zu Kosten in der Verarbeitung des Zielreizes, da die räumliche Aufmerksamkeit erst auf die Position des Zielreizes gerichtet werden musste. Diese Kosten zeigten sich auch im Vergleich zu Bedingungen in denen kein Hinweisreiz gezeigt wurde, weshalb zumindest einige ForscherInnen davon ausgehen, dass unabhängig vom informativen Wert eines Reizes, die Aufmerksamkeit immer zunächst von diesem angezogen wird (für ein Review siehe Theeuwes, 2010). Im Gegensatz dazu zeigten Folk, Remington und Johnston (1992), mit dem Phänomen der bedingten Aufmerksamkeitsanziehung (engl. *contingent capture*), dass der Einfluss eines Hinweisreizes auf die Verarbeitung eines Zielreizes auch von den kognitiven Absichten einer Person und somit von der zu bearbeitenden Aufgabenstellung abhängt. In der Studie von Folk et al. mussten die Vpn entweder nach einem plötzlich auftretenden peripheren Zielreiz oder nach einem durch eine bestimmte Farbe definierten Zielreiz suchen. Die Verarbeitung des Zielreizes wurde durch einen plötzlich auftretenden peripheren Hinweisreiz nur dann gestört, wenn die Vpn auch nach einem plötzlich auftretenden peripheren Zielreiz gesucht hatten. Der Hinweisreiz beeinflusste die Verarbeitung des Zielreizes jedoch nicht, wenn die Vpn nach einem durch seine Farbe bestimmten Zielreiz suchen mussten. Unter bestimmten Umständen ist es Personen also möglich, aufgaben-irrelevante visuelle Reize auszublenden. Unter welchen Bedingungen relevante von irrelevanten Reizen getrennt werden können, wird weiterhin ausgiebig untersucht. Da es in der vorliegenden Studie jedoch hauptsächlich um die Rolle der Aufmerksamkeit bei der Verarbei-

tung von auditiven Reizen geht, wird in den nächsten Absätzen versucht, die gerade gezogenen Schlussfolgerung auf die auditive Modalität zu übertragen.

Unterschiede zwischen visueller und auditiver Modalität

Aufgrund der fundamentalen Unterschiede in der Verarbeitung von visueller und auditiver Information, liegt es nahe, auch von Unterschieden in den zugrundeliegenden Aufmerksamkeitsmechanismen auszugehen. Während in der visuellen Wahrnehmung der räumlichen Selektivität der Reize eine bedeutende Rolle zukommt, nimmt die Verarbeitung der räumlichen Position in der auditiven Modalität laut zum Beispiel Dalton und Hughes (2014) eine untergeordnete Funktion ein. Sie verwiesen in ihrer Studie auf Kubovy (1981), der die zeitliche Erstreckung und den Unterschied in der spektral-temporalen Signalstruktur für die Differenzierung auditiver Reize als wesentlicher erachtet. Allerdings relativierte Kubovy selbst seine Auffassung in einer späteren Studie und verwies auf die Gemeinsamkeiten von auditiven und visuellen Objekten, die sich jeweils zeitlich und räumlich formen (Kubovy & Valkenburg, 2001). Laut Bregman (1990) erfolgt die auditive Selektion durch die Einteilung einer wahrgenommenen auditiven Szene in sogenannte auditive „Streams“, auf welche in weiterer Folge die Aufmerksamkeit gelenkt werden kann. Dabei profitiert die auditive Wahrnehmung von einer räumlichen Separation dieser auditiven „Streams“. Beispielsweise fällt es Vpn schwerer sich auf eine von zwei Stimmen zu fokussieren, wenn diese von der gleichen räumlichen Position ertönen, als wenn sie räumlich voneinander getrennt erklingen (Best, Gallun, Ihlefeld, & Shinn-Cunningham, 2006). Auch Shinn-Cunningham (2008) beschreibt die Formierung von Objekten sowohl in der visuellen als auch in der auditiven Modalität als wesentliche Voraussetzung für die selektive Aufmerksamkeit. Nur wenn die gewünschten Objekte aus einer komplexen Wahrnehmungsszene extrahiert werden können, kann die absichtsgesteuerte Aufmerksamkeit darauf gerichtet werden. Diese Formierung geschieht anhand von Objektmerkmalen wie räumliche Position, Tonhöhe, Klangfarbe und gelernte Bedeutung. Wird die Aufmerksamkeit absichtsvoll von einem auditiven Objekt oder „Stream“¹ auf ein anderes auditives Objekt verlagert, resultiert daraus eine Verzögerung der Verarbeitung des nun beachteten Objektes (Koch, Lawo, Fels, & Vorländer, 2011; Lawo, Fels, Oberem, & Koch, 2014). Ähnlich wie in der visuellen Modalität benötigt die räumliche Aufmerksamkeitsverlagerung also auch in der auditiven Modalität eine gewisse Zeit und profitiert von

¹ Obwohl die Definition von auditiven Objekten unklar erscheint und diese eventuell besser als auditive „Streams“ verstanden werden sollten, verwenden wir in Übereinstimmung mit Shinn-Cunningham (2008) in der vorliegenden Arbeit den Begriff auditive Objekte, um sie einfacher mit visuellen Objekten vergleichen zu können.

einer Voreinstellung der räumlichen Aufmerksamkeit auf die Position eines auditiven Zielreizes. Im nächsten Absatz wird die unwillkürliche Beeinflussung der Aufmerksamkeit durch auditive Reize diskutiert.

Das Hinweisparadigma in der auditiven Modalität

Trotz der weiter oben erwähnten Unterschiede in der Wahrnehmung von visuellen und auditiven Reizen, wurde das Hinweisreizparadigma auf die auditive Modalität übertragen (Mondor & Amirault, 1998; Mondor, Breau, & Milliken, 1998; Mondor, Zatorre, & Terrio, 1998; Roberts, Summerfield, & Hall, 2009; Teshiba et al., 2012). Beispielsweise präsentierten Mondor und Amirault (1998) ihren Vpn in mehreren Durchgängen zwei aufeinanderfolgende Schallsignale, welche zufällig entweder von der gleichen räumlichen Position oder von verschiedenen räumlichen Positionen ertönten. Der Hinweisreiz bestand aus einem Breitbandrauschsignal (engl. white noise burst) und der Zielreiz aus einem von zwei möglichen Glissandotönen, deren Frequenz entweder von 1000 zu 10000 Hz oder von 10000 zu 1000 Hz reichte. Nach jedem Durchgang mussten die Vpn angeben, ob der Zielreiz ein Glissando nach oben oder nach unten beinhaltete. Genauso wie in der visuellen Modalität profitierte die Verarbeitung des Zielreizes in den Bedingungen, in denen beide Schallsignale von der gleichen räumlichen Position ertönten. Entgegen der Behauptung, dass die räumliche Dimension auditiver Reize nur bedeutsam wird, wenn sie für die Bearbeitung einer Aufgabe relevant ist (McDonald & Ward, 1999; Spence & Driver, 1994), spricht auch die Studie von Roberts et al. (2009). In einer Reihe von Experimenten zeigten Roberts et al., dass die verdeckte räumliche Aufmerksamkeitsverlagerung auch in nicht-räumlichen Aufgabenstellungen für die Verarbeitung auditiver Reize relevant ist.

Wie sich die auditive Präsentation eines Hinweisreizes auf die Verarbeitung eines Zielreizes auswirkt, hängt allerdings von mehreren Faktoren ab. Beispielsweise zeigten Mondor, Breau et al. (1998), dass sich der positive räumliche Kongruenzeffekt (PRK) bei einer Verlängerung des Hinweis-Zielreiz-Intervalls (engl. stimulus onset asynchrony, SOA) in einen negativen räumlichen Kongruenzeffekt (NRK) umgekehrt. Das heißt, der gleiche Hinweisreiz erleichterte in Bedingungen, in denen beide Reize von derselben räumlichen Position erklingen einerseits die Verarbeitung des Zielreizes bei einer kurzen SOA und verzögerte diese andererseits bei einer längeren SOA. Mit Referenz auf einen ähnlichen Effekt in der visuellen Modalität (engl. Inhibition of Return, IOR; für ein Review siehe Klein, 2000), wurde dieses Ergebnis als räumliche Hemmung der zuvor beachteten Position interpretiert. Außerdem variierten Mondor, Breau et al. in einer anderen Experimentalbedingung die Ähnlichkeit zwischen Hinweis- und Zielreiz. In mehreren Durchgängen wurden den Vpn zwei aufei-

einanderfolgende Töne zufällig entweder mit der gleichen oder mit einer unterschiedlichen Frequenz dargeboten, wobei beide von der gleichen räumlichen Position ertönt. Die Vpn konnten schneller angeben, ob es sich bei dem Zielreizton um einen hohen oder tiefen Ton handelte, wenn Hinweis- und Zielreizton die gleiche Frequenz aufwiesen. Nicht zuletzt spielt auch die Darbietung von Distraktoren zusätzlich zum Zielreiz gerade in Bezug auf eine realistischere Umweltsituation eine entscheidende Rolle für auditive Aufmerksamkeitsprozesse (Eramudugolla, McAnally, Martin, Irvine, & Mattingley, 2008).

Die selektive Aufmerksamkeit in der auditiven Modalität

Obwohl bei der Erforschung der selektiven Aufmerksamkeit mithilfe des Hinweisreizparadigmas in der visuellen Modalität sowohl Hinweis- als auch Zielreize fast ausschließlich in Kombination mit zusätzlichen Reizen dargeboten werden, verwenden Studien in der auditiven Modalität häufig nur isolierte Reize. Dies liegt zum einen sicherlich daran, dass die Erstellung auditiver Szenen mit mehreren Reizen wesentlich komplexer ist als in der visuellen Modalität und zum anderen daran, dass viele ForscherInnen davon ausgehen, die Relevanz der räumlichen Information wäre in komplexen auditiven Szenen zweitrangig (z. B. Kubovy, 1981). Laut Dalton und Lavie (2004) ist daher die zeitlich rasch aufeinanderfolgende Präsentation von Tönen besser geeignet als die räumliche Anordnung simultan dargebotener Reize, um Prozesse der selektiven Aufmerksamkeit in einer auditiven Suchaufgabe zu untersuchen. In der Studie von Dalton und Lavie mussten die Vpn aus einer Sequenz von 5-7 Tönen, die entweder die gleiche Frequenz, Intensität oder Dauer aufwiesen, auf einen Zielreiz reagieren, der sich in einer der drei genannten Dimensionen von den anderen Reizen unterschied. Gleichzeitig variierte einer der Töne zufällig auf einer aufgaben-irrelevanten Dimension. In einer Reihe von Experimenten konnten sie zeigen, dass diese Distraktoren die Verarbeitung des Zielreizes verzögerten (Dalton & Lavie, 2004). Allerdings konnte dieses Phänomen der unwillkürlichen Aufmerksamkeitsanziehung (engl. *attentional capture*) durch eine Anpassung der Suchstrategie unterdrückt werden (Dalton & Lavie, 2007). In Übereinstimmung mit Folk et al. (1992) wird die selektive Aufmerksamkeit also auch in der auditiven Modalität sowohl von reizgetriebenen als auch von absichtsgetriebenen Prozessen gesteuert. Aufgrund der weiter oben diskutierten Studien erscheint es jedoch plausibel, dass derartige Effekte auch bei der räumlichen Trennung von simultanen auditiven Reizen auftreten. Zur Untersuchung solcher Effekte, wurde für die vorliegende Studie ein neuartiges experimentelles Verfahren entwickelt.

Experiment 1

Anforderungen an das experimentelle Design

Das experimentelle Design sollte folgenden Anforderungen gerecht werden. Zunächst nutzen die meisten früheren Studien in der auditiven Modalität eher künstliche Signale wie Sinustöne oder Rauschsignale (Mondor & Amirault, 1998; Mondor, Breau et al., 1998; Mondor, Zatorre et al., 1998; Roberts et al., 2009; Teshiba et al., 2012). Um die Generalisierbarkeit der Effekte zu überprüfen, wurden in der vorliegenden Studie Sprachsignale verwendet. Einerseits kommen Sprachsignale in der Realität häufig vor und andererseits lassen sich damit zum Cocktail-Party-Problem ähnlichere Aufgabenstellungen kreieren. Die Aufgabenstellung bestand aus einer Sprachdiskriminationsaufgabe ähnlich wie in den Studien von Koch et al. (2011) und Lawo et al. (2014). Die Vpn mussten in diesen Studien auf eine von zwei gleichzeitig präsentierten Stimmen unterschiedlichen Geschlechts achten und angeben, ob die geäußerte Zahl größer oder kleiner fünf war. Dieses Antwortformat hat den weiteren Vorteil, dass die Antworteingabe unabhängig von der räumlichen Anordnung der auditiven Reize erfolgt und somit Antwort-Priming weitgehend ausgeschlossen werden kann (Kunde, Kiesel, & Hoffmann, 2003; Norris & Kinoshita, 2008). Im Unterschied zu den oben erwähnten Studien benutzten wir in unserem Design jedoch keinen visuellen Hinweisreiz, der entweder die Position der Stimme oder das Geschlecht der Stimme angab (vgl. Koch et al., 2011; Lawo et al., 2014), sondern hielten die zu beachtende Stimme im Vorhinein über alle Durchgänge hinweg konstant. Durch diese Vorgehensweise konnten wir den unwillkürlichen Einfluss eines auditiven Hinweisreizes auf die Verarbeitung des Zielreizes untersuchen und mögliche cross-modale Interaktionen weitgehend ausschließen. Als weitere Modifikation nutzen wir in unserer Studie nicht eine männliche und eine weibliche Stimme, sondern zwei weibliche Stimmen, die sich in ihrer Grundfrequenz nur zwei Halbtöne voneinander unterschieden und räumlich entlang der horizontalen Ebene voneinander getrennt waren. Die damit verbundene größere Ähnlichkeit zwischen Ziel- und Distraktorreizen sollte die Aufgabenstellung erschweren und die Verwendung räumlicher Hinweisreize forcieren. Den Hinweisreiz erzeugten wir aus einer Überlagerung zufällig ausgewählter Sprachsegmente aus dem gesamten Sprachmaterial der höheren Stimme, sodass er einen ähnlichen Tonhöhenbereich wie die Stimme der höheren Sprecherin abdeckte. Aus früheren Studien sowohl in der visuellen als auch der auditiven Modalität ist bekannt, dass eine höhere Ähnlichkeit zwischen Hinweis- und Zielreiz die Effektivität von Hinweisreizen erhöht (z. B. Folk & Remington, 1998, in der visuellen Modalität; Mondor, Breau et al., 1998; Mondor, Zatorre et al., 1998, in der auditiven Modalität). Die Vpn wurden zufällig in zwei Gruppen eingeteilt, wobei eine Gruppe auf die höhere

(Gruppe H) und die andere Gruppe auf die tiefere (Gruppe T) Stimme achten sollte. Damit sollte die Ähnlichkeit zwischen Hinweis- und Zielreiz in der Gruppe, die auf die höhere Stimme achten musste, höher sein, als in der Gruppe, die auf die tiefere Stimme achten musste.

Ablauf und erwartete Ergebnisse

In einer Vielzahl von Durchgängen mussten die Vpn angeben, ob die von der höheren beziehungsweise tieferen Stimme geäußerte Zahl größer oder kleiner fünf war. Vor dem Zahlenpaar der beiden Stimmen wurde den Vpn ein uninformativer auditiver Hinweisreiz entweder an der gleichen räumlichen Position wie der Zielsprecherin (kongruent) oder an der gegenüberliegenden und somit gleichen räumlichen Position wie der Distraktorsprecherin (inkongruent) präsentiert. Es wurde vermutet, dass die Vpn der Gruppe H aufgrund der größeren Ähnlichkeiten zwischen Hinweis- und Zielreiz in kongruenten Bedingungen von der unwillkürlichen Voreinstellung der räumlichen Aufmerksamkeit profitieren und daher in kongruenten Durchgängen schnellere Reaktionszeiten aufweisen als in inkongruenten. Je nachdem wie bedeutsam der Unterschied in der Hinweis-Zielreiz-Ähnlichkeit (HZÄ) zwischen Gruppe H und T ist, sollten die Vpn der Gruppe T weniger stark von der räumlichen Übereinstimmung von Hinweis- und Zielreiz profitieren. Das heißt, die Reaktionszeiten in den kongruenten Durchgängen sollten insgesamt schneller sein als in den inkongruenten Durchgängen, im Vergleich zu Gruppe H sollte dieser Unterschied jedoch geringer ausfallen. Jeweils ähnliche Ergebnismuster sollten sich in den Trefferquoten widerspiegeln.

Methode

Versuchspersonen. Es wurden 9 Vpn (5 weibliche) im Labor des Instituts für Schallforschung der Österreichischen Akademie der Wissenschaften in Wien getestet, wobei die Altersspanne von 22 bis 39 reichte ($M = 25.7$; $SD = 5.2$). Alle Vpn wiesen eine normale Hörfähigkeit und normale oder dazu korrigierte Sehfähigkeit auf. Eine der weiblichen Vpn erhielt zunächst eine falsche Instruktion und hatte dadurch im folgenden Experiment anfänglich deutlich größere Schwierigkeiten als alle anderen Vpn. Diese Auffälligkeit wird in der statistischen Analyse noch weiter diskutiert.

Apparatur und Reize. Das Experiment wurde in einer doppelwandigen, schallisolierenden Kammer durchgeführt. Die Vpn wurden auf einen bequemen Stuhl vor einen Monitor gesetzt, wobei sie den Abstand zum Bildschirm frei wählen durften. Die Präsentation der auditiven Reize erfolgte durch einen zirkumauralen Kopfhörer (Sennheiser HD 518). Zur Bearbeitung der Aufgabe hielten die Probanden ein Gamepad (Logitech WingMan Action Game Pad) in der Hand. Die auditiven Zielreize bestanden aus gesprochenen Zahlwörtern von eins

bis neun ohne die fünf. Dazu wurde das bereits aufgenommene Sprachmaterial einer der weiblichen Sprecherinnen von Fels, Oberem und Koch (2016) verwendet. Für unsere Zwecke wurden die gesprochenen Zahlwörter in ihrer Tonhöhe modifiziert, um sowohl Ziel- als auch Distraktorreize zu erhalten, die sich um zwei Halbtonschritte voneinander unterscheiden. Diese Festlegung auf einen Tonhöhenunterschied von zwei Halbtonschritten erfolgte nach umfangreichen Vortests, mit dem Ziel, eine Erkennungsrate zwischen 80% bis 90% zu erzielen. Die Tonhöhenmodifikation wurde mit der Funktion „Tonhöhe ändern“ in der Freeware Audacity 2.1.0 durchgeführt, wobei das Stimmmaterial jeweils um einen Halbton nach unten und einen Halbton nach oben verändert wurde. Zusätzlich erzeugten wir einen Hinweisreiz, der mittels Überlagerung von Sprachsegmente, die mit zufällig gewählter zeitlicher Position aus allen möglichen Zahlwörtern der höheren Stimme erzeugt wurden (d. h. der Hinweisreiz sollte der höheren Stimme ähnlicher sein als der tieferen). Der Hinweisreiz wurde räumlich lateralisiert entweder von links oder rechts dargeboten. Dies wurde durch eine gezielte Manipulation der beiden binauralen Orientierungsreize (interaurale Laufzeitdifferenz; ITD = +/-300 μ s und interauraler Pegelunterschied; ILD = +/-5 dB SPL) gewährleistet². Die genannten Werte von ITD und ILD entsprachen einer räumlichen Auslenkung von ca. 45° entweder nach halblinks oder -rechts. In gleicher Weise wurden auch Ziel- und Distraktorreize räumlich dargeboten, so dass der Zielreiz entweder von der gleichen Position wie der Hinweisreiz oder von der gegenüberliegenden Position ertönte und der Distraktorreiz von der jeweils dem Zielreiz gegenüberliegenden Position. Alle Reize wurden mit einem durchschnittlichen Schalldruckpegel von 64 dB SPL präsentiert. Die Kalibration des Kopfhörersignals wurde mithilfe eines Schallpegelmessers (2260, Brüel & Kjær), verbunden mit einem künstlichen Ohr (4153, Brüel & Kjær), durchgeführt.

Aufgabe. Alle Vpn wurden zufällig in zwei Gruppen eingeteilt, wobei die eine Gruppe auf die höhere und die andere auf die tiefere Stimme achten sollte. Die Vpn mussten bei je-

² Die räumliche Darbietung auditiver Reize kann auf verschiedene Weise realisiert werden. Während frühere Studien in der auditiven Modalität meistens monaurale Darbietungen der einzelnen Schallsignale verwendeten, das heißt ein Schallsignal nur auf dem linken und ein weiteres nur auf dem rechten Ohr (z. B. Cherry, 1953), verwenden neuere Studien entweder freistehende Lautsprecheranordnungen oder eine gezielte Manipulation der binauralen Merkmale der auditiven Orientierungsreize über einen Kopfhörer. Laut Fels, Oberem, Karnbach, Lawo und Koch (2013) ist aber von keinem großen Unterschied auf die Effekte der selektiven Aufmerksamkeit zwischen monauraler oder binauraler Darbietung auditiver Reize auszugehen (siehe dazu auch Roberts et al., 2009). Da die binaurale Darbietung eher einer realistischen Situation entspricht, wurde diese ausgewählt.

dem Durchgang angeben, ob die von der höheren bzw. tieferen Stimme gesprochene Zahl größer oder kleiner fünf war. Bei einer Zahl größer fünf mussten die Vpn die Richtungstaste „oben“ und bei einer Zahl kleiner fünf die Richtungstaste „unten“ auf dem Gamepad drücken. Den Vpn wurde erlaubt, während der Aufgabenbearbeitung die Augen zu schließen.

Prozedur. Alle Vpn erhielten die gleichen mündlichen Instruktionen und starteten mit einem Trainingsblock von 32 Durchgängen. Nach dem Trainingsblock erhielten die Vpn visuelles Feedback über ihre Trefferquote. Wenn die Trefferquote in den kongruenten Durchgängen über 75% lag, wurde das eigentliche Experiment gestartet. Ansonsten wurde der Trainingsblock solange wiederholt, bis das geforderte Kriterium erreicht war. Das Experiment bestand aus fünf Blöcken zu je 128 Durchgängen, wobei nach jedem Block die Möglichkeit einer Pause gegeben war. Der Ablauf eines Durchgangs ist schematisch in Abbildung 1 dargestellt. Alle Durchgänge starteten zunächst mit der Präsentation eines Fixationskreuzes. Nach einem Prä-Stimulus-Intervall von 200 ms wurde dann ein Hinweisreiz mit einer Dauer von 100 ms auditiv präsentiert. Der Hinweisreiz ertönte, zufällig angeordnet, in der einen Hälfte der Durchgänge von links und in der anderen Hälfte von rechts. Nach einem Inter-Stimulus-Intervall von 100 ms ertönte ein Zielreiz-Distraktor-Paar mit einer Dauer von 600 ms. Der Zielreiz wurde dabei zufällig aus dem Zahlenvorrat der höheren Stimme und der Distraktor aus dem Zahlenvorrat der tieferen Stimme ausgewählt (bzw. umgekehrt in Gruppe T), wobei beide Zahlen aus verschiedenen Kategorien (größer/kleiner fünf) stammen mussten. Genau wie der Hinweisreiz wurde der Zielreiz in der Hälfte der Durchgänge von links und in der anderen Hälfte von rechts dargeboten. Der dadurch zustande kommende räumliche Kongruenzfaktor war damit über alle Durchgänge zufällig variiert. Gleichzeitig mit der Präsentation des Zielreiz-Distraktor-Paares änderte sich der Bildschirm vom Fixationskreuz in ein Antwortdisplay, auf dem die Aufgabenstellung sowie die möglichen Antwortalternativen zu sehen waren. Die Antworteingabe war ab diesem Zeitpunkt möglich, das heißt die im weiteren Verlauf angegebenen Reaktionszeiten wurden ab dem Beginn des Zielreizes bis zum Zeitpunkt der Antworteingabe gemessen. Wenn die Vpn nicht innerhalb eines Zeitintervalls von 2600 ms antworteten, erschien das visuelle Feedback „Bitte schneller antworten!“ und die Vpn mussten einen gelben Knopf drücken, um den nächsten Durchgang zu starten. Wenn die Vpn eine falsche Antwort gaben, erschien das visuelle Feedback „Nicht richtig!“ und der nächste Durchgang wurde ebenfalls durch Drücken des gelben Knopfes gestartet. Bei korrekten Antworten wurde der nächste Durchgang automatisch nach der Reaktion der Vpn gestartet. Das gesamte Experiment dauerte zwischen 20 und 30 min.

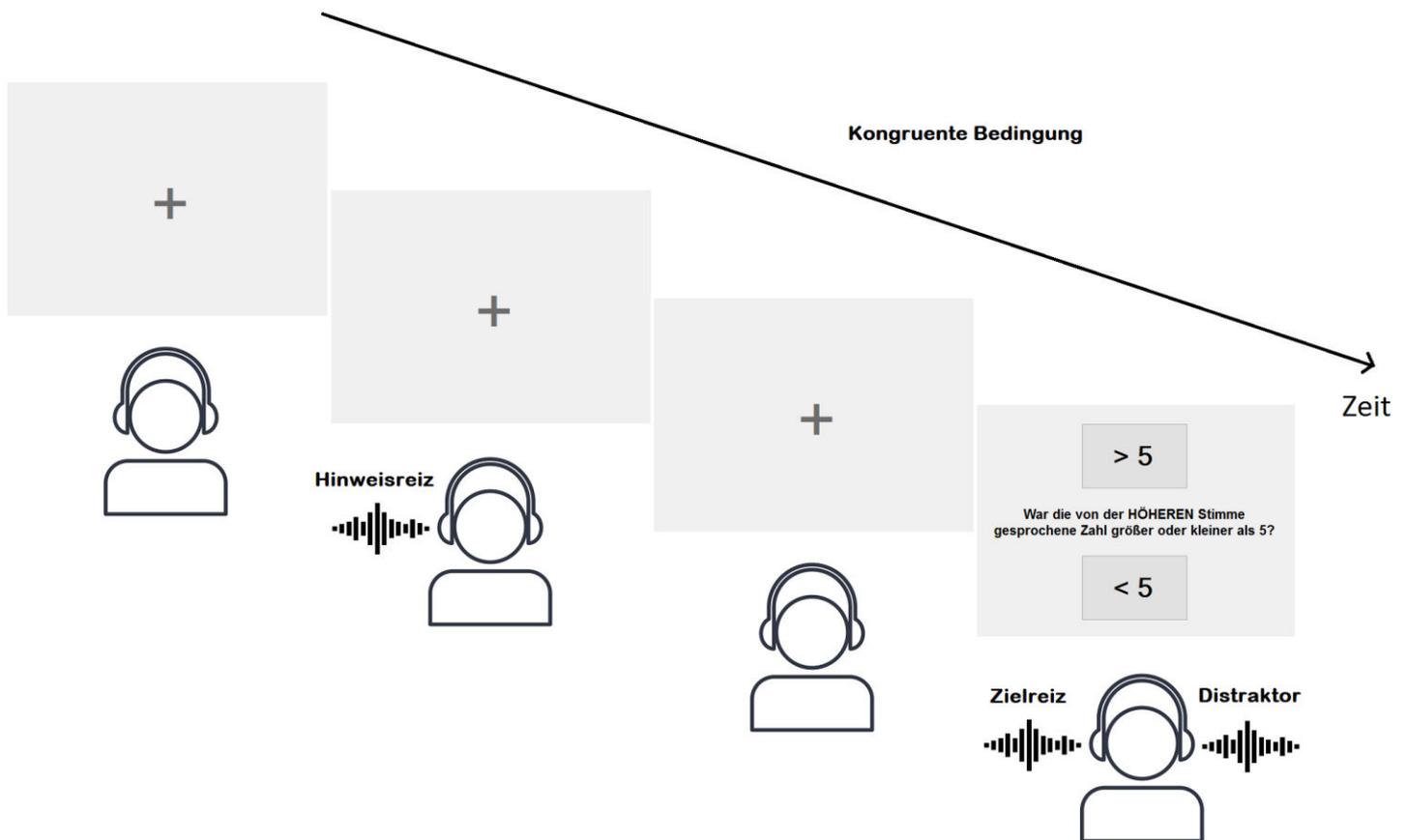


Abbildung 1. Schematische Abbildung eines Experimentaldurchganges (der kongruenten Bedingung). Der Hinweisreiz wurde zufällig entweder von links oder rechts dargeboten (in diesem Fall von links). Genauso wurde der Zielreiz zufällig entweder von links oder rechts dargeboten (in diesem Fall von links). Gleichzeitig mit dem Zielreiz ertönte jeweils ein Distraktorreiz von der gegenüberliegenden Seite (in diesem Fall von rechts). Daraus ergaben sich in der einen Hälfte des Experimentes räumlich kongruente Durchgänge, in denen Hinweis- und Zielreiz von der gleichen Position ertönten und in der anderen Hälfte des Experimentes räumlich inkongruente Durchgänge, in denen Hinweis- und Zielreiz von gegenüberliegenden Positionen ertönten.

Ergebnisse

Von der Gesamtanzahl von 5760 Durchgängen wurden 22 Durchgängen (0.38%) aufgrund der Überschreitung des Zeitlimits von der weiteren Analyse ausgeschlossen. Für die Analyse der Reaktionszeiten wurden 137 Durchgänge (2.4%) aufgrund falscher Antworten und 284 Durchgänge (4.9%) aufgrund der individuellen Ausreißerkorrektur ausgeschlossen. Die durchschnittliche Trefferquote (d. h. der Prozentsatz korrekter Antworten) über alle Bedingungen und Personen betrug 97.6% ($SD = 1.9\%$) und die durchschnittliche Reaktionszeit aller berücksichtigten Durchgänge betrug 1156 ms ($SD = 161$ ms).

Reaktionszeiten. In diesem und im nächsten Experiment wurden für die statistische Analyse der Reaktionszeiten alle Durchgänge, die mehr als zwei *SD* vom Durchschnitt der jeweiligen Person innerhalb der jeweiligen Bedingung abwichen, von der weiteren Analyse ausgeschlossen. Der Shapiro-Wilk-Test ergab keine signifikanten Abweichungen von einer Normalverteilung (NV) der Reaktionszeiten in kongruenten und inkongruenten Durchgängen (alle $p > .753$). Ein einfacher t-Test für verbundene Stichproben mit den Reaktionszeiten und dem Innersubjektfaktor *Kongruenz* (kongruent vs. inkongruent) ergab einen knapp signifikanten Effekt ($T(8) = 2.314$; $p = .049$; $d_z = .771$). Überraschenderweise waren die durchschnittlichen Reaktionszeiten in den inkongruenten Durchgängen ($M = 1141$ ms) kürzer als in den kongruenten Durchgängen ($M = 1171$ ms). Das heißt, der auditive Hinweisreiz erzeugte nicht wie erwartet einen PRK, sondern einen deutlichen NRK (siehe Abbildung 2). Für den Vergleich der beiden Gruppen wurde ein Levene-Test zur Überprüfung der Varianzhomogenität durchgeführt, welcher nicht signifikant ausfiel (alle $p > .744$). Eine weitere wechselseitige ANOVA für wiederholte Messungen der Reaktionszeiten mit dem Innersubjektfaktor *Kongruenz* (kongruent vs. inkongruent) und dem Zwischensubjektfaktor *Gruppe* (höhere Stimme vs. tiefere Stimme) ergab keine signifikante Interaktion zwischen den Faktoren *Kongruenz* \times *Gruppe* ($F < 1$; $p = .413$). Dieses Ergebnis deutet darauf hin, dass unabhängig davon, ob die Vpn auf die höhere oder tiefere Stimme reagieren sollten, der Hinweisreiz an der gleichen Position wie der Zielreiz zu einer Verzögerung in der Aufgabenbearbeitung führte. Außerdem war in dieser wechselseitigen ANOVA auch der Haupteffekt *Kongruenz* knapp nicht mehr signifikant ($p = .067$). Wird allerdings die weiter oben erwähnte Vpn, welche zunächst eine falsche Instruktion erhielt und in ihrer Leistung mit mehr als 2 *SD* vom Mittelwert aller Trefferquoten abwich, von der Analyse ausgeschlossen, ergibt sich wieder ein signifikanter Haupteffekt im Faktor *Kongruenz* ($p = .011$). Auch der scheinbare numerische Trend, der Unterschiede zwischen beiden Gruppen vermuten ließe, verschwindet unter Ausschluss der besagten Vpn (siehe dazu Tabelle 1). Insgesamt lassen die Ergebnisse daher auf keinen Unterschied zwischen den Gruppen bezüglich der Stärke des räumlichen Hinweisreizeffektes schließen.

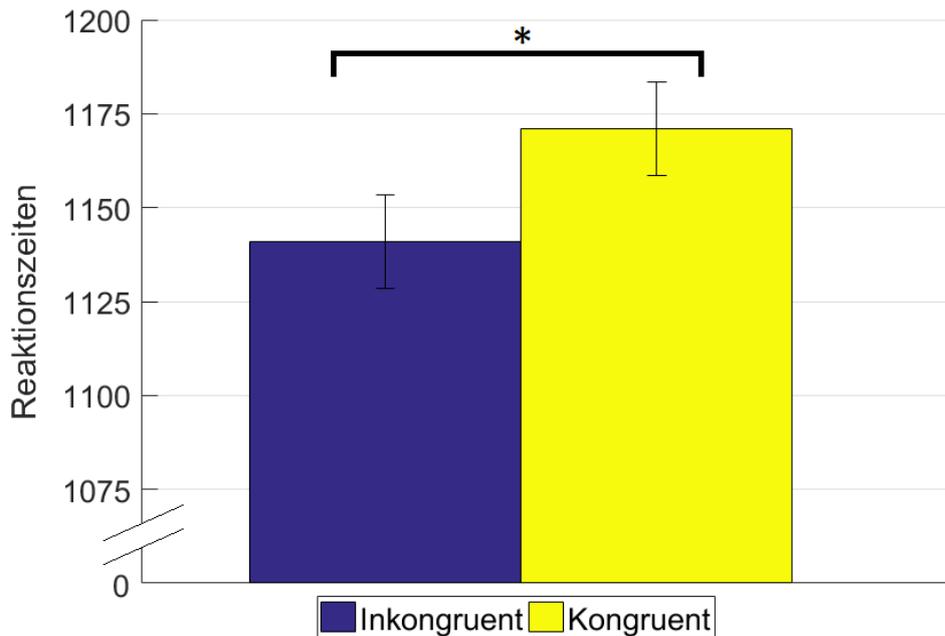


Abbildung 2. In der Abbildung sind die durchschnittlichen Reaktionszeiten (in ms) der inkongruenten und kongruenten Durchgänge dargestellt. Zusätzlich werden die standardisierten 95%-Konfidenzintervalle³ der jeweiligen Bedingungen angezeigt. * $p < .05$. ** $p < .01$. *** $p < .001$.

Tabelle 1

Numerischer Trend der möglichen Unterschiede zwischen beiden Gruppen.

	$n = 9$ (ohne Ausschluss)				$n = 8$ (mit Ausschluss)			
	Gruppe H ($n = 4$)		Gruppe T ($n = 5$)		Gruppe H ($n = 3$)		Gruppe T ($n = 5$)	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Inkongruent	1148	11	1136	8	1137	4	1136	8
Kongruent	1164	11	1176	8	1175	4	1176	8

Anmerkung. n steht für die Anzahl an Vpn in den jeweiligen Gruppen. Dargestellt sind die durchschnittlichen Reaktionszeiten aufgeteilt nach inkongruenten und kongruenten Durchgängen. Zusätzlich dazu werden die z-transformierten Standardfehler *SE* angezeigt. Es ist

³ Werden Mittelwertunterschiede in einem Innersubjektfordesign miteinander verglichen, verlieren normale 95%-Konfidenzintervalle ihre Aussagekraft. Allerdings können standardisierte Mittelwertunterschiede berechnet werden, deren 95%-Konfidenzintervall wiederum ähnlich wie in einem Zwischen-subjektfordesign interpretiert werden können (vgl. Loftus & Masson, 1994).

deutlich zu erkennen, dass sich in der statistischen Analyse unter Ausschluss der kritischen Vpn kein Unterschied zwischen Gruppe H und T zeigt.

Trefferquoten. Der Shapiro-Wilk-Test ergab eine signifikante Abweichung von einer NV der Trefferquoten in den inkongruenten Durchgängen ($p < .001$). Ein einfacher t-Test für verbundene Stichproben mit den Trefferquoten und dem Innersubjektfaktor *Kongruenz* ergab keinen signifikanten Effekt ($T(8) = .165$; $p = .873$). Der Vergleich mit dem nicht-parametrischen Wilcoxon-Vorzeichen-Rang-Test für verbundene Stichproben zeigte ein ähnliches Ergebnis ($p = .859$). Angesichts der allgemein sehr hohen Trefferquoten ist davon auszugehen, dass sich aufgrund von Deckeneffekten keine Unterschiede in den Trefferquoten finden ließen. Auch hier wurde für den Vergleich der beiden Gruppen ein Levene-Test zur Überprüfung der Varianzhomogenität durchgeführt, der diese bestätigte (alle $p > .055$). Wie in den Reaktionszeiten ergab die wechselseitige ANOVA für wiederholte Messungen der Trefferquoten mit dem Innersubjektfaktor *Kongruenz* und dem Zwischensubjektfaktor *Gruppe* keine signifikanten Effekte (alle $p > .285$). Hier ergaben sich auch unter Ausschluss der Daten der kritischen Vpn keine signifikanten Effekte (alle $p > .419$).

Diskussion

Überraschenderweise waren die Reaktionszeiten der Vpn signifikant langsamer, wenn Hinweis- und Zielreiz von derselben Position ertöntem als wenn sie von unterschiedlichen Positionen ertöntem. In den Trefferquoten spiegelte sich dieser Effekt wahrscheinlich aufgrund von Deckeneffekten nicht wider. Einerseits deuten die Ergebnisse also darauf hin, dass sich mithilfe des neu entwickelten experimentellen Design räumliche Kongruenzeffekte mit rein auditiven Reizen in einer Sprachdiskriminationsaufgabe nachweisen lassen. Andererseits bleibt zunächst unklar, warum sich statt des erwarteten PRK ein NRK zeigte. Wie in der theoretischen Einführung bereits erwähnt, hängt die Art des räumlichen Kongruenzeffektes von mehreren Faktoren ab. Beispielsweise hat die Länge der SOA (Mondor, Breau et al., 1998; Teshiba et al., 2012), die Art der Aufgabenstellung (Mondor, Breau et al., 1998), die Wahrscheinlichkeit, mit der der Hinweisreiz die Position des Zielreizes korrekt vorhersagt (Mondor & Bryden, 1992; Roberts et al., 2009), sowie die Ähnlichkeit zwischen Hinweis- und Zielreiz (Mondor, Breau et al., 1998; Mondor, Zatorre et al., 1998; Dalton & Lavie, 2004) einen entscheidenden Einfluss auf den zu erwartenden Kongruenzeffekt. Üblicherweise wurde ein derartiger NRK in auditiven Studien bei einer SOA von 700 ms beobachtet (Mondor, Breau et al., 1998; Teshiba et al., 2012) und als IOR, der ursprünglich in der visuellen Modalität beobachtet wurde, interpretiert. Allerdings beschreibt Klein (2000), zumindest in der visuellen

Modalität, die SOA zwischen 200-300 ms als kritischen Wendepunkt von einem PRK in einen NRK, also im Bereich der in Experiment 1 verwendeten SOA von 200 ms. Gleichzeitig verschiebt sich dieser zeitliche Wendepunkt bei komplexeren Aufgabenstellungen jedoch eher zu längeren SOAs (Lupiáñez, 2010; Mondor, Breau et al., 1998). Aufgrund der vergleichsweise komplexen Aufgabenstellung in Experiment 1, erscheint die Erklärung des Effekts als IOR unplausibel. Des Weiteren ergab der Vergleich der durchschnittlichen Reaktionszeiten in den kongruenten und inkongruenten Durchgängen keinen signifikanten Unterschied zwischen den Gruppen, obwohl sich die Ähnlichkeit zwischen Hinweis- und Zielreiz in beiden Gruppen unterschied. Da der Hinweisreiz aus dem Sprachmaterial der hohen Sprecherin erzeugt wurde, sollte die Ähnlichkeit zwischen Hinweis- und Zielreiz in der Gruppe, die auf die höhere Stimme achten musste, höher sein als in der Gruppe, die auf die tiefere Stimme achten sollte. Wie bereits erwähnt, hat die Ähnlichkeit zwischen Hinweis- und Zielreiz einen Einfluss auf die Effektivität des Hinweisreizes (z. B. Mondor, Breau et al., 1998; Mondor, Zatorre et al., 1998). Wenn die Suchschablone der Vpn die Merkmale des Hinweisreizes nicht enthält, zieht dieser die Aufmerksamkeit nicht auf sich und hat daher keinen Einfluss auf die Verarbeitung des Zielreizes (z. B. Folk & Remington, 1998). Einige Studien in der visuellen Modalität fanden sogar Hinweise darauf, dass wenn sich Hinweis- und Zielreiz sehr unähnlich sind und beide an der gleichen Position präsentiert werden, dies zu einer Verzögerung der Verarbeitung der Zielreize führen kann (Carmel & Lamy, 2014; Carmel & Lamy, 2015; Lamy, Leber, & Egeth, 2004). Da sich zwischen beiden Gruppen kein Unterschied in der Richtung des räumlichen Kongruenzeffektes zeigte, wäre eine mögliche Erklärung, dass die Ähnlichkeit zwischen Hinweis- und Zielreiz in *beiden* Gruppen zu gering war und sich deshalb ein NRK zeigte. Bei der Erzeugung des Hinweisreizes mittels Überlagerung zufällig gewählter zeitlicher Ausschnitte aus allen Zahlwörtern war die Intention, einerseits die allgemeine Charakteristik von Sprachsignalen zu erhalten und andererseits nicht ein einzelnes Exemplar der Zahlwörter zu bevorzugen. Die retrospektive auditive Überprüfung des Kluges zeigte allerdings, dass sich die Klangfarbe des Hinweisreizes deutlich von jener der einzelnen Zahlwörter unterscheidet, was auf Phaseneffekte bei der Überlagerung der einzelnen Sprachsegmente zurückzuführen ist. Die Erklärung des NRK auf Basis der HZÄ wurde in Experiment 2 überprüft.

Experiment 2

Beschreibung und erwartete Ergebnisse

Erstaunlicherweise wurde der Zusammenhang zwischen HZÄ und der Übereinstimmung der räumlichen Position von Hinweis- und Zielreiz in der auditiven Modalität bislang noch wenig untersucht. Obwohl in den Studien von Mondor, Breau et al. (1998) und Mondor,

Zatorre et al. (1998) sowohl die Frequenz als auch die räumliche Position von Hinweis- und Zielreiz zufällig variiert wurden, analysierten die AutorInnen die Ergebnisse nicht hinsichtlich der unterschiedlichen Ähnlichkeit zwischen Hinweis- und Zielreiz in Abhängigkeit ihrer räumlichen Übereinstimmung. In der vorliegenden Studie wurde dies realisiert, indem die HZÄ durch die Verwendung von drei unterschiedlichen Hinweisreizarten systematisch variiert wurde. Dazu wurde zusätzlich zum in Experiment 1 verwendeten Hinweisreiz ein Hinweisreiz getestet, der dem Sprachsignal möglichst ähnlich sein sollte und ein weiterer Hinweisreiz, der nicht-sprachlich und somit minimale HZÄ repräsentiert. In einem zu Experiment 1 ähnlichen experimentellen Design wurde der Einfluss der jeweiligen Hinweisreiztypen auf den räumlichen Kongruenzeffekt untersucht.

Es wurde vermutet, dass eine hohe HZÄ zu einem positiven und eine geringe HZÄ zu einem negativen räumlichen Kongruenzeffekt führt. Das heißt, Vpn profitieren im Falle des sprachlichen Hinweisreizes von einer Voreinstellung der Aufmerksamkeit und können somit in kongruenten Durchgängen schneller auf den Zielreiz reagieren und erzielen höhere Trefferquoten als in inkongruenten Durchgängen. Umgekehrt führen jedoch die sprachunähnlichen Hinweisreize zu einer Verzögerung der Verarbeitung des Zielreizes und die Vpn reagieren langsamer und erzielen niedrigere Trefferquoten in kongruenten als in inkongruenten Durchgängen. Damit könnte zum einen gezeigt werden, dass die HZÄ auch in der auditiven Modalität eine entscheidende Rolle für das Zustandekommen eines räumlichen Kongruenzeffektes spielt und zum anderen würde eine solche Abhängigkeit zu einem besseren Verständnis der Steuerung der selektiven Aufmerksamkeit innerhalb einer auditiven Szene führen.

Methode

Versuchspersonen. Es wurden 19 Vpn (10 weibliche) im Labor des Instituts für Schallforschung der Österreichischen Akademie der Wissenschaften in Wien getestet, wobei die Altersspanne von 21 bis 32 Jahren reichte ($M = 25.7$; $SD = 3.1$). Alle Vpn wiesen normale Hörfähigkeit und normale oder dazu korrigierte Sehfähigkeit auf. Allerdings berichtete eine der Vpn die Wahrnehmung von Tinnitus. Eine Analyse ohne deren Daten ergab jedoch keinen Unterschied in den Signifikanztests. Außerdem wurde eine weibliche Vpn von der weiteren Analyse ausgeschlossen, da sie die geforderte Trefferquote in der Trainingsphase auch nach sehr vielen Wiederholungen nicht erreichte und diese sich im Bereich der Zufallswahrscheinlichkeit befand. Eine mögliche Erklärung wäre, dass sie als einzige Vpn nicht Muttersprachlerin war und ihr deshalb die Aufgabenstellung schwerer fiel.

Apparatur und Reize. Die verwendete Apparatur war identisch mit Experiment 1. Auch wurden die gleichen auditiven Reize wie in Experiment 1 verwendet, außer dass zusätz-

lich zu dem bereits vorhandenen Hinweisreiz zwei weitere Arten von Hinweisreizen getestet wurden. Die drei getesteten Hinweisreiztypen sollten sich in ihrer Ähnlichkeit zu den Zielreizen unterscheiden. Der Hinweisreiztyp mit der größten HZÄ (Vokalisation, im Folgenden auch „sprachlicher“ Hinweisreiz genannt) bestand aus den ersten 100 ms der gesprochenen Zahlwörter „eins“ und „acht“ und zwar von der in ihrer Tonhöhe erhöhten Stimme. Es wurden diese beiden Zahlwörter ausgewählt, da sie in den ersten 100 ms phonetisch nur sehr schwer voneinander zu unterscheiden sind und außerdem jeweils aus einer der beiden Antwortkategorien stammten. In der entsprechenden Experimentalbedingung wurde dann zufällig einer der beiden Reize (d. h. entweder die ersten 100 ms des Zahlworts „eins“ oder die ersten 100 ms des Zahlworts „acht“) ausgewählt, um mögliches semantisches Priming gleichmäßig auf die beiden Antwortkategorien zu verteilen.⁴ Der zweite Hinweisreiztyp (Sprachgemisch) bestand wie in Experiment 1 aus einer Überlagerung mehrerer Sprachsegmente, die zufällig aus allen möglichen Zielreizen ausgewählt wurden, das heißt, dieser wurde auch aus dem Stimmmaterial der höheren Stimmen gebildet. Dieser Hinweisreiztyp wurde verwendet, da mit ihm in Experiment 1 der NRK bereits gezeigt wurde, der in diesem Experiment nun repliziert werden sollte. In Experiment 1 wurde der NRK auf die geringe HZÄ zurückgeführt. Der dritte Hinweisreiztyp (Rauschen) bestand aus einem weißen Breitbandrauschsignal. Da er nicht auf Sprachmaterial beruhte, wurde minimale HZÄ angenommen. Alle Hinweisreize wiesen eine Dauer von 100 ms auf und wurden mit einem durchschnittlichen Schalldruckpegel von 64 dB SPL präsentiert. Allerdings ergab eine genauere post-hoc-Messung der individuellen Schalldruckpegel eine Abweichung. Der Rauschsignalpegel war um ca. 4 dB SPL höher als für die beiden anderen Hinweisreiztypen. Mögliche Konsequenzen auf die inhaltliche Interpretation werden im Diskussionsteil angesprochen.

Aufgabe und Prozedur. Die Aufgabestellung war identisch mit Experiment 1, außer dass nun alle Vpn auf die höhere Stimme achten sollten. Die Prozedur änderte sich im Vergleich zu Experiment 1 wie folgt. Alle Vpn starteten mit einem Trainingsblock von 24 Durchgängen. Nach erfolgreichem Abschluss der Trainingsphase startete das Hauptexperiment. Diese bestand aus sechs Blöcke zu je 192 Durchgängen, wobei nach jedem Block die Mög-

4 Eine separate qualitative Auswertung der durchschnittlichen Reaktionszeiten der berücksichtigten Durchgänge mit spezifischen Zahlenwerten als Zielreiz ergab schnellere Reaktionszeiten in Durchgängen, bei denen die Zahl 1 oder 8 als Zielreiz erklang (siehe Tabelle 2 im Anhang), was auf semantisches Priming hinweist. Allerdings zeigten sich in einer separaten 3 x 2 ANOVA der Reaktionszeiten ohne die Durchgänge mit 1 oder 8 als Zielreiz keine Unterschiede in den Signifikanztests (siehe Abbildung 5 im Anhang) in Bezug auf die Kongruenzeffekte.

lichkeit einer Pause gegeben war. Alle Durchgänge starteten zunächst mit der Präsentation eines Fixationskreuzes. Nach einem Prä-Stimulus-Intervall von 500 ms wurde ein Hinweisreiz mit einer Dauer von 100 ms auditiv präsentiert. In jeweils einem Drittel der Durchgänge bestand der Hinweisreiz aus der Vokalisation, dem Sprachgemisch oder dem Rauschen, wobei die Hinweisreiztypen zufällig über alle Durchgänge variiert wurden. Außerdem ertönte der Hinweisreiz, in zufälliger Reihenfolge, in der einen Hälfte der Durchgänge von links und in der anderen Hälfte von rechts. Nach einem Inter-Stimulus-Intervall von 100 ms ertönte ein Zielreiz-Distraktor-Paar. Der Zielreiz wurde dabei zufällig aus dem Zahlenvorrat der höheren Stimme und der Distraktorreiz aus dem Zahlenvorrat der tieferen Stimme ausgewählt, wobei beide Zahlen aus verschiedenen Kategorien stammen mussten. Genau wie der Hinweisreiz wurde der Zielreiz in der Hälfte der Durchgänge von links und in der anderen Hälfte von rechts dargeboten. Wie in Experiment 1 wurde der dadurch zustande kommende räumliche Kongruenzfaktor über alle Durchgänge zufällig verteilt. Das gesamte Experiment dauerte aufgrund der deutlich größeren Anzahl an Durchgängen zwischen 50-70 min.

Ergebnisse

Von der Gesamtanzahl von 20736 Durchgängen wurden 141 Durchgänge (0.68%) aufgrund der Überschreitung des Zeitlimits von der weiteren Analyse ausgeschlossen. Für die Analyse der Reaktionszeiten wurden 1324 Durchgänge (6.4%) aufgrund falscher Antworten und 892 Durchgänge (4.3%) aufgrund der individuellen Ausreißerkorrektur ausgeschlossen. Die durchschnittliche Trefferquote über alle Bedingungen und Personen betrug 93.5% ($SD = 6.75\%$) und die durchschnittliche Reaktionszeit aller berücksichtigten Durchgänge betrug 1327 ms ($SD = 156$ ms).

Reaktionszeiten. Der Shapiro-Wilk-Test ergab eine signifikante Abweichung von einer NV der Reaktionszeiten in den inkongruenten Durchgängen des Hinweisreiztyps Vokalisation ($p = .01$). Da es sich bei der ANOVA aber um ein robustes Verfahren handelt und in allen anderen Bedingungen die Reaktionszeiten nicht signifikant von einer NV abwichen, wurde trotzdem ein ANOVA gerechnet. Eine wechselseitige 3 x 2 ANOVA für wiederholte Messungen der Reaktionszeiten mit den Innersubjektfaktoren *Hinweisreiztyp* (Vokalisation vs. Sprachgemisch vs. Rauschen) und *Kongruenz* (kongruent vs. Inkongruent) ergab einen signifikanten Haupteffekt für den Faktor *Hinweisreiztyp* ($F = 7.641$; $p = .002$; $\eta^2_p = .310$). Außerdem war die Interaktion zwischen den Faktoren *Hinweisreiztyp* x *Kongruenz* ($F = 19.631$; $p < .001$; $\eta^2_p = .536$) signifikant. Die geplanten Vergleiche mittels t-Tests für verbundene Stichproben der jeweiligen Bedingungen ergaben signifikante Mittelwertunterschiede bei allen drei Hinweisreiztypen (siehe Abbildung 3): Konsistent mit unseren Hypothesen

konnten die Vpn unter Einfluss des sprachlichen Hinweisreiztyps (Vokalisation) den Zielreiz in kongruenten ($M = 1314$ ms) schneller verarbeiten als in inkongruenten ($M = 1372$ ms) Durchgängen. Im Gegensatz dazu führten die beiden anderen Hinweisreiztypen zu langsameren Reaktionszeiten in kongruenten als in inkongruenten Durchgängen, sowohl für den Hinweisreiztyp Sprachgemisch ($M = 1347$ ms vs. $M = 1308$ ms) als auch für den Hinweisreiztyp Rauschen ($M = 1322$ ms vs. $M = 1304$ ms). Übereinstimmend mit unseren Hypothesen führte eine geringe HZÄ also zu einem NRK und eine hohe HZÄ zu einem PRK.

Zusätzlich wurden für die in der Diskussion gezogenen Schlussfolgerungen zwei weitere post-hoc-Tests mittels t-Test für verbundene Stichproben durchgeführt, um diese statistisch zu untermauern. Der Vergleich zwischen den kongruenten Durchgängen des Hinweisreizes „Vokalisation“ mit den jeweiligen inkongruenten Durchgängen der Hinweisreize „Sprachgemisch“ und „Rauschen“ ergab keine signifikanten Unterschiede in den durchschnittlichen Reaktionszeiten (alle $p > .234$). Genauso unterschieden sich die inkongruenten Durchgänge von den kongruenten Durchgängen sowohl im Hinweisreiztyp Sprachgemisch als auch im Hinweisreiztyp Rauschen gleich stark voneinander. Dies wurde überprüft, indem für jede Vpn die durchschnittliche Reaktionszeit der inkongruenten Durchgänge von jenen der kongruenten Durchgänge subtrahiert wurde und zwar separat für den jeweiligen Hinweisreiztyp. Die so gewonnenen Werte wurden zwischen den beiden Hinweisreiztypen Sprachgemisch und Rauschen miteinander verglichen ($p = .0685$).

Da die Belegung der Antworttasten zwischen den Vpn nicht variiert wurde, wurde eine weitere wechselseitige 2×2 ANOVA für wiederholte Messungen mit den Innersubjektfaktoren *Zielreizposition* (links vs. rechts) und *Antworttaste* (oben vs. unten) gerechnet, um orthogonale Simon-Effekte (vgl. dazu Proctor & Cho, 2006) auszuschließen. Keiner der Effekte war signifikant (alle $p > .206$).⁵

Trefferquoten. Der Shapiro-Wilk-Test ergab in allen Bedingungen eine signifikante Abweichung von einer NV der Trefferquoten. Diesem Umstand wurde Rechnung getragen, indem, soweit möglich, ein zusätzliches nicht-parametrisches Verfahren angewendet wurde. Die wechselseitige ANOVA für wiederholte Messungen der Trefferquoten mit den gleichen Innersubjektfaktoren ergab einen signifikanten Haupteffekt für den Faktor *Hinweisreiztyp* ($F = 7.226$; $p = .002$; $\eta^2_p = .298$). Außerdem war die Interaktion zwischen den Faktoren *Hinweisreiztyp* \times *Kongruenz* ($F = 5.142$; $p = .011$; $\eta^2_p = .232$) signifikant. Die geplanten Vergleiche mittels t-Tests für verbundene Stichproben der jeweiligen Bedingungen ergaben jedoch nur

⁵ Siehe Abbildung 6 im Anhang.

einen signifikanten Mittelwertunterschied im Hinweisreiztyp Vokalisation: Konsistent mit den Ergebnissen der Reaktionszeiten machten die Vpn weniger Fehler in kongruenten ($M = .934$) als in inkongruenten ($M = .917$) Durchgängen, wenn ihnen vor dem Zielreiz ein sprachlicher Hinweisreiz präsentiert wurde. Obwohl in beiden anderen Hinweisreiztypen die Unterschiede in den Trefferquoten zwischen kongruenten und inkongruenten Durchgängen nicht das geforderte Signifikanzniveau erreichten ($p = .195$ für das Sprachgemisch und $p = .0645$ für das Rauschsignal), zeigte der numerische Trend ein zu den Reaktionszeiten passendes Bild (siehe Abbildung 4). Der Vergleich mit dem nicht-parametrischen Wilcoxon-Vorzeichen-Rang-Test für verbundene Stichproben zeigte keine Unterschiede in den Signifikanztests.

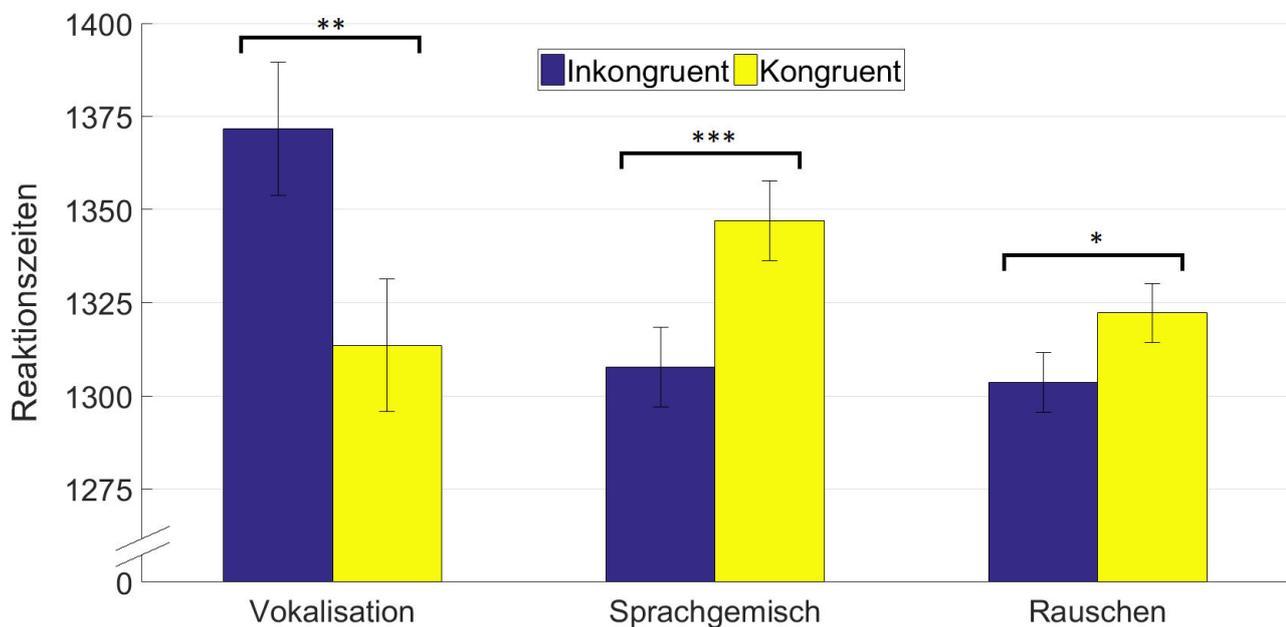


Abbildung 3. In der Abbildung sind die durchschnittlichen Reaktionszeiten (in ms) aufgeteilt nach Hinweisreiztypen dargestellt. Zusätzlich werden die standardisierten 95%-Konfidenzintervalle der jeweiligen Bedingungen angezeigt. Die inkongruenten Durchgänge unterscheiden sich in allen drei Hinweisreiztypen signifikant von den kongruenten Durchgängen. * $p < .05$. ** $p < .01$. *** $p < .001$.

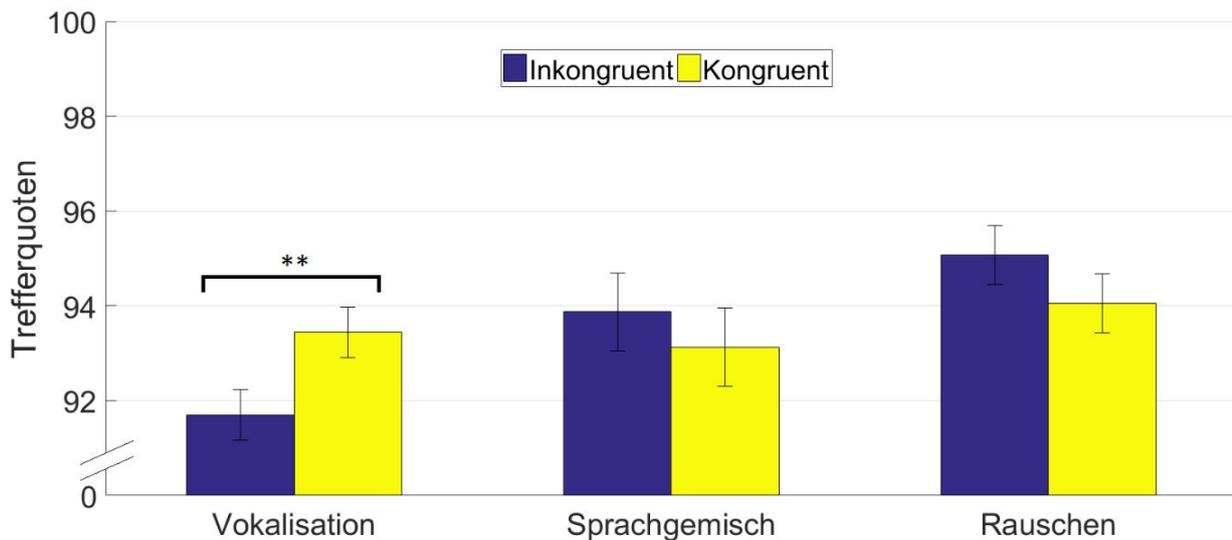


Abbildung 4. In der Abbildung sind die durchschnittlichen Trefferquoten (in %) aufgeteilt nach Hinweisreiztypen dargestellt. Zusätzlich werden die standardisierten 95%-Konfidenzintervalle der jeweiligen Bedingungen angezeigt. Nur im Falle des Hinweisreiztyps Vokalisation unterscheiden sich die inkongruenten signifikant von den kongruenten Durchgängen. * $p < .05$. ** $p < .01$. *** $p < .001$.

Generelle Diskussion

Gemäß unseren Hypothesen variierte die Richtung des Kongruenzeffektes in den Reaktionszeiten je nach Art des Hinweisreizes. Bestand der Hinweisreiz aus einem sprachlichen Signal (Vokalisation) ergab sich ein positiver Kongruenzeffekt, das heißt schnellere Reaktionszeiten in Durchgängen, in denen Hinweis- und Zielreiz am selben Ort ertönten. Bestand der Hinweisreiz hingegen aus einem sprachunähnlichen oder nicht-sprachlichen Signal (Sprachgemisch oder Rauschen), kehrte sich der positive in einen negativen Kongruenzeffekt um, das heißt langsamere Reaktionszeiten in Durchgängen, in denen Hinweis- und Zielreiz am selben Ort ertönten. Die zeitliche Verarbeitung des Zielreizes war also abhängig von der Ähnlichkeit zwischen Hinweis- und Zielreiz. In den Trefferquoten zeigte sich ein ähnliches Bild, obwohl der räumliche Kongruenzeffekt nur in der Bedingung des sprachlichen Hinweisreizes signifikant war. Bei den Trefferquoten in den Bedingungen der beiden anderen Hinweisreiztypen stimmte zwar die Richtung des Kongruenzeffektes mit den Reaktionszeiten überein, aber die Effekte erreichten nicht das geforderte Signifikanzniveau. Ähnlich wie in Experiment 1 könnte dies wiederum auf Deckeneffekte zurückgeführt werden.

Damit wurde zum einen der in früheren auditiven Studien üblicherweise aufgetretene Erleichterungseffekt repliziert, der die schnellere Verarbeitung von Zielreizen an Orten, an denen zuvor ein Hinweisreiz ertönte, propagiert (Mondor & Amirault, 1998; Mondor, Breau et al., 1998; Mondor, Zatorre et al., 1998; Roberts et al., 2009; Teshiba et al., 2012) und zum

anderen der in Experiment 1 aufgetretene NRK, bei dem die Verarbeitung des Zielreizes durch zuvor und am gleichen Ort erklungene Hinweisreize verlangsamt wurde. Gleichzeitig kann der NRK nicht auf den zu Beginn erwähnten IOR zurückgeführt werden, da die SOA über alle Bedingungen konstant gehalten wurde. Offen bleibt daher die Frage, wie die Übereinstimmung oder Nicht-Übereinstimmung der Merkmale von Hinweis und Zielreiz einerseits Erleichterungs- und andererseits Hemmungseffekte produzieren können.

Erklärungen innerhalb der auditiven Modalität

Zunächst erscheint es einfacher, den PRK zu erklären. Der sprachliche Hinweisreiz zieht die Aufmerksamkeit auf sich, weshalb in weiterer Folge die Verarbeitung des vom gleichen Ort kommenden Zielreizes erleichtert wird. Wird der Hinweisreiz hingegen an der räumlich gegenüberliegenden Seite dargeboten, kann die Verarbeitung des Zielreizes nicht von einer Voreinstellung der Aufmerksamkeit profitieren, da diese erst auf die Position des Zielreizes gerichtet werden muss. Durch diese notwendige Aufmerksamkeitsverlagerung verzögert sich die Wahrnehmung und somit auch die Enkodierung des Zielreizes, was sich in langsameren Reaktionszeiten niederschlägt (Kanai, Ikeda, & Tayama, 2007; Mondor & Amirault, 1998).

Unklarer ist hingegen die Interpretation des NRK. Ein ähnlicher Effekt findet sich beispielsweise in Experiment 2B bei Mondor, Zatorre et al. (1998). Hierbei wurden den Vpn zwei aufeinanderfolgende Töne an der gleichen räumlichen Position dargeboten, die in der einen Hälfte der Durchgänge die gleiche und in der anderen Hälfte eine unterschiedliche Frequenz aufwiesen. Die Durchgänge wurden zufällig ausgewählt und sowohl der Hinweis- als auch der Zielreizton konnte fünf unterschiedliche Frequenzen aufweisen. Außerdem variierte die Signalanstiegszeit des Zielreizes (zweiter Ton) zufällig zwischen schnell und langsam. Die Aufgabe der Vpn bestand in der Angabe eben dieser Signalanstiegszeit. Dabei zeigte sich, dass die Vpn für die Verarbeitung des Zielreizes mehr Zeit benötigten, wenn Hinweis- und Zielreiz unterschiedliche Frequenzen aufwiesen. Mondor, Zatorre et al. erklären ihren Befund folgendermaßen. Zwischen Hinweisreiz und den Parametern der Aufmerksamkeitschablone (AS, engl. attentional template) herrscht eine dynamische Interaktion. Wenn der oder die Hörende die AS nicht schon im Voraus festsetzen kann, da beispielsweise zu wenig Informationen über die Identität des Zielreizes vorhanden sind, beeinflusst der Hinweisreiz diese Festsetzung. Der Grad der Passung zwischen dem Zielreiz und der AS (und damit dem Hinweisreiz) bestimmt dann die Geschwindigkeit, mit der die Selektion des Zielreizes vollzogen werden kann.

Erleichterungseffekte entstehen daher bei guter Hinweis-Zielreiz-Passung und Verzögerungseffekte bei schlechter Passung. Die Frage ist, ob in der vorliegenden Studie genügend Informationen über die Identität des Zielreizes vorhanden waren, um von der eben beschriebenen Interaktion auszugehen. Immerhin sollte nach einer gewissen Anzahl von Durchgängen die Stimme der höheren Sprecherin bekannt erscheinen und sich somit die AS über die Zeit schärfen.⁶ Zumindest schließen Mondor, Zatorre et al. (1998) aus ihrem dritten Experiment, bei dem die Vpn einen vordefinierten Zielreizton in einer Sequenz mit mehreren Distraktortönen erkennen mussten, dass sowohl die Übereinstimmung der Position als auch die der Frequenz zwischen Distraktor- und Zielreizton einen Einfluss auf die Selektion auditiver Informationen hat, unabhängig davon, ob die AS reiz- oder absichtsgesteuert gebildet wird. Wie im Falle einer absichtsgesteuerten AS die entsprechenden Effekte entstehen, wird allerdings nicht näher erläutert.

In einer Studie von Mayr, Buchner, Möller und Hauke (2011) wurde auch die Interaktion zwischen der Übereinstimmung der Position und der Ähnlichkeit von Hinweis- und Zielreiz untersucht. Den Vpn wurden in mehreren Durchgängen zwei aufeinanderfolgende Tierlaute in Kombination mit einem weiteren Distraktor-Tierlaut dargeboten, wobei Ziel- und Distraktorreiz aus zwei von vier unterschiedlichen räumlichen Positionen ertönten. Vor jeder Kombination von Tierlauten erschien ein visueller Hinweisreiz in Form eines Bildes des Tieres, auf welches die Aufmerksamkeit gerichtet werden sollte, welcher sich von der ersten zur zweiten Tierlautkombination änderte. Die Aufgabe der Vpn bestand in der Angabe der Position des zu beachtenden Tierlautes. Wichtig hierbei ist, dass der nicht zu beachtende Distraktor-Tierlaut von der ersten Tierlautkombination (D1) entweder mit dem Zielreiz aus der zweiten Tierlautkombination (T2) übereinstimmte oder nicht und zwar sowohl hinsichtlich der räumlichen Position als auch der Art des Tierlautes. Es zeigte sich, dass die Vpn auf T2 schneller reagieren konnten, wenn D1 aus demselben Tierlaut bestand und von derselben Position ertönte, als wenn er von einer anderen Position ertönte. Wenn D1 allerdings aus einem von T2 verschiedenen Tierlaut bestand, reagierten Vpn langsamer auf T2 von derselben Position und schneller auf T2 auf einer anderen Position. Das heißt, der positive kehrte sich in einen negativen räumlichen Kongruenzeffekt um, wenn D1 statt aus einem passenden aus einem unpass-

⁶ In Abbildung 7 und 8 im Anhang sind die durchschnittlichen Reaktionszeiten und Trefferquoten der jeweiligen Hinweisreiztypen in ihrer zeitlichen Entwicklung über die sechs Experimentalsitzungen abgebildet.

senden Hinweisreiz bestand. Interessant ist hierbei, dass dieser Effekt auftrat, obwohl die Aufmerksamkeit explizit vom Distraktor-Tierlaut in der ersten Tierlautkombination abgewendet war. Mayr et al. (2011) erklären ihren Befund mithilfe der Theorie der Objektbindung (engl. object-binding). Auditive Merkmale, wie eine bestimmte räumliche Position oder spektral-temporale Signalstruktur, werden dabei spontan in episodische auditive Objekte integriert, welche im Kurzzeitgedächtnis gespeichert werden. Diese Episoden werden als Ganzes nur dann wieder abgerufen, wenn zumindest eines ihrer Elemente im späteren Verlauf erneut vorgefunden wird. Der Abruf findet hingegen nicht statt, wenn keines der Merkmale wiederholt wird. Dies wäre der Fall in der Bedingung, in welcher sowohl die Art des Tierlautes als auch die Position des Distraktors verschieden vom Zielreiz-Tierlaut war. In der Bedingung, in der die Art des Tierlautes und die Position von Distraktor und Zielreiz identisch waren, ist der Abruf des Objektes unproblematisch oder sogar vorteilhaft für die nachfolgende Verarbeitung. Wird jedoch nur eines der Merkmale (und nicht alle) wiederholt, führt dies zu einer Reaktivierung von irreführender und widersprüchlicher Information, welche die Verarbeitung und damit die Reaktion verlangsamt. Dies war der Fall, wenn sich entweder die Art des Tierlautes änderte jedoch nicht die Position oder die Art des Tierlautes gleichblieb, sich aber die Position, jeweils zwischen Distraktor und Zielreiz, änderte. Trotz der teilweise starken Unterschiede im experimentellen Design, lassen sich die Ergebnisse der vorliegenden Studie in ähnlicher Weise erklären. Übereinstimmend damit zeigte sich kein signifikanter Unterschied in den Reaktionszeiten zwischen der kongruenten Bedingung des sprachlichen und der inkongruenten Bedingung des sprach-unähnlichen beziehungsweise des nicht-sprachlichen Hinweisreiztyps (siehe Ergebnisteil S. 21). Variierte jedoch nur entweder die Übereinstimmung der Position oder nur die Ähnlichkeit zwischen Hinweis- und Zielreiz, kam es zu entsprechenden zeitlichen Kosten in der Verarbeitung des Zielreizes. Offen ist, ob in der vorliegenden Studie die Hinweisreize die Aufmerksamkeit auf sich gezogen haben oder nicht. Weitere Studien könnten dies durch Hinzufügen einer Kontrollbedingung, in der keine Hinweisreize gezeigt werden, überprüfen. Gleichzeitig steht aber eine Aufmerksamkeitszuwendung zum Hinweisreiz nicht im Widerspruch zu der zuvor beschriebenen Objektbindungstheorie. Da auch in der visuellen Modalität ähnliche Erklärungsansätze herangezogen werden, diskutiert der nächste Abschnitt mögliche Parallelen zwischen beiden Modalitäten.

Vergleich zwischen auditiver und visueller Modalität

Das experimentelle Design der vorliegenden Studie wurde unter anderem vom Phänomen der bedingten Aufmerksamkeitsanziehung inspiriert. Mehrere Studien in der visuellen Modalität zeigten, dass Hinweisreize die Aufmerksamkeit nur dann auf sich ziehen, wenn sie

Merkmale besitzen, nach denen gesucht wird (Folk & Remington, 1998; Folk et al., 1992). Beispielsweise mussten Vpn in einer Studie von Folk und Remington (1998) nach einem roten Zielreiz suchen. Ein irrelevanter farblicher Hinweisreiz zog die Aufmerksamkeit nur auf sich, wenn er die gleiche Farbe wie der Zielreiz trug, beeinflusste die Bearbeitung der Aufgabe jedoch nicht, wenn er in einer anderen Farbe präsentiert wurde. Darüber hinaus konnten Lamy et al. (2004) in einer ähnlichen Studie Hinweise darauf finden, dass zuvor gezeigte irrelevante Hinweisreize, die eine andere Farbe als der Zielreiz hatten, die Verarbeitung der Zielreize sogar erschwerten, wenn sie am gleichen Ort präsentiert wurden. Carmel und Lamy (2015) führten diese scheinbar widersprüchlichen Ergebnisse auf die unterschiedlichen Aufgabenstellungen beider Studien zurück. Des Weiteren zeigten sie in einer Reihe von Experimenten, dass sich die Erleichterungs- und Hemmungseffekte innerhalb des Hinweisreizparadigmas aus einer Kombination dreier separater Effekte verstehen lassen, wobei für die Zwecke dieser Studie nur die ersten zwei von Belang sind. Erstens, ein Vorteil für Reize an der gleichen räumlichen Position, welcher durch die Ähnlichkeit zwischen Hinweis- und Zielreiz bestimmt wird und mit dem Phänomen der bedingten Aufmerksamkeitsanziehung zusammenhängt. Zweitens, Kosten für Reize an der gleichen räumlichen Position, bei geringer Ähnlichkeit zwischen Hinweis- und Zielreiz, die aber erst nach der Selektion des Zielreizes auftreten und mit Prozessen des visuellen Arbeitsgedächtnisses zusammenhängen. Auch wenn in jenem Artikel die zugrundeliegenden Mechanismen dieses Kosteneffekts als noch nicht ausreichend belegt klassifiziert wurden, beschreiben frühere Artikel diese Kosten als Folge einer Objektaktualisierung (engl. object-updating; Carmel & Lamy, 2014, Lamy, Alon, Carmel, & Shalev, 2015). Die episodische Repräsentation eines Objektes (engl. object files; Kahneman, Treisman, & Gibbs, 1992) im visuellen Arbeitsgedächtnis muss aktualisiert werden, wenn sich das Objekt ändert. Diese Erklärung wird auch für eine Reihe von anderen Phänomenen herangezogen, wie zum Beispiel das negative Priming (Park & Kanwisher, 1994) oder die Merkmalsbindung (engl. feature binding; Hommel, 1998), und weist somit starke Ähnlichkeiten mit der in der auditiven Modalität beschriebenen Theorie der Objektbindung auf. In Übereinstimmung damit legen einige Studien nahe, dass Aufmerksamkeits- und Gedächtnisprozesse in der visuellen Modalität den gleichen neuronalen Ressourcen wie in der auditiven Modalität unterliegen (Mayer, Harrington, Stephen, Adair, & Lee, 2007; Shinn-Cunningham, 2008; Teshiba et al., 2012).

Limitationen und weiterführende Forschung

Erstens wurde die unterschiedliche Ähnlichkeit zwischen den verschiedenen Hinweisreiztypen und dem Zielreiz nur angenommen, diese ist jedoch nicht so augenscheinlich wie

zum Beispiel bei der Betrachtung von Farben. Obwohl sich die Effektgrößen von den Hinweisreiztypen Sprachgemisch und Rauschen nicht signifikant voneinander unterscheiden, lässt sich ein numerischer Trend erkennen, der auf einen geringeren Effekt im Hinweisreiztyp Rauschen deutet. Wie im Methodenteil beschrieben, unterschied sich das Rauschsignal allerdings auch in seiner Intensität von den beiden anderen Signalen. Eine mögliche Erklärung wäre, dass sich das Rauschsignal von den anderen beiden Signalen bezüglich der spektral-temporalen Signalstruktur und Intensität so stark abhebt, so dass Hinweis- und Zielreiz im Sinne der Objektbindungstheorie nicht mehr als zum selben episodischen Objekt gehörig wahrgenommen werden. Es ist jedoch fraglich, ob es sich bei der Formierung auditiver episodischer Objekte, um einen graduellen Prozess handelt. Da die in der vorliegenden Studie verwendeten Reize komplexe Schallsignale darstellen, könnten zukünftige Studien einfachere Signale verwenden, um die Ähnlichkeit gezielter manipulieren und somit ihren jeweiligen Einfluss besser untersuchen zu können.

Zweitens, wie oben schon angedeutet, wäre die Erhebung einer Kontrollbedingung mit Durchgängen ohne die auditive Präsentation von Hinweisreizen für die theoretische Interpretation interessant. Dadurch könnte festgestellt werden, ob die irrelevanten Hinweisreize die Aufmerksamkeit auf sich ziehen oder nicht. Außerdem könnte mittels des Vergleichs der Reaktionszeiten in Durchgängen ohne und mit Hinweisreizen überprüft werden, ob die Verarbeitung der Zielreize durch die auditive Präsentation von Hinweisreizen in spezifischen Bedingungen wirklich verbessert beziehungsweise verschlechtert wird.

Drittens beschreiben Carmel und Lamy (2015) die Dauer der Hinweisreize sowie die Integration von Hinweis- und Zielreiz in ein zugehöriges Objekt (Carmel & Lamy, 2014) als kritische Faktoren für das Auftreten von Objektaktualisierungskosten. Zukünftige Studien könnten durch systematische Variation der Hinweisreizdauer und der konsistenten beziehungsweise nicht-konsistenten Objektzugehörigkeit überprüfen, ob sich diese Zusammenhänge auch in der auditiven Modalität nachweisen lassen.

Viertens variierten die durchschnittlichen Reaktionszeiten und auch die Stärke der jeweiligen räumlichen Kongruenzeffekte zwischen den einzelnen Vpn sehr stark. Schon in früheren Studien zeigte sich eine größere Variabilität in auditiven im Vergleich zu visuellen räumlichen Hinweisreizeffekten (vgl. dazu Roberts et al., 2009; Roberts, Summerfield, & Hall, 2006). Auch in unseren Experimenten zeigte sich eine hohe Variabilität zwischen den Vpn, was in einigen Bedingungen sogar zu den Gesamtergebnissen widersprüchlichen Einzelleistungen führte (die Analyse der Zwischensubjektfaktoren Alter und Geschlecht ergab keine signifikanten Effekte). Daher ist es möglich, dass einzelne Vpn unterschiedliche Lösungsstra-

tegien für die Bearbeitung der Aufgabe verwendet haben. Weitere Studien sind nötig, um diese Mutmaßungen zu überprüfen. Eine mögliche, nicht wünschenswerte Alternativstrategie wäre beispielsweise, über das gesamte Experiment hinweg nur eines der beiden Ohren zu beachten. Dadurch, dass es nur zwei Antwortalternativen gab, wäre so eine zweckmäßige Lösung der Aufgabe möglich. Es ist jedoch fraglich, ob es sich dabei um eine effiziente Lösungsstrategie handelt. Außerdem hat keiner der Vpn auf Rückfrage eine derartige Vorgehensweise angegeben. Durch Hinzufügen weiterer Antwortalternativen könnte diese Art von Lösungsstrategie in zukünftigen Studien im Vorhinein ausgeschlossen werden.

Fünftens könnte in weiteren Studien durch eine zusätzliche Manipulation der Aufgabenstellung der Einfluss von absichtsgesteuerten Prozessen auf die Informationsverarbeitung untersucht werden. Ähnlich wie in Experiment 1 könnte beispielweise die Hälfte der Vpn auf die hohe und die andere Hälfte auf die tiefe Stimme achten. Der sprachliche Hinweisreiz bestehend aus der hohen Stimme sollte dann nur die Aufmerksamkeit auf sich ziehen, wenn die Vpn auch nach der hohen Stimme suchen. Damit könnte der Anteil von reiz- beziehungsweise absichtsgetriebenen Aufmerksamkeitsprozessen in dieser Sprachdiskriminationsaufgabe weiter differenziert werden.

Schlussfolgerung

Ähnlich wie in der visuellen Modalität ließen sich Hinweise darauf finden, dass die Verarbeitung eines auditiven Reizes von dessen Ähnlichkeit zu einem kurz zuvor präsentierten Reiz abhängt. Die vorliegende Studie legt nahe, dass sensorische Informationen sowohl hinsichtlich der räumlichen Position als auch der spektral-temporale Signalstruktur für die auditive Aufmerksamkeitssteuerung relevant sind. Im Speziellen verdeutlichen die hier gezeigten Hinweisreizeffekte die Wichtigkeit der Kombination von übereinstimmender Position und übereinstimmender Signaleigenschaften von Hinweis- und Zielreiz in einer Cocktail-Party ähnlichen Aufgabenstellung. Wie von einem plausibilitätsorientierten System zu erwarten, ist die Aufmerksamkeit am stärksten auf ein Signal in einem Signalgemisch, wenn kurz zuvor ein ähnlich klingendes Signal am gleichen Ort präsentiert wurde. Interessanterweise ist trotz hoher Rechenleistung die Extraktion eines unbekanntes Schallsignals aus einer Kombination mit anderen Schallsignalen, noch immer ein generelles Problem in der Signalverarbeitung (Deville, 2016). Ein besseres Verständnis der menschlichen Fähigkeit, die Aufmerksamkeit auf nur eine Schallquelle zu fokussieren, kann somit zur Verbesserung von Hörgeräten und Cochleaimplantaten sowie anderen elektronischen Geräten zur Sprach- oder Schallverarbeitung beitragen.

Literaturverzeichnis

- Ansorge, U., & Leder, H. (2017). Wahrnehmung und Aufmerksamkeit. In *Wahrnehmung und Aufmerksamkeit*, 1-17. Wiesbaden: Springer Fachmedien Wiesbaden GmbH.
- Best, V., Gallun, F. J., Ihlefeld, A., & Shinn-Cunningham, B. G. (2006). The influence of spatial separation on divided listening. *The Journal of the Acoustical Society of America*, *120*(3), 1506-1516.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Carmel, T., & Lamy, D. (2014). The same-location cost is unrelated to attentional settings: An object-updating account. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1465-1478.
- Carmel, T., & Lamy, D. (2015). Towards a resolution of the attentional-capture debate. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1772-1782.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484-1525.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975-979.
- Dalton, P., & Hughes, R. W. (2014). Auditory attentional capture: Implicit and explicit approaches. *Psychological Research*, *78*(3), 313-320.
- Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 180.
- Dalton, P., & Lavie, N. (2007). Overriding auditory attentional capture. *Attention, Perception, & Psychophysics*, *69*(2), 162-171.
- Deville, Y. (2016). Blind source separation and blind mixture identification methods. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 1-33.
- Eramudugolla, R., McAnally, K. I., Martin, R. L., Irvine, D. R., & Mattingley, J. B. (2008). The role of spatial location in auditory search. *Hearing Research*, *238*(1), 139-146.
- Fels, J., Oberem, J., Karnbach, B., Lawo, V., & Koch, I. (2013). Comparison of dichotic and binaural reproduction in an experiment on selective auditive attention. In *Proceedings of the International Conference on Acoustics*, Merano, Italien.
- Fels, J., Oberem, J., & Koch, I. (2016). Examining auditory selective attention in realistic, natural environments with a newly designed paradigm, in '22nd International Congress on Acoustics', Buenos Aires, Argentina.

- Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 847-858.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1030-1044.
- Haykin, S., & Chen, Z. (2005). The cocktail party problem. *Neural Computation*, 17(9), 1875-1902.
- Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, 5(1-2), 183-216.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175-219.
- Kanai, K., Ikeda, K., & Tayama, T. (2007). The effect of exogenous spatial attention on auditory information processing. *Psychological Research*, 71(4), 418-426.
- Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, 4(4), 138-147.
- Koch, I., Lawo, V., Fels, J., & Vorländer, M. (2011). Switching in the cocktail party: Exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 37(4), 1140.
- Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*, 55-96. Hillsdale, NJ: Erlbaum.
- Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, 80(1), 97-126.
- Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition*, 88(2), 223-242.
- Lamy, D., Alon, L., Carmel, T., & Shalev, N. (2015). The role of conscious perception in attentional capture and object-file updating. *Psychological Science*, 26(1), 48-57.
- Lamy, D., Leber, A., & Egeth, H. E. (2004). Effects of task relevance and stimulus-driven salience in feature-search mode. *Journal of Experimental Psychology: Human Perception and Performance*, 30(6), 1019-1031.
- Lawo, V., Fels, J., Oberem, J., & Koch, I. (2014). Intentional attention switching in dichotic listening: Exploring the efficiency of nonspatial and spatial selection. *The Quarterly Journal of Experimental Psychology*, 67(10), 2010-2024.

- Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476-490.
- Lupiáñez, J. (2010). Inhibition of return. In A. C. Nobre & J. T. Coull (Eds.), *Attention and Time*, 17-34. Oxford: Oxford University Press.
- Mayer, A. R., Harrington, D. L., Stephen, J., Adair, J. C., & Lee, R. R. (2007). An event-related fMRI study of exogenous facilitation and inhibition of return in the auditory modality. *Journal of Cognitive Neuroscience*, *19*(3), 455-467.
- Mayr, S., Buchner, A., Möller, M., & Hauke, R. (2011). Spatial and identity negative priming in audition: Evidence of feature binding in auditory spatial memory. *Attention, Perception, & Psychophysics*, *73*(6), 1710-1732.
- McDonald, J. J., & Ward, L. M. (1999). Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1234–1252.
- Mondor, T. A., & Amirault, K. J. (1998). Effect of same- and different-modality spatial cues on auditory and visual target identification. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 745.
- 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.
- Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 66-79.
- Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, *137*(3), 434.
- Park, J., & Kanwisher, N. (1994). Negative priming for spatial locations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(3), 613-623.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X: Control of Language Processes*, *32*(10), 531-556.
- Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416-442.

- Roberts, K. L., Summerfield, A. Q., & Hall, D. A. (2006). Presentation modality influences behavioral measures of alerting, orienting, and executive control. *Journal of the International Neuropsychological Society*, *12*(4), 485-492.
- Roberts, K. L., Summerfield, A. Q., & Hall, D. A. (2009). Covert auditory spatial orienting: An evaluation of the spatial relevance hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1178.
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182-186.
- 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.
- Teshiba, T. M., Ling, J., Ruhl, D. A., Bedrick, B. S., Peña, A., & Mayer, A. R. (2012). Evoked and intrinsic asymmetries during auditory attention: Implications for the contralateral and neglect models of functioning. *Cerebral Cortex*, *23*(3), 560-569.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77-99.

Abbildungsverzeichnis

Abbildung 1. Schematische Abbildung eines Experimentaldurchgangs.....	14
Abbildung 2. Durchschnittliche Reaktionszeiten der inkongruenten und kongruenten Durchgänge (Exp. 1).....	16
Abbildung 3. Durchschnittliche Reaktionszeiten aufgeteilt nach Hinweisreiztypen (Exp. 2).....	23
Abbildung 4. Durchschnittliche Trefferquoten aufgeteilt nach Hinweisreiztypen (Exp. 2)...	23
Abbildung 5. Interaktion der Reaktionszeiten zwischen den Faktoren <i>Hinweisreiztyp</i> x <i>Kongruenz</i> ohne die Durchgänge mit 1 oder 8 als Zielreiz.....	37
Abbildung 6. Interaktion der Reaktionszeiten in den Faktoren <i>Zielreizposition</i> x <i>Antwortastete</i>	38
Abbildung 7. Durchschnittliche Reaktionszeiten pro Experimentalsitzung aufgeteilt nach den drei verschiedenen Hinweisreiztypen.....	39
Abbildung 8. Durchschnittliche Trefferquoten pro Experimentalsitzung aufgeteilt nach den drei verschiedenen Hinweisreiztypen.....	40

Tabellenverzeichnis

Tabelle 1. Numerischer Trend der möglichen Unterschiede zwischen beiden Gruppen.....	16
Tabelle 2. Durchschnittliche Reaktionszeiten der Durchgänge mit spezifischen Zahlenwerten als Zielreiz.....	36

Abkürzungsverzeichnis

Vpn	Versuchspersonen
PRK	positiver räumlicher Kongruenzeffekt
NRK	negativer räumlicher Kongruenzeffekt
SOA	stimulus onset asynchrony (def.: Intervall zwischen Beginn des Hinweis- und Beginn des Zielreizes)
IOR	inhibition of return (def.: Hemmungseffekt einer zuvor beachteten räumlichen Position)
HZÄ	Hinweis-Zielreiz-Ähnlichkeit
ITD	interaural time difference (deut.: interaurale Laufzeitdifferenz)
ILD	interaural level difference (deut.: interaurale Pegeldifferenz)
NV	Normalverteilung
AS	Aufmerksamkeitsschablone

Anhang

Tabelle 2

Durchschnittliche Reaktionszeiten der Durchgänge mit spezifischen Zahlenwerten als Zielreiz

Zahlenwert des Zielreizes	k	M	SD
1	2244	1278	275
2	2374	1323	260
3	2313	1331	274
4	2302	1322	249
6	2115	1403	281
7	2398	1336	254
8	2333	1261	288
9	2301	1312	256

Anmerkung. k gibt die Anzahl an Durchgängen mit dem in der ersten Spalte angegebenen spezifischen Zahlenwert an (berücksichtigt werden nur die korrekten Durchgänge). In den letzten beiden Spalten sind die Mittelwerte und Standardabweichungen (jeweils in ms) der spezifischen Durchgänge angegeben.

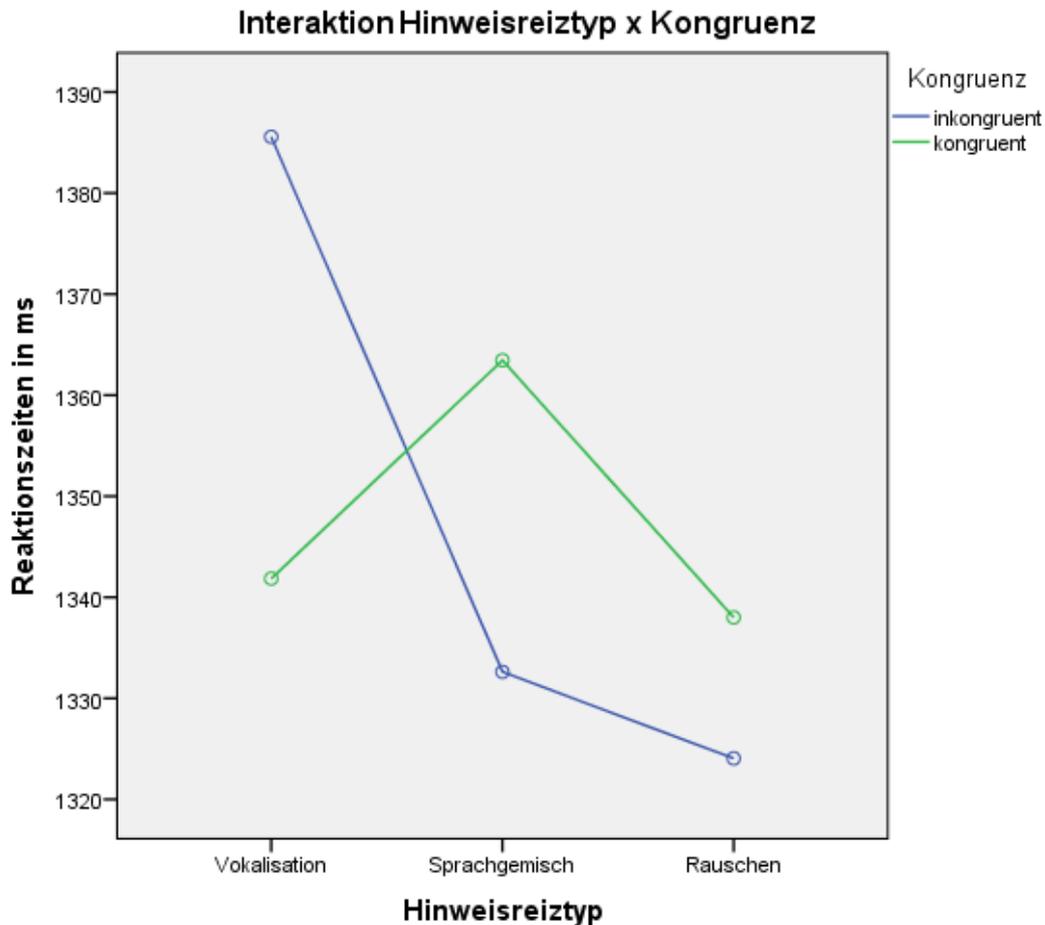


Abbildung 5. Interaktion der Reaktionszeiten zwischen den Faktoren *Hinweisreiztyp* x *Kongruenz* ohne die Durchgänge, bei denen die Zahl 1 oder 8 als Zielreiz vorkam. Die paarweisen Vergleiche ergaben bei allen drei Hinweisreiztypen die gleichen signifikanten Unterschiede wie unter Berücksichtigung aller korrekten Durchgänge, wenn auch mit etwas kleineren p -Werten (alle $p < .05$).

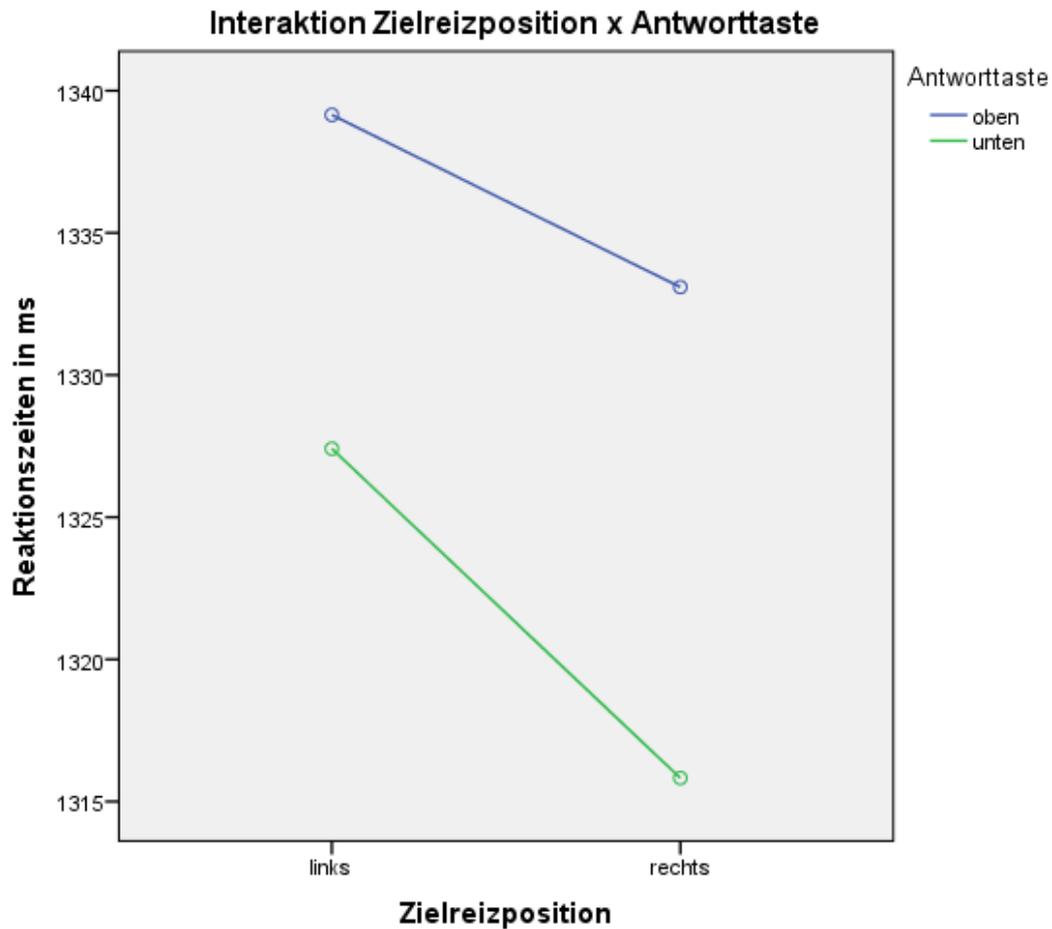


Abbildung 6. Interaktion der Reaktionszeiten in den Faktoren *Zielreizposition* x *Antworttaste*. Die dargestellte Interaktion sowie die Haupteffekte der 2 x 2 ANOVA der Reaktionszeiten waren nicht signifikant (alle $p > .206$).

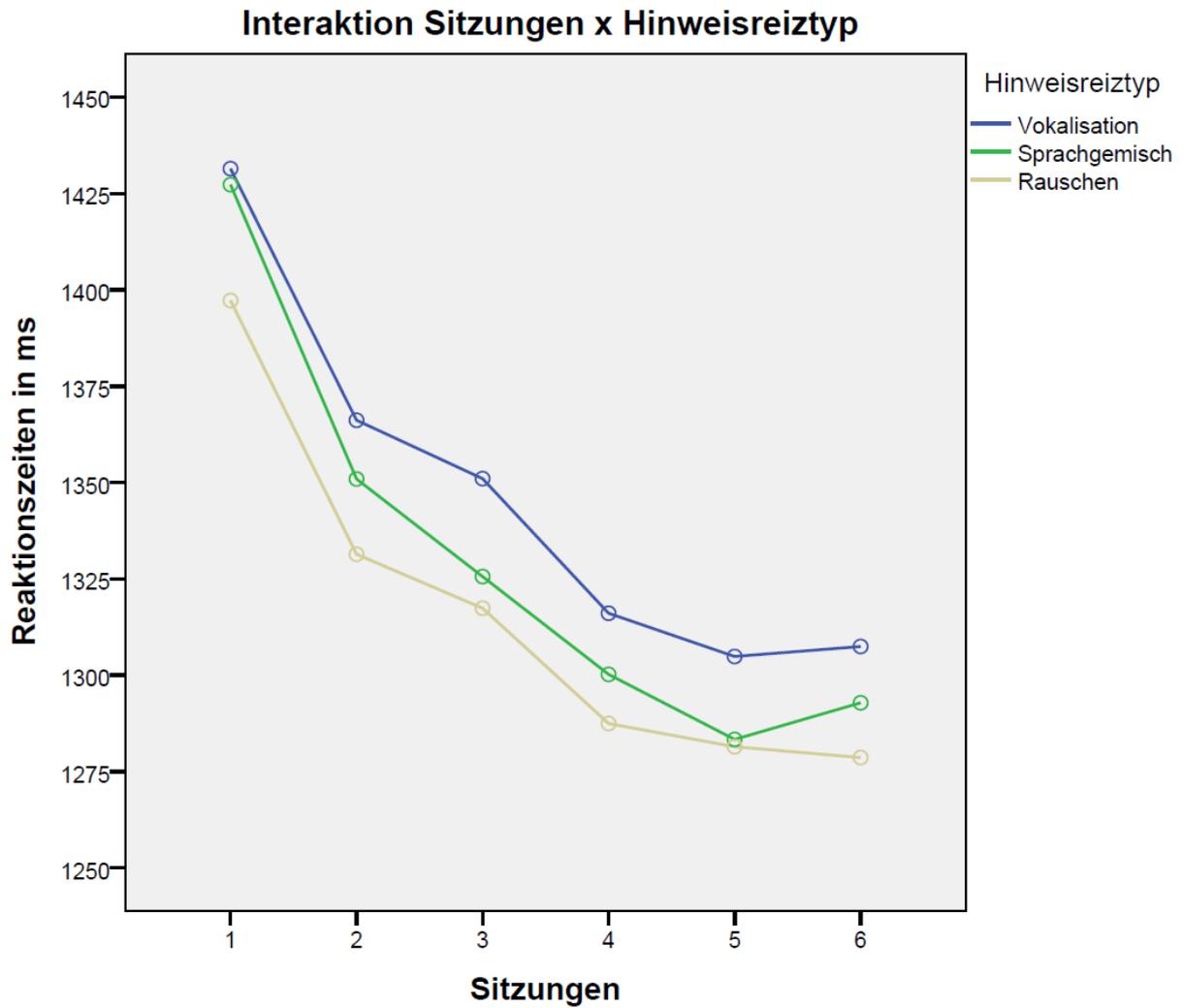


Abbildung 7. Durchschnittliche Reaktionszeiten pro Experimentalsitzung aufgeteilt nach den drei verschiedenen Hinweisreiztypen.

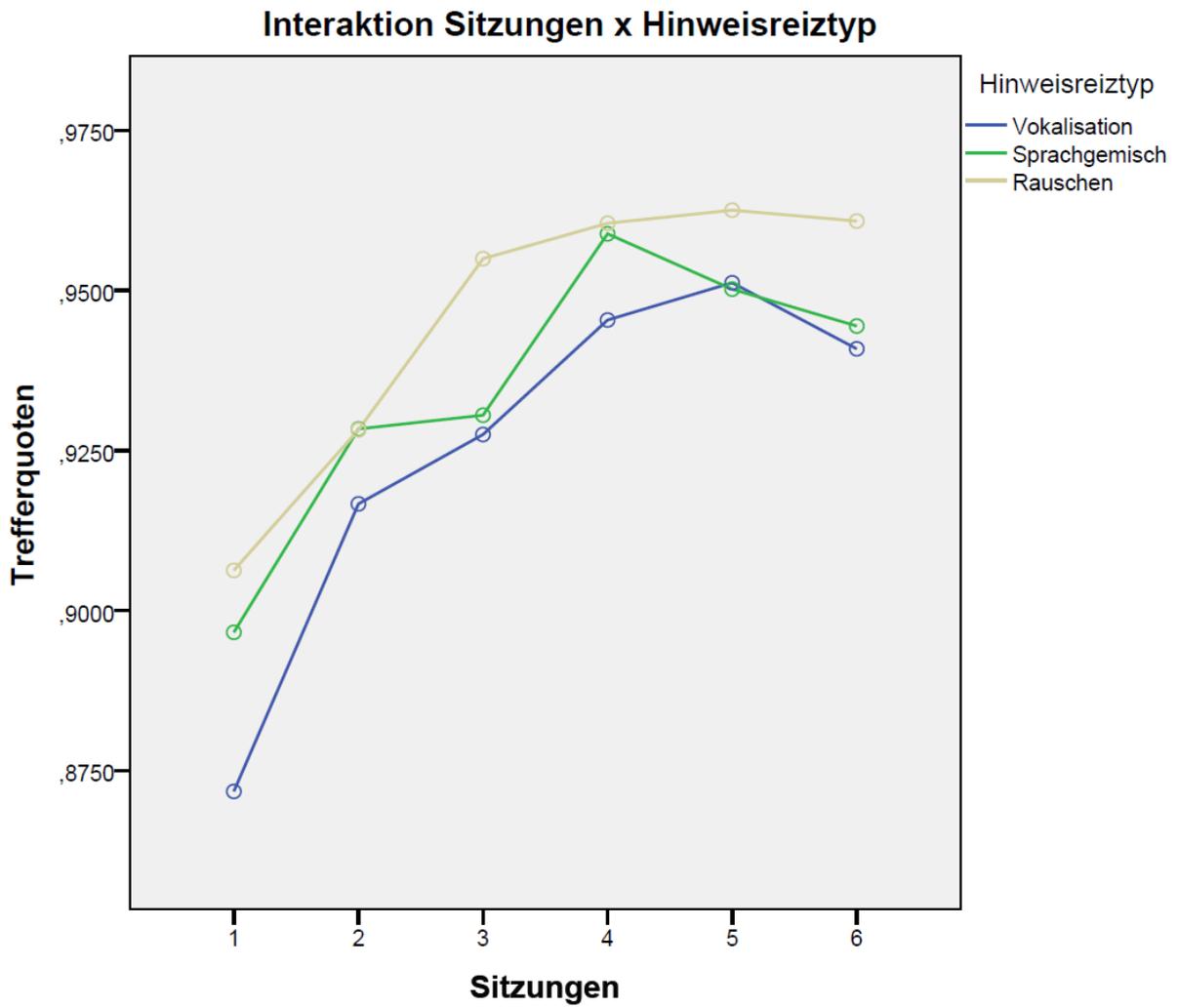


Abbildung 8. Durchschnittliche Trefferquoten pro Experimentalsitzung aufgeteilt nach den drei verschiedenen Hinweisreiztypen.