

# **Neuronale Mechanismen zeitabhängiger Prozesse bei audiovisueller Verarbeitung**

**Dissertation**

zur Erlangung des akademischen Grades

**doctor rerum naturalium  
(Dr. rer. nat.)**

genehmigt durch die Fakultät für Naturwissenschaften  
der Otto-von-Guericke Universität Magdeburg

von Dipl.-Psych. Daniel Bergmann

geb. am 22.02.1981 in Leipzig

Gutacher: PD Dr. Tömmе Noesselt  
Dr. Uta Noppeneу

Eingereicht am: 18.06.2009

Verteidigt am: 27.01.2010

# DANKSAGUNG

Ich danke PD Dr. Tömme Nösselt für seine hilfsbereite, intensive, lehrreiche und geduldige Unterstützung meiner Arbeit an den hier vorgestellten Projekten. Im Rahmen eines freundlichen und angenehmen Arbeitsambientes gelang es ihm stets, durch kritische Reflexionen erheblich zur Unterstützung und Entwicklung meiner wissenschaftlichen Arbeit beizutragen. Darüber hinaus möchte der gesamten Arbeitsgruppe des ZENIT und der Neurologischen Universitätsklinik, die zur Realisierung der hier vorgestellten Studien beigetragen haben danken.

Ich danke Frau Dr. Uta Noppeney für die Begutachtung meiner Arbeit.

# ZUSAMMENFASSUNG

Verarbeitungsmechanismen verschiedener Sinne wurden in der Vergangenheit intensiv untersucht. Dabei wurden jedoch meist einzelne Sinneskanäle betrachtet (z. B. Sehen oder Hören oder Tastsinn). Durch Ereignisse in der realen Welt werden hingegen meist mehrere Sinne gleichzeitig angesprochen. Darüber hinaus entstammen einzelne Wahrnehmungsanteile der über unterschiedliche Sinneskanäle eintreffenden Information ein und demselben Objekt oder Ereignis, wenn wir beispielsweise eine Person reden sehen und hören. Die Kombination der Informationen mehrerer Sinneskanäle bestimmt dabei perzeptuelle Urteile. So werden Informationen über das zeitliche Verhältnis von Ereignissen zweier Sinneskanäle perzeptuell nicht immer gemäß der tatsächlichen zeitlichen Relation beider Reize abgebildet. Vielmehr entscheiden implizite Annahmen, Stimuluseigenschaften und dynamische Prozesse des Wahrnehmungsapparates darüber, ob zwei Ereignisse als zeitlich synchron wahrgenommen werden. Studien zeigten, dass ein variables Zeitfenster existiert, innerhalb dessen bimodale Ereignisse zusammengefügt werden können. Diese Arbeit befasst sich mit den zugrundeliegenden neuronalen Prozessen audiovisueller Zeitwahrnehmung. So werden im Rahmen dreier fMRT-Experimente sowohl für semantisch belegte als auch für semantisch freie Reize neuronale Modulationen wahrgenommener Zeitrelationen identifiziert. Dabei wird Augenmerk auf die Beteiligung unisensorischer und multisensorischer Kortices gelegt. Mit Hilfe eines Adaptationsparadigmas werden neuronale Grundlagen plastischer Mechanismen audiovisueller Zeitwahrnehmung untersucht, um stimulusabhängige und perzeptuelle Effekte wahrgenommener Zeitrelationen zu differenzieren. Schließlich werden anhand ökologisch valider Sprachreize unterschiedliche Zeitperzepte erfasst und innerhalb multisensorischer Hirnareale funktionell zugeordnet. Die Ergebnisse zeigen unterschiedliche Aktivierungsmuster für wahrgenommene Synchronizität im Vergleich zu Asynchronizität. Darüber hinaus konnte gezeigt werden, dass unterschiedliche kortikale Netzwerke an einer dynamischen Anpassung zeitlicher Wahrnehmungsprozesse in Richtung Asynchronizität bzw. Synchronizität beteiligt sind. Schließlich wurde demonstriert, dass multisensorische Gebiete funktionell nach Zeitperzepten getrennt werden können. Insgesamt deuten diese Ergebnisse darauf hin, dass stabile Wahrnehmungen synchroner und asynchroner audiovisueller

Zeitinformationen an getrennte multisensorische Hirnareale gekoppelt zu sein scheinen und Aktivierungsmuster dabei chronotopisch angeordnet sind.

# SUMMARY

In the past, mechanisms of sensory and perceptual processes have been studied intensively by both psychologists and neuroscientists. However, recent research investigated mostly just single senses (i.e. vision or audition or touch, whereas real-world events often stimulate more than one single modality concurrently. Moreover, parts of information arising from one and the same object need to be joint together across distinct sensory modalities, as when we both feel and see an object in our hand or both see and hear someone speak. The way how information is bound across modalities determines perceptual judgments. For example, temporal perception of external bimodal events does not coercively correspond to their actual physical relations. Rather, perception of subjective simultaneity is determined by implicit knowledge, stimulus properties and dynamics of the perceptual apparatus. Previous studies demonstrated the existence of a temporal window that can flexibly be widened or tightened. Bimodal events will be temporally bound as long as they both fall within this temporal window. This thesis investigates the neural basis of audiovisual perception of temporal relations. Therefore, by means of three fMRI-studies neural modulations of perceived temporal relations were identified when participants judged both semantic and non-semantic stimuli. In doing so, influences of unisensory and bimodal cortices were investigated. Furthermore, neural correlates of temporal plasticity during audiovisual perception were studied with the aid of an adaptation paradigm. This allowed for separating perceptual and stimulus-driven effects of audiovisual temporal perception. Finally, temporal percepts of audiovisual speech were captured by ecologically valid stimuli and the according functional basis was located within bimodal areas. Results revealed distinct activation patterns for perceptual synchrony compared to asynchrony. Moreover, it was shown that different cortical networks are involved in the establishing of subjective synchrony vs. asynchrony by means of dynamic adaptation. In the end, it was demonstrated that bimodal areas can be separated functionally by different time percepts. To conclude, these results indicate that audiovisual temporal information of stable synchrony percepts and stable asynchrony percepts seem to be related to enhanced activity in distinct multisensory cortices. Beyond, the relevant activation patterns were found to be chronotopically arranged.

# INHALTSVERZEICHNIS

<b>DANKSAGUNG</b>	<b>II</b>
<b>ZUSAMMENFASSUNG</b>	<b>III</b>
<b>SUMMARY</b>	<b>V</b>
<b>INHALTSVERZEICHNIS</b>	<b>VI</b>
<b>ABKÜRZUNGSVERZEICHNIS</b>	<b>XI</b>
<b>0 AUFBAU DER ARBEIT</b>	<b>1</b>
<b>1 ALLGEMEINE EINLEITUNG</b>	<b>1</b>
<b>1.1 Zeitwahrnehmung</b>	<b>2</b>
1.1.1 Zeitwahrnehmung verschiedener Sinnesmodalitäten	6
1.1.2 Neuronale Grundlagen der Zeitwahrnehmung	6
1.1.3 Zusammenfassung der Zeitwahrnehmung	10
<b>1.2 Multimodale Integration</b>	<b>10</b>
1.2.1 Informationsverarbeitungstheorien multimodaler Integration	12
1.2.2 Statistische Ansätze multimodaler Verarbeitung	15
1.2.3 Dynamische Systeme	18
1.2.4 Neuronale Mechanismen multimodaler Integration	20
1.2.5 Synästhesie	25
1.2.6 Studien zur multimodalen Integration	28
1.2.6.1 Psychophysik	28
1.2.6.2 Neuronale Grundlagen	32
1.2.6.3 Tierstudien	32
1.2.6.4 Humanstudien	33
<b>1.3 Fragestellungen</b>	<b>38</b>
1.3.1 Neuronale Grundlagen audiovisueller Zeitwahrnehmung einfacher Reize	39
1.3.2 Neuronale Grundlagen audiovisueller Zeitlichkeitsadaptation	40
1.3.3 Funktionelle Korrelate von Zeitprozessen audiovisueller Sprache	40
<b>2 ALLGEMEINE METHODIK</b>	<b>42</b>

<b>2.1</b>	<b>Psychophysische Maße</b>	<b>42</b>
2.1.1	JND und PSS	42
<b>2.2</b>	<b>Funktionelle Magnetresonanztomographie</b>	<b>44</b>
<b>2.3</b>	<b>Analyse der fMRT-Daten</b>	<b>46</b>
2.3.1	Das Allgemeine Lineare Modell	46
2.3.2	Auswertung mit SPM	47
2.3.2.1	Vorverarbeitung	47
2.3.2.2	Gruppenanalyse	48
<b>3</b>	<b>EXPERIMENTE</b>	<b>49</b>
<b>3.1</b>	<b>Experiment 1</b>	<b>49</b>
3.1.1	Einführung	49
3.1.2	Vorexperiment 1	50
3.1.2.1	Methodik	50
3.1.2.2	Fragestellung und Hypothesen	53
3.1.2.3	Ergebnisse des Vorexperiments 1	55
3.1.2.4	Synchronizitätsurteile	55
3.1.2.5	JNDs und PsSS	56
3.1.2.6	Bewertung der Ergebnisse des Vorexperiments 1	58
3.1.3	Hauptexperiment	59
3.1.3.1	Methodik	59
3.1.3.2	Ergebnisse	65
3.1.3.2.1	Verhaltensdaten	65
3.1.3.3	fMRT-Ergebnisse	66
3.1.3.3.1.1	Lokalisationsdurchgang	66
3.1.3.3.1.2	Experimentelle Kontraste	67
3.1.3.3.1.2.1	<i>Physikalische (Un-)Gleichzeitigkeit</i>	67
3.1.3.3.1.2.2	<i>Licht vor Ton vs. Ton vor Licht (physikalisch)</i>	68
3.1.3.3.1.2.3	<i>(A)synchrones Antwortverhalten</i>	69
3.1.3.3.1.2.4	<i>(A)Synchrone Wahrnehmung</i>	70
3.1.3.3.1.2.5	<i>Synchrone Wahrnehmung von Licht vor Ton, Synchronizität und Ton vor Licht</i>	71
3.1.3.3.1.2.6	<i>Asynchrone Wahrnehmung von Licht vor Ton, Synchronizität und Ton vor Licht</i>	74
3.1.3.4	Diskussion der Ergebnisse des fMRT-Experiments 1	75
3.1.3.4.1	Verhaltensdaten	76

---

3.1.3.4.2	fMRT-Daten	77
3.1.3.4.2.1	Physikalische (A)Synchronizität	78
3.1.3.4.2.1.1	(A)synchrone Perzeption	85
3.1.3.4.2.2	(a)synchrone Antwortverhalten	95
3.1.3.4.2.3	Zusammenfassung	96
<b>3.2</b>	<b>Experiment 2</b>	<b>98</b>
3.2.1	Einführung	98
3.2.2	Fragestellungen und Hypothesen zu den Voruntersuchungen	100
3.2.3	Vorexperiment 2.1	101
3.2.3.1	Methodik	101
3.2.3.2	Ergebnisse des Vorexperiments 2.1	104
3.2.3.2.1	Synchronizitätsurteile	104
3.2.3.2.2	JNDs und PsSS	107
3.2.3.3	Bewertung der Ergebnisse von Vorexperiment 2.1	110
3.2.3.3.1	Synchronizitätsurteile	110
3.2.3.3.2	JNDs und PsSS	110
3.2.4	Vorexperiment 2.2	115
3.2.4.1	Methodik	115
3.2.4.2	Ergebnisse des Vorexperiments 2.2	117
3.2.4.2.1	Synchronizitätsurteile	117
3.2.4.2.2	JNDs und PsSS	119
3.2.4.3	Bewertung der Ergebnisse von Vorexperiment 2.2	120
3.2.5	Vorexperiment 2.3	125
3.2.5.1	Methodik	125
3.2.5.2	Ergebnisse des Vorexperiments 2.3	126
3.2.5.2.1	Synchronizitätsurteile	126
3.2.5.2.2	JNDs und PsSS	128
3.2.5.3	Bewertung der Ergebnisse von Vorexperiment 2.3	129
3.2.6	Zusammenfassung der Vorexperimente 2.1, 2.2 und 2.3	134
3.2.7	Hauptexperiment	135
3.2.7.1	Einführung	135
3.2.7.2	Methodik	136
3.2.7.3	Ergebnisse	142
3.2.7.3.1	Verhaltensdaten	142
3.2.7.3.1.1	Synchronizitätsurteile	142

---

3.2.7.3.1.2	JNDs und PsSS	144
3.2.7.3.2	fMRT-Daten	147
3.2.7.3.2.1	Experimentelle Kontraste	147
3.2.7.3.2.1.1	3.>1. Minute bei asynchroner Exposition innerhalb perzeptueller Synchronizität	147
3.2.7.3.2.1.2	1.>3. Minute bei asynchroner Exposition innerhalb perzeptueller Asynchronizität	148
3.2.7.3.2.1.3	Vergleich der Expositionsphase außerhalb perzeptueller Areale	149
3.2.7.3.2.1.4	BOLD-Signale	150
3.2.7.4	Diskussion	152
3.2.7.4.1	Verhaltensdaten	152
3.2.7.4.2	fMRT-Daten	154
<b>3.3</b>	<b>Experiment 3</b>	<b>161</b>
3.3.1	Einführung	161
3.3.2	Vorexperiment 3	162
3.3.2.1	Methodik	162
3.3.2.2	Ergebnisse des Vorexperiments	165
3.3.2.3	Bewertung der Ergebnisse des Vorexperiments 3	166
3.3.3	Hauptexperiment	167
3.3.3.1	Methodik	167
3.3.3.2	Ergebnisse	173
3.3.3.2.1	Verhaltensdaten	173
3.3.3.2.2	fMRT-Daten	174
3.3.3.2.2.1	Lokalisationsdurchgang	174
3.3.3.2.2.2	Experimentelle Kontraste	175
3.3.3.2.2.2.1	Physikalische (A)Synchronizität	175
3.3.3.2.2.2.2	(A)Synchrone Perception	175
3.3.3.2.2.2.3	Kohärente Perzepte	176
3.3.3.2.2.2.4	Region-of-Interest-Analyse	178
3.3.3.2.2.2.5	Psychophysiologische Interaktionen	182
3.3.3.3	Diskussion	183
<b>4</b>	<b>ALLGEMEINE DISKUSSION</b>	<b>188</b>
<b>5</b>	<b>LITERATURVERZEICHNIS</b>	<b>193</b>
<b>6</b>	<b>ANHANG</b>	<b>222</b>

<b>LEBENS LAUF</b>	<b>256</b>
<b>EIGENSTÄNDIGKEITSERKLÄRUNG</b>	<b>258</b>

# ABKÜRZUNGSVERZEICHNIS

<b>Abb.</b>	Abbildung
<b>ACC</b>	anteriorer cingulärer Kortex
<b>ALM</b>	allgemeines lineares Modell
<b>Ang</b>	Gyrus angularis
<b>aSTS</b>	anteriorer Sulcus temporalis superior
<b>AV</b>	abhängige Variable
<b>BOLD</b>	blood oxygen level dependent
<b>cd</b>	Candela
<b>cm</b>	Zentimeter
<b>dB (SPL)</b>	Dezibel (Schalldruckpegel)
<b>DLPFC</b>	dorsolateraler Präfrontalkortex
<b>EEG</b>	Elektroenzephalographie
<b>EKP</b>	ereigniskorrelierte Potentiale
<b>EPI</b>	echo planar image
<b>FFA</b>	fusiformes Gesichtsareal
<b>FG</b>	Gyrus fusiformis
<b>fMRT</b>	funktionelle Magnetresonanztomographie
<b>HG</b>	Heschl'scher Gyrus
<b>HRF</b>	hemodynamic response function
<b>Hz</b>	Hertz
<b>IFG</b>	Gyrus frontalis inferior
<b>IOG</b>	Gyrus okzipitalis inferior
<b>IPL</b>	inferiorer Parietallappen
<b>IPS</b>	Sulcus intraprietalis
<b>ISI</b>	Interstimulusintervall
<b>ITG</b>	Gyrus temporalis inferior
<b>JND</b>	just noticeable difference
<b>LED</b>	light emitting diode
<b>LG</b>	Gyrus lingualis
<b>LG</b>	Gyrus lingualis

---

<b>LIP</b>	laterale intraparietale Region
<b>MEG</b>	Magnetenzephalographie
<b>MFG</b>	Gyrus frontalis medius
<b>min</b>	Minuten
<b>MNI</b>	Montreal Neurological Institute
<b>MOG</b>	Gyrus okzipitalis medius
<b>ms</b>	Millisekunden
<b>mSTS</b>	mittlerer Sulcus temporalis superior
<b>MT</b>	medial temporales Areal
<b>MTG</b>	Gyrus temporalis medius
<b>PC</b>	personal computer
<b>PET</b>	Positronen-Emissionstomographie
<b>PFC</b>	Präfrontalkortex
<b>PPI</b>	psychophysiologische Interaktion
<b>PSS</b>	point of subjective simultaneity
<b>pSTS</b>	posteriorer Sulcus temporalis superior
<b>PT</b>	Planum temporale
<b>ROI</b>	region of interest
<b>SAPS</b>	Stimulusanzahl pro Sequenz
<b>SC</b>	Colliculi superiores
<b>Sek.</b>	Sekunden
<b>SMG</b>	Gyrus supramarginalis
<b>SOA</b>	stimulus onset asynchrony
<b>SOG</b>	Gyrus okzipitalis superior
<b>STG</b>	Gyrus temporalis superior
<b>STS</b>	Sulcus temporalis superior
<b>Tab.</b>	Tabelle
<b>TE</b>	echo time
<b>TI</b>	inversion time
<b>TMS</b>	Transkranielle Magnetstimulation
<b>TOJ</b>	Temporal order judgment

---

<b>TPO</b>	temporo-parieto-okzipitale Region
<b>TR</b>	repetition time
<b>UV</b>	unabhängige Variable
<b>VIP</b>	Ventral intraparietale Region
<b>VLDFC</b>	ventrolateraler Präfrontalkortex

# 0 Aufbau der Arbeit

Die vorliegende Arbeit befasst sich mit dem Einfluss wahrnehmungs- und stiumulusabhängiger Prozesse auf die Verarbeitung audiovisueller Gleichzeitigkeit und deren neuronalen Grundlagen. Die Arbeit beginnt mit einer Einführung in die Thematik der Zeitverarbeitung und multimodalen Integration. Dabei werden zuerst konzeptuelle Hintergründe beider Phänomene in einem Überblick dargestellt und anschließend aktuelle Befunde dieser Forschungsbereiche besprochen, anhand derer sich Fragestellungen ergeben, die im Rahmen dieser Arbeit behandelt werden sollen. In Kapitel 2 werden Methoden vorgestellt, mit deren Hilfe die hier präsentierten Daten ermittelt wurden. In Kapitel 3 werden drei aufeinander aufbauende Untersuchungen, die mittels funktioneller Magnetresonanztomographie (fMRT) durchgeführt wurden inklusive ihrer Voruntersuchungen aufgeführt und besprochen. Abschließend werden die vorgestellten Befunde in Kapitel 4 zusammenfassend diskutiert.

## 1 Allgemeine Einleitung

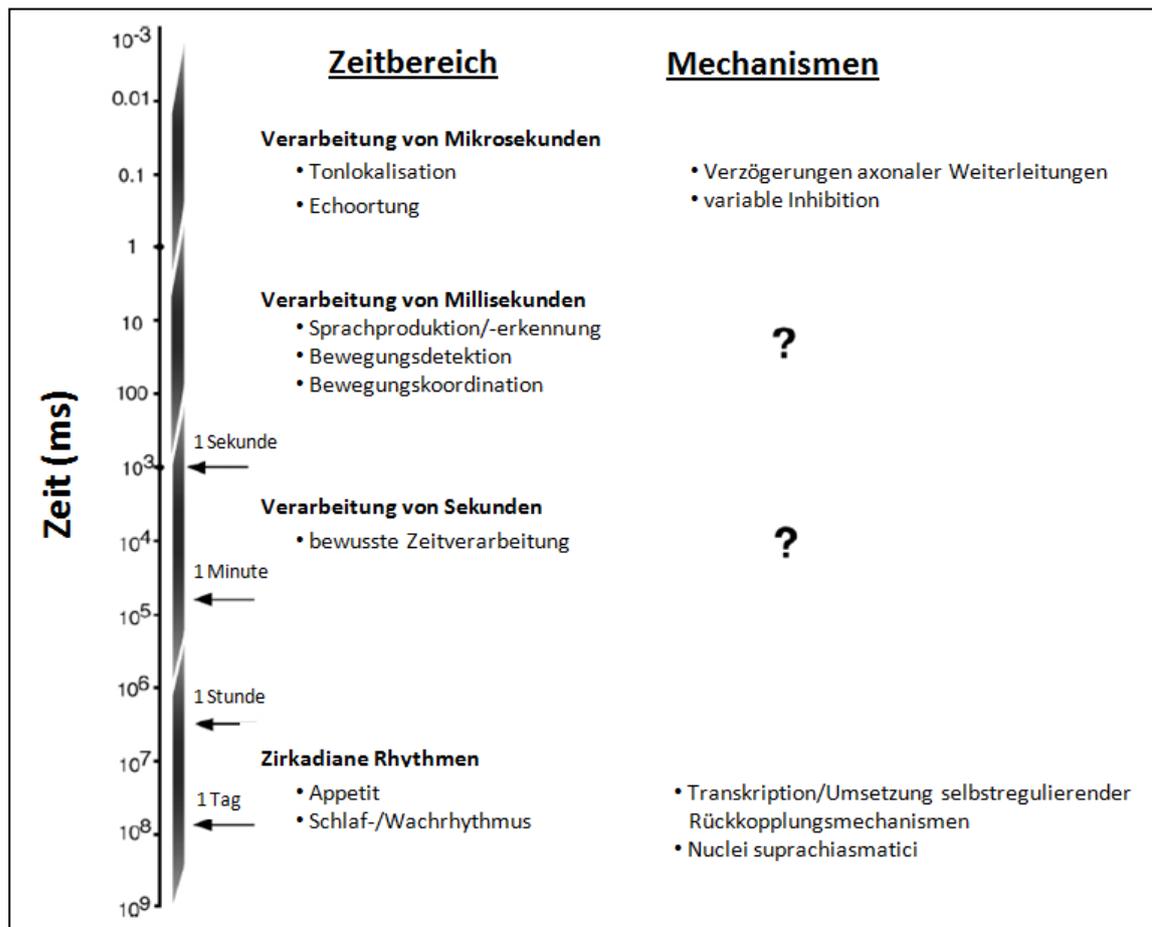
Ereignisse im Alltag erstrecken sich über unterschiedliche Zeitspannen und haben unterschiedliche zeitliche Abfolgen. Dementsprechend müssen kognitive und neuronale Prozesse zeitliche Abfolgen und Zeitspannen analog zur Umgebung adäquat abbilden. Die Verarbeitung von zeitlicher Information ist ein zentraler Teil unseres Moment-zu-Moment-Funktionierens. Im Gegensatz zu anderen Wahrnehmungsmodalitäten (z. B. Hören oder Sehen) verfügen wir jedoch über kein spezialisiertes sensorisches System (Auge oder Ohr) oder einen Rezeptor, der die Erfahrung von Zeit erklären kann (Fraise, 1984; Lalonde & Hannequin, 1999). Nichtsdestotrotz können Menschen Zeitintervalle bis auf wenige Millisekunden diskriminieren (Grondin & Rousseau, 1991; Rousseau et. al, 1983). Auch gelingt das Vorbereiten zeitlich präziser Bewegungen wie unter anderem beim Tennis Spielen adäquat. Jedoch ist Zeitwahrnehmung variabel und kann beispielsweise durch Drogen (Meck, 1983) oder cerebrale Läsionen (Harrington et al., 1998a; Verstichel, 2003) beeinflusst werden. Zudem können Zeitwahrnehmungsprozesse interindividuell variieren. So kam es 1796 am königlichen Observatorium in Greenwich

zu Beobachtungsfehlern beim Erfassen von Himmelsereignissen: mangels exakter Zeitmesser wurden damals vorbeiziehende Himmelskörper nach der ‚Auge-und-Ohr-Methode‘ registriert um später deren Rektaszension zu berechnen (Mollon & Perkins, 1996). Diese Methode sah vor, dass der Astronom die Anzahl von Klickgeräuschen zählte, die er in einer Hand erzeugte, während der Himmelskörper an der Teleskoplinse vorbeizog. Die Zeitangaben David Kinnebrooks, dem damaligen Assistenten des königlichen Astronomen Nevil Maskelyne, wichen teilweise um 800 Millisekunden (ms) von den Schätzungen seines Vorgesetzten ab. Diese Abweichungen führten zum Begriff des ‚persönlichen Ausgleichs‘, der die Zeit beschreibt, die zu einer individuellen Messung addiert bzw. subtrahiert werden muss, um sie in Einklang mit den Messungen anderer zu bringen. Das Interesse von Astronomen am ‚persönlichen Ausgleich‘ führte zu Untersuchungen von Reaktionszeiten und Urteilen über zeitliche Abfolgen, die das erste Labor experimenteller Psychologie in Leipzig dominierten (Wundt, 1887).

Die vorliegende Arbeit untersucht mit Hilfe psychophysischer Methoden und bildgebender Verfahren die Funktionsmechanismen zeitlicher Integration von Licht und Ton und deren neuronale Grundlagen.

## 1.1. Zeitwahrnehmung

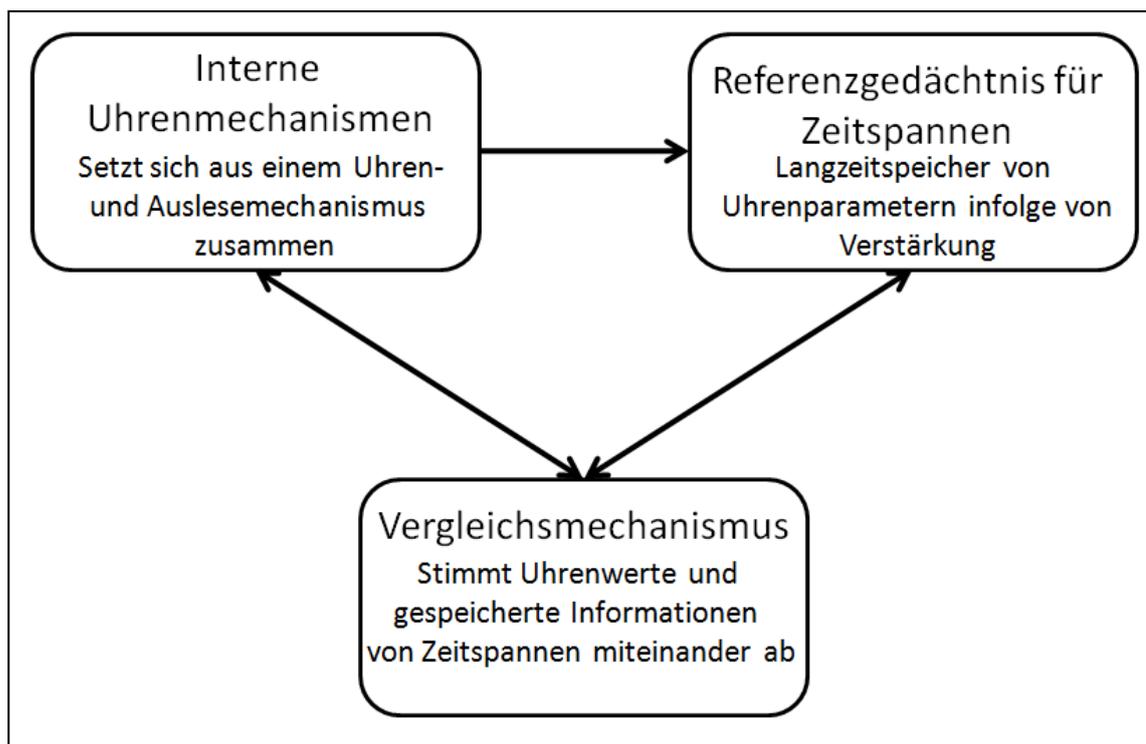
Das oben aufgeführte Beispiel verdeutlicht, dass die objektive Wahrnehmung von Zeit ohne externe Hilfen nicht einheitlich bzw. zuverlässig gelingt. Dennoch existieren mehrere zeitliche Ebenen, die Menschen verarbeiten können (s. Abb. 1). Sie reichen von zirkadianen Oszillationen, die im Bereich von mehreren Stunden schwanken bis hin zu Prozessen, die im Mikrosekunden-Bereich angesiedelt sind. Während zirkadiane Schwankungen wie der Schlaf-Wach-Rhythmus von subthalamischen Regionen gesteuert werden (z. B. King & Takahashi, 2000), können Laufzeitunterschiede von Mikrosekunden über Unterschiede axonaler Weiterleitung ausgewertet werden (Carr, 1993). Eindeutige Mechanismen, die Zeitprozesse von Millisekunden und Sekunden verarbeiten sind bisher nicht eindeutig definiert worden, jedoch existieren einige Modelle, die versuchen, Zeitwahrnehmung innerhalb dieses Bereichs zu erklären.



**Abbildung 1.** Unterschiedliche Zeitskalen, die vom menschlichen Wahrnehmungssystem verarbeitet werden. Dabei reichen die detektierbaren Zeitbereiche von Mikrosekunden bis hin zu täglichen physiologischen Schwankungen. Dies ermöglicht einerseits das Entdecken von Zeitunterschieden, die ein Tonreiz benötigt, um von einem Ohr zum anderen zu gelangen, und andererseits das Umsetzen biologischer Rhythmen. Zirkadiane Zyklen werden durch biochemische Prozesse kontrolliert. Die Verarbeitung von Zehntel- und Hundertstelsekunden ist hochkomplex und Grundlage motorischer und sprachbezogener Prozesse. Das Einschätzen von Zeitspannen oder -intervallen gliedert sich in den Bereich von Sekunden und Minuten ein und wird als bewusste Wahrnehmung von Zeit betrachtet. *Abbildung nach Mauk & Buonomano (2004).*

Alle Modelle der Zeitwahrnehmung verfügen über drei Komponenten: eine innere Uhr, eine Gedächtniskomponente und eine Entscheidungs- bzw. Vergleichskomponente (nach Matell & Meck, 2000; s. Abb. 2). Dabei bringt die Uhrkomponente in Abhängigkeit von Zeitprozessen Veränderungen mit sich, die von einem anderen Mechanismus aufgenommen oder evaluiert werden. Die Werte des Uhrprozesses werden im Referenzgedächtnis abgelegt, so dass für nachfolgende Ereignisse Vorhersagen getroffen werden und können. Der Vergleichsapparat stellt die aktuellen Messwerte den zuvor im Referenzgedächtnis gespeicherten Werten gegenüber. Wenn beide Werte eine Gleichartigkeitsschwelle überschreiten, wird ein Beobachter eine entsprechende Antwort

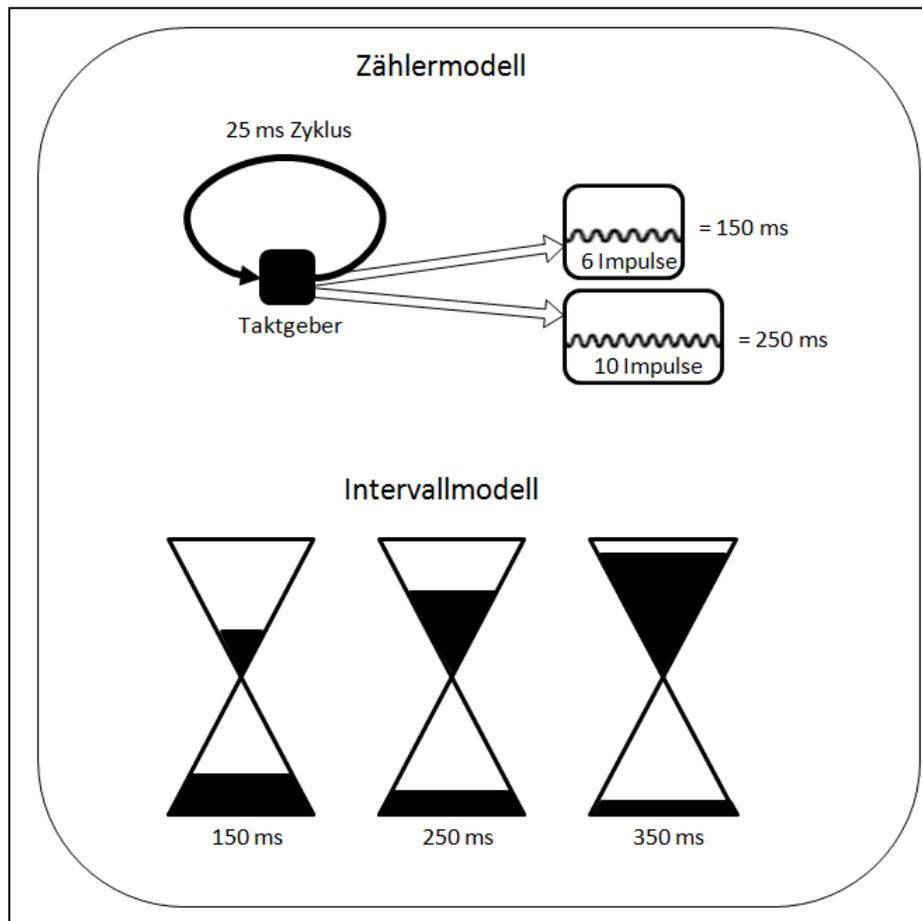
geben (z. B. über das Ablufen eines bestimmten Zeitereignisses). Unsere Wahrnehmung von Zeit hängt jedoch auch von dem Ausmaß der Aufmerksamkeit ab, die dem Vergehen von Zeit geschenkt wird (Coull et al., 2004). Wenn Aufmerksamkeit geteilt wird, tritt eine systematische Verkürzung der subjektiven Dauer auf. Dies bedeutet, dass Impulse des Uhrenmechanismus verloren gehen (Brown, 1985). Daher resultiert eine subjektive Repräsentation von Zeit aus dem Zusammenspiel zwischen zeitgebenden und aufmerksamkeitsbezogenen Mechanismen.



**Abbildung 2.** Interne Uhrenmodelle benötigen drei primäre Komponenten: eine Uhr, ein Gedächtnis und einen Vergleichsmechanismus. Die Uhrkomponente verändert ihr Substrat über die Zeit, das je nach Modell diese Veränderungen registriert. Infolge von verstärkenden Prozessen werden diese Werte im Referenzgedächtnis abgelegt und können für zukünftige Erwartungen verwandt werden. Die Vergleichskomponente gleicht Uhrenprozesse mit gespeicherten Daten über Zeitspannen ab und initiiert gegebenenfalls eine Antwort über das Ablufen einer Zeitdauer. *Abbildung nach Matell & Meck (2000).*

Über die Funktionsweise der inneren Uhr gibt es verschiedene Modelle: Taktgeber-Speichermodelle, Prozess-Verfallmodelle und Oszillator/Koinzidenz-Entdeckungsmodelle. Taktgeber-Speicher Modelle legen einen kontinuierlich laufenden Taktgeber und einen Speicher zugrunde, die als Zähler fungieren. Der Speicher, der zu Beginn eines Signals aktiviert wird, zählt die Schläge des Taktgebers, die sich über eine Dauer erstrecken (z. B. Ivry, 1996; Matell & Meck, 2000; s. Abb. 2). Prozess-Verfallmodelle

zeichnen das Abfallen neuronaler Aktivität in bezug auf den Beginn eines Signals auf (Grossberg & Schmajuk, 1987; Staddon & Higa, 1999). Das Oszillator/Koinzidenz-Entdeckungsmodell erfasst spezifische Kombinationen periodischer neuronaler Ereignisse, die bei Signalbeginn aktiviert werden. Zeit wird hier enkodiert, indem diese kombinierten Aktivierungsmerkmale speziellen Laufzeiten zugeordnet werden (Miall, 1989).



**Abbildung 3.** Verschiedene Mechanismen, die Zeitinformationen repräsentieren können: Zählermodelle (oben im Bild) gehen von einem Taktgeber aus, der Daten an einen Zähler abgibt. Je höher die Informationsdichte (Anzahl Impulse), desto größer wird eine Zeitspanne geschätzt. Intervall-basierte Modelle (unten im Bild) implizieren, dass unterschiedliche Intervalle von verschiedenen Elementen dargestellt werden, die jeweils eine spezifische Zeitspanne abbilden. *Abbildung nach Ivry (1996).*

In Anbetracht verschiedener Zeitskalen, dem das menschliche Wahrnehmungssystem ausgesetzt ist, wird vermutet, dass nicht nur ein einziger Uhrenmechanismus existiert, sondern es im Sinne eines intervallbasierten Mechanismus

verschiedene Kreisläufe für spezifische Zeitintervalle gibt (Ivry, 1996; s. Abb. 3). Darüber hinaus wird diskutiert, ob Zeitwahrnehmung zentral verarbeitet wird oder vielmehr verteilt ist und von der Wahrnehmungsmodalität abhängig ist, die beurteilt wird (Mauk & Buonomano, 2004).

### **1.1.1 Zeitwahrnehmung verschiedener Sinnesmodalitäten**

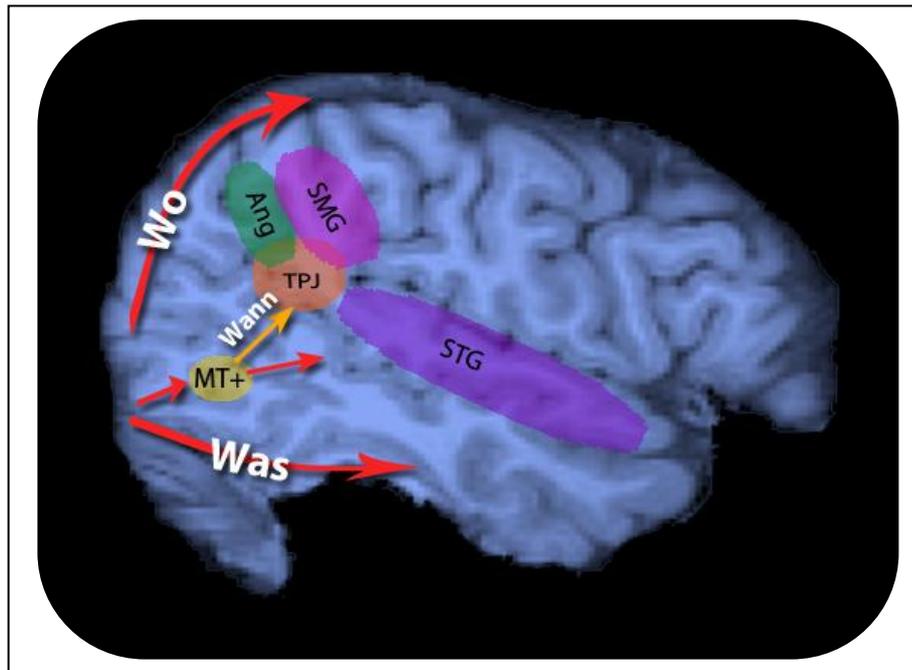
Durchschnittlich beträgt die Geschwindigkeit des visuellen Taktgebers 92 % der auditorischen Geschwindigkeit (Wearden et al., 1998), wobei die Unterschiede bei Kindern größer sind als bei Erwachsenen (Droit-Volet et al., 2007). Wearden et al. (1998) schlagen einen Schaltmechanismus zwischen Taktgeber und Speicher vor, der Beginn und Ende visueller Reize und den Vergleich mit gespeicherten Zeitspannen anders verwertet als bei auditorischen Reizen und damit zu unterschiedlichen Wahrnehmungen der Zeitspanne gleichlanger Stimuli beider Modalitäten führt. Penney und Tourret (2005) gehen von einem amodalen (modalitätsübergreifenden) Taktgeber und amodalen Gedächtnisrepräsentationen von Zeitspannen aus, aber postulieren modalitätsspezifische Speicher und Schaltmechanismen, die die Unterschiede in bezug auf die Zeitwahrnehmung zwischen beiden Modalitäten erklären.

Van Wassenhove et al. (2008) zeigten, dass die Darbietung kombinierter inkongruenter visueller Reize in bezug auf Zeitspanne und Salienz zusammen mit auditorischen zu einer Dehnung bzw. einem Zusammenziehen der subjektiv wahrgenommenen Zeitspannen führen konnte. Dies belegt, dass sich die Wahrnehmungsmodalitäten hinsichtlich der Zeitwahrnehmungsprozesse gegenseitig beeinflussen. Die Autoren mutmaßten, dass ein multimodaler Kontrasteffekt vorliegende Informationen über Zeitunterschiede zwischen beiden Modalitäten vergrößert.

### **1.1.2 Neuronale Grundlagen der Zeitwahrnehmung**

Untersuchungen, die einen zentralen Zeitmechanismus vermuten, ordnen diesen anhand von Ergebnissen bildgebender Verfahren und Läsionsstudien zu Zeitdiskriminationsaufgaben einem fronto-parietalen Netzwerk (Belin et al., 2002; Harrington, 1998a), den Basalganglien (Coull & Nobre, 2008; Harrington, 1998b; Ivry, 1996) und dem Cerebellum (Buonomano & Karmarkar, 2002; Ivry, 1996; Ivry & Spencer,

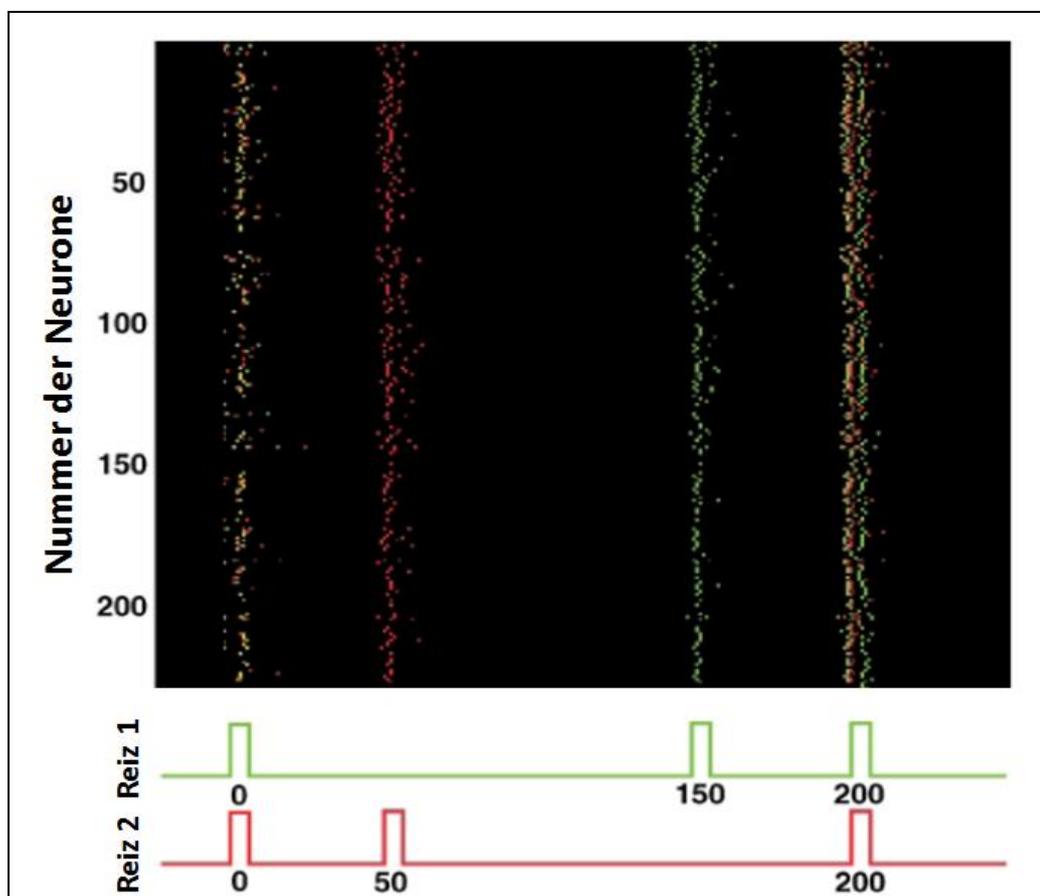
2004; Rao et al., 2001) zu. Diese Areale wurden durch das Einschätzen der Dauer von Ereignissen bzw. von Abständen zwischen zwei Ereignissen stärker neuronal moduliert als während einer zeitunabhängigen Vergleichsaufgabe, bzw. zeigten Patienten mit spezifischen Störungen der erwähnten Gebiete schlechtere Leistungen hinsichtlich dieser Aufgaben als gesunde Probanden. Andere Befunde weisen darauf hin, dass der Kortex analog zur Anordnung des visuellen Systems (z. B. Matlin & Foley, 1997) eine hierarchische Ordnung größer werdender zeitlicher rezeptiver Felder aufweist, in denen von frühen zu höheren kortikalen Arealen zunehmend längere Zeitspannen verarbeitet werden (Hasson et al., 2008). Anhand von Läsionsstudien (Battelli et al., 2007) und Erkenntnissen aus fMRT-Studien (Rao et al., 2001) wurde die Bedeutung des inferioren Parietallappens für die Wahrnehmung zeitlicher Aspekte von Stimuli hervorgehoben. Dieser war in die Diskrimination zeitlicher Intervalle sowohl von auditorischen als auch von visuellen Reizen involviert. Die Autoren gehen daher davon aus, dass neben dorsalem ‚Wo‘- und ventralem ‚Was‘-Pfad, die jeweils Informationen über räumliche und objektbezogene verarbeiten (Mishkin & Ungerleider, 1982) ein zeitlicher ‚Wann‘-Pfad existiert, der sich von visuellen Gebieten über den medialen Temporallappen (MT+), inferioren Parietallappen (IPL) und Sulcus temporalis superior (STS) erstreckt (s. Abb. 4). Die Beteiligung dieses dritten Pfades an der Zeitwahrnehmung lässt sich auch daraus erklären, dass in ihm visuelle Bewegungsprozesse kodiert werden (Beer et al., 2002; Howard et al., 1996; Watson et al., 1993; Zeki et al., 1991), die sich, im Gegensatz zu statischen Objekten in der Zeit entfalten. Battelli et al. (2007) schließen weiterhin, dass auditorische und visuelle Zeitprozesse denselben neuronalen Mechanismen unterliegen könnten. Im Widerspruch dazu gehen einige Autoren anhand der Grundlage von Befunden, die zeigen, dass visuelle Reize als kürzer wahrgenommen werden als auditorische (Droit-Volet et al., 2007; Efron, 1970; Penney et al., 2000; Penney & Turret, 2005; Wearden et al., 1998), davon aus, dass der Uhrenmechanismus für beide Modalitäten unterschiedlich getaktet ist.



**Abbildung 4.** Der Wann-Pfad verarbeitet zeitabhängige Informationen lateralisiert in der rechten Hemisphäre. Gemäß bisheriger Studien werden Informationen aus dem visuellen Kortex entlang eines dorsalen (räumliche Wahrnehmung) oder ventralen (Objekterkennung) Pfades verarbeitet. Es wird vermutet, dass ein weiterer Pfad vom visuellen Kortex ausgehend Zeitinformationen nutzt, um Objekte zu identifizieren (beispielsweise das Erscheinen oder Verschwinden von Objekten). Dies soll vor allem unter der Beteiligung der temporo-parietalen Verbindung (TPJ) geschehen. Darüber umfasst gemäß Battelli et al. (2007) der ‚Wann-Pfad‘ den Gyrus supramarginalis (SMG), den Gyrus angularis (Ang), den Gyrus temporalis superior (STG) und das Gebiet MT+ (medio-temporale Region). *Abbildung nach Battelli et al. (2007).*

Buonomano und Merzenich (1995) schlagen in Widerspruch zu einem zentralisierten Uhrenmechanismus vor, dass zeitliche Prozesse keine zentralen Verrechnungseinheiten benötigen, sondern Zeitinformationen vielmehr durch ein kortikales Netzwerk zugänglich werden, das durch ihm inhärente Eigenschaften Zeitinformationen verarbeitet, da durch vorangegangene Informationsflüsse zeitabhängige Veränderungen in bezug auf den Zustand des Netzwerkes geschehen (Mauk & Buonomano, 2004). Diese zeitabhängigen Eigenschaften könnten auf neuronaler Ebene durch kurzzeitige synaptische Plastizität oder inhibitorische postsynaptische Potentiale gegeben sein, die die Weiterleitung nachfolgender Aktionspotentiale beeinflussen. Aufgrund dieser zeitabhängigen Veränderungen wird das Netzwerk 50, 100 und 200 Millisekunden (ms) nach Eintreffen eines Reizes jeweils in einem anderen

Zustand sein (s. Abb. 5). Dasselbe Ereignis, das nach 100 ms auftritt, wird das Netzwerk in einem anderen Zustand erreichen, als das Ereignis, das nach 0 ms eintraf. Dies führt dazu, dass derselbe Reiz je nach Ereignisvergangenheit unterschiedliche Subpopulationen von Neuronen aktiviert. Diese Unterschiede der Populationsaktivität, die der zweite und erste Reiz hervorriefen, können genutzt werden, um den zeitlichen Abstand der zwei Ereignisse zu enkodieren. Im Vergleich zu den oben aufgeführten Modellen werden die zeitabhängigen Prozesse nicht bestimmten Zeitabständen zugeordnet, sondern die unterschiedlichen Zeitzustände drücken vielmehr die Wahrscheinlichkeit aus, mit der verschiedene Neurone aktiviert werden (s. Abb. 5).



**Abbildung 5.** Zustandsabhängiges Modell der Intervallerkennung. Die Punktwolken zeigen verschiedene Neurone an, die auf Intervallsequenzen der Art ‚lang-kurz‘ (grün) und ‚kurz-lang‘ (rot) antworten. Neurone, die auf beide Intervalltypen antworten sind in gelb dargestellt. Jeder Reizimpuls aktiviert eine Neuronenpopulation und wird dadurch Kurzzeitplastizität hervorrufen. Dies führt dazu, dass das Netzwerk beim Eintreffen des zweiten Reizimpulses, je nach dem, ob er den Beobachter nach 150 (grün) oder 50 ms (rot) erreicht, in einem anderen Zustand ist. Bei beiden Reizarten kommt der dritte Impuls nach 200 ms an, das Netzwerk wird jedoch zu diesem Zeitpunkt, je nach Stimulustyp, in einem jeweils anderen Zustand sein. Dies ermöglicht dem Wahrnehmungssystem auf denselben Reizimpuls unterschiedlich zu reagieren. *Abbildung nach Mauk & Buonomano (2004).*

### 1.1.3 Zusammenfassung der Zeitwahrnehmung

Die Fähigkeit des Menschen, sich zeitlichen Gegebenheiten der Umgebung anzupassen, kann eine fundamentale Voraussetzung für erfolgreiches Verhalten sein. Es wird vermutet, dass die Fähigkeit, Ereignisse auch über Wahrnehmungsmodalitäten hinaus zeitlich zu verbinden (z. B. die Verbindung von Licht und Ton) und Gleichzeitigkeit bzw. Ungleichzeitigkeit zwischen ihnen zu entdecken von Geburt an verfügbar ist und damit eine angeborene Eignung darstellt, auf der assoziatives Lernen basiert (Bahrick, 1992; Spelke, 1979). Darüber hinaus impliziert die Analyse von zeitlichen Abläufen Schlüsse über Ursache und Wirkung (Fonlupt, 2003), ermöglicht, Sprachinformationen zu verarbeiten (Xu et al., 2005; Zacks et al., 2001) und ist Voraussetzung für soziale Interaktionen und damit verbundene empathische Fähigkeiten („Theory of Mind“; Saxe et al., 2004).

Die Erfahrung von Zeit scheint ein wichtiges Prinzip zu sein, um Sinneseindrücke zu ordnen und komplexes Verhalten und Kognitionen zu ermöglichen (Dawson, 2004). Oben wurden Modelle beschrieben, die Hinweise darauf geben, wie Abstände zwischen Ereignissen oder Zeitspannen von Reizen verarbeitet werden könnten. Alltägliche Ereignisse erfordern jedoch die zeitliche Verbindung oder Trennung von Ereignissen, die aus unterschiedlichen Sinnesmodalitäten hervorgehen. Die Fähigkeit des Menschen, Ereignisse in einer Sinnesmodalität mit einer anderen Modalität zeitlich abzugleichen, machten sich interessanterweise auch die britischen Astronomen zunutze (s. Kapitel 1), anstatt auf das Vermögen eines einzigen Sinneskanals zu vertrauen. Die Interaktionen zweier Modalitäten in Bezug auf Zeitwahrnehmungsphänomene wurden kurz besprochen. Grundlegende Konzepte der multimodalen Wahrnehmung sollen im nächsten Abschnitt dargestellt werden.

## 1.2. Multimodale Integration

Der zeitliche Abstand zweier Ereignisse, die unterschiedliche Wahrnehmungssysteme ansprechen (z. B. Licht und Ton) kann darüber entscheiden, ob diese beiden Ereignisse perzeptuell einem einheitlichen Ursprung zugeordnet werden, oder als getrennt wahrgenommen werden. Dabei steht das Gehirn bei multisensorischen Reizen vor dem besonderen Problem, Informationen, die zunächst in unterschiedlichen

---

kortikalen Arealen repräsentiert sind, zusammenzuführen. Die Erfahrung beispielsweise, während eines Gewitters, den Blitz vor dem dazugehörigen Donner als unabhängiges visuelles Ereignis wahrzunehmen, ist allgemein bekannt. Diese Diskrepanz zwischen unserer Wahrnehmung und einem physikalisch synchronen, wenn auch entfernten, multisensorischem Ereignis, entsteht in diesem Beispiel durch physikalisch bedingte Zeitunterschiede hinsichtlich der Zeit, die beide Stimuli benötigen, um Auge und Ohr zu erreichen. Licht bewegt sich weit schneller durch den Raum als Schall: mit 300.000.000 im Vergleich zu 330 Metern pro Sekunde. Unterschiede in bezug auf den Zeitpunkt des Eintreffens zweier Reize unterschiedlicher Modalität treten auch schon bei Ereignissen auf, die sehr viel näher bei uns stattfinden und dennoch bemerken wir diese Zeitdifferenzen selten. Dies liegt unter anderem daran, dass die mechanische Transduktion von Schallwellen im Ohr weniger Zeit benötigt als die chemische Transduktion von Licht in der Retina (King & Palmer, 1985). Diese physikalischen und biophysikalischen Zeitunterschiede von Licht und Schall heben sich gegenseitig auf, wenn sich Reize ungefähr zehn Meter von uns entfernt, am sogenannten ‚Horizont der Gleichzeitigkeit‘, befinden (Pöppel, 1985). Da die meisten audiovisuellen Ereignisse nicht aus dieser „optimalen“ Entfernung wahrgenommen werden, befassen sich Psychologen, Neurowissenschaftler und Philosophen schon lange mit der Frage, warum die Wahrnehmung multisensorischer Synchronizität in unserem Alltagsleben vorherrscht (Dennett & Kinsbourne, 1992; King, 2005).

Das Phänomen, dass zwei Ereignisse unter bestimmten Kontexten perzeptuell miteinander verbunden werden und unter anderen Gegebenheiten nicht, wird als das Bindungsproblem bezeichnet. In diesem Rahmen wird der Frage nachgegangen, wie neuronale Prozesse die Wahrnehmung einer Einheit einzelner Ereignisse erklären können, wenn daneben mehrere andere mögliche Wahrnehmungszustände existieren.

Die Verbindung multimodaler Ereignisse zu einem einheitlichen Perzept hängt nicht nur von zeitlichen (z. B. Dixon & Spitz, 1980), sondern auch von räumlichen (z. B. Vroomen & De Gelder, 2004), semantischen (z. B. Calvert, 2004) und stimuluspezifischen Merkmalen (z. B. Stein et al., 1996) der Ereignisse ab.

Traditionellerweise wurde angenommen, dass Informationen aus verschiedenen Sinneskanälen abstrahiert, zusammengeführt und in höheren kortikalen Gebieten integriert werden. Dieser Ansatz wurde z. B. von Piaget (1974) vorgeschlagen; er

mutmaßte, dass Kinder langsam durch Interaktionen mit ihrer Umwelt hierarchisch übergeordnete Zuordnungsfunktionen zwischen angeborenen, voneinander abgegrenzten Modalitäten erlernen. Dieser Standpunkt geht auf Helmholtz (1896) zurück, der bereits glaubte, dass frühkindliche Sinne von Geburt an getrennt sind und nur durch Erfahrung miteinander verknüpft werden können. Neuere Erkenntnisse lassen im Gegensatz dazu darauf schließen, dass multimodale Ereignisse bereits in frühen subkortikalen (Stein & Meredith, 1993) und unisensorischen (z. B. Noesselt et al., 2007) Gebieten verarbeitet werden und Integration mit Kommunikation unisensorischer Areale untereinander (Falchier et al., 2002) und mit höheren Arealen einhergeht (Driver & Noesselt, 2008). Allerdings gibt es anhand einer Tierstudie Hinweise darauf, dass Zellen in subkortikalen Strukturen, die im ausgewachsenen Tier multisensorische Informationen verarbeiten durch postnatale Einflüsse höherer Areale erst heranreifen (Wallace, 2004). Dies bedeutet, dass schnelle bzw. frühe Verarbeitung multisensorischer Informationen nicht angeboren ist, sondern erlernt wird, wobei bis heute umstritten ist, inwieweit subkortikale Modulationen durch den Kortex gesteuert werden (Jiang et al., 2001; 2007; s. 1.2.6.3).

Prinzipiell bietet multisensorische Integration den Vorteil, dass unterschiedliche Sinne Informationen über ein Ereignis liefern. So wird Sprache beispielsweise leichter verstanden, wenn die dazugehörigen Lippenbewegungen gesehen werden können (Sumbly & Pollack, 1954). Dies verändert das Signal-zu-Rausch-Verhältnis um bis zu 20 dB. Zudem werden schwache auditorische und visuelle Reize schneller entdeckt, wenn sie gemeinsam dargeboten werden (Stein & Meredith, 1993). In der Literatur werden verschiedene Ansätze verfolgt, um diese multimodalen Verarbeitungsmechanismen zu erklären. Diese sollen in den folgenden Abschnitten erläutert werden.

### **1.2.1 Informationsverarbeitungstheorien multimodaler Integration**

Die im Folgenden aufgeführten Konzepte multimodaler Integration resultieren aus Studien, in denen den Sinneskanälen inkongruente Informationen dargeboten wurden. Ein prominentes Beispiel diesbezüglich ist der McGurk-Effekt (McGurk & MacDonald, 1976): während einem Probanden ein Gesicht dargeboten wurde, dessen Lippen die Silbe ‚ga‘ formten, wurde die Silbe ‚ba‘ auditorisch dargeboten. Durch crossmodale Interaktionseffekte wurde diese Stimuluskombination als ‚da‘

wahrgenommen. Dies belegt, dass visuelle Erfahrungen die auditorische Wahrnehmung beeinflussen bzw. umgekehrt und warf die Frage auf, unter welchen Bedingungen die Sinnesmodalitäten einander in welchem Ausmaß beeinflussen.

Um die intersensorischen Verzerrungseffekte, wie sie beispielsweise beim McGurk-Effekt auftreten, zu erklären, wurden einige Verarbeitungskonzepte multisensorischer Integration erörtert: Die *Modalitätseignungshypothese* (Welch & Warren, 1980) geht davon aus, dass verschiedene Sinne unterschiedlich gut geeignet sind, bestimmte Reizmerkmale zu verarbeiten. So kann das auditorische System zeitliche Muster gut voneinander differenzieren, während das visuelle System räumliche Reize gut zuordnen kann. Rubinstein und Gruenberg (1971) zeigten beispielsweise, dass ein perzeptueller Vorteil hinsichtlich der Beurteilung hochfrequenter auditorischer Reize im Vergleich zu gleichartigen visuellen Stimuli besteht. Eine weitere Studie zeigte außerdem, dass Probanden ein Flimmern von Leuchtdioden mit einer bestimmten Frequenz in der Frequenz eines simultan präsentierten auditorischen Flatterns mit einer anderen Frequenz wahrnahmen (Shipley, 1964). Dies bedeutet, dass auditorische Perzeption hinsichtlich der Beurteilung zeitlicher Muster zuverlässiger ist als die visuelle; sie ist dafür besser ‚geeignet‘. Dieses als ‚auditory driving‘ in die Literatur eingegangene Phänomen konnte in zahlreichen Studien belegt werden (z. B. auch Welch et al., 1986). Untersuchungen, die anhand von Prismenbrillen haptische mit visuellen Informationen konkurrieren ließen ergaben, dass eigene Körperteile näher an der Position wahrgenommen wurden, an der sie durch die Prismenbrille gesehen wurden, als an der gefühlten Position (s. Welch & Warren, 1980 für eine Zusammenfassung). Dies belegt die Überlegenheit des visuellen Systems in bezug auf räumliche Informationen im Vergleich zum taktilen und auditorischen System.

Die *Hypothese gerichteter Aufmerksamkeit* (z. B. Welch & Warren, 1980) geht hingegen davon aus, dass diejenige Modalität über andere dominiert, die stärker attendiert wird. Dabei können Unterschiede bezüglich der Salienz zwischen beiden Modalitäten oder Prädispositionen des Beobachters einer Modalität mehr Aufmerksamkeit zu schenken als einer anderen, von Bedeutung sein. Aus psychophysischen Untersuchungen gibt es verschiedene Hinweise darauf, dass eine explizite Anweisung, eine bestimmte Modalität zu attendieren zu anderen Urteilen führten, als wenn Probanden beide Modalitäten gleichzeitig beachteten. So beeinflusste

beispielsweise die Anzahl kurzer auditorischer Signaltöne die wahrgenommene Anzahl visueller Blitze, nicht jedoch, wenn Probanden nur die auditorischen Reize beachteten (Andersen et al., 2004).

Shams et al. (2000, 2002) konnten zeigen, dass ein kurzer Lichtreiz, wenn er mit zwei oder mehr kurzen Tonsignalen kombiniert wird, perzeptuell verdoppelt wird. Das heißt die Illusion eines zusätzlichen Lichtblitzes entsteht durch den Einfluss auditorischer Signale. Dieser Teilung des visuellen Reizes konnte jedoch keine Bindung mehrerer dargebotener visueller Reize gegenübergestellt werden, die mit einem Tonsignal kombiniert wurden. Der Modalitätseignungshypothese gemäß würde auch im letzten Fall der auditorische Reiz die visuelle Perception beeinflussen. Da dies nicht nachgewiesen werden konnte, formulierten Shams et al. (2002) die *Unterbrechungshypothese*. Dementsprechend würde der Einfluss der Modalität, die unterbrochene Reize verarbeitet, größer sein als der derjenigen, die kontinuierliche verarbeitet. Neuere Studien deuten jedoch darauf hin, dass diese Unterbrechungshypothese zu kurz greift, da in weiteren Studien auch eine ‚Teilung‘ des visuellen Reizes berichtet wurde (Mishra et al., 2007; Watkins et al., 2008).

Ein weiterer Ansatz geht davon aus, dass die dominante Modalität nicht allein durch ihre Eignung definiert ist, sondern dadurch, wie zuverlässig die jeweiligen Informationen verarbeitet werden können (*Informationsreliabilitätshypothese*; Warren, 1979). Dabei kann die Reliabilität einer Modalität durch kognitive Faktoren beeinflusst werden, wenn Probanden beispielsweise manipuliert werden, einem Sinneskanal mehr zu vertrauen als einem anderen. Sie kann jedoch auch durch Stimuluseigenschaften beeinflusst werden. So fanden Bresciani und Ernst (2007) heraus, dass der Einfluss der auditorischen auf die taktile Wahrnehmung sinkt, wenn die Intensität des akustischen Signals verringert wird.

Neuere Erkenntnisse lassen darauf schließen, dass die oben erwähnten Hypothesen nicht exklusiven Erklärungsanspruch haben, sondern dass alle aufgeführten Faktoren je nach Stimuluskonfiguration und Aufgabenanforderungen dazu beitragen können, dass eine Modalität relativ zu einer anderen dominiert (Andersen et al., 2004).

## 1.2.2 Statistische Ansätze multimodaler Verarbeitung

In Anlehnung an das Konzept der Informationsreliabilitätshypothese wurde untersucht, wie die Schätzungen über Informationen unterschiedlicher Modalitäten miteinander kombiniert werden. Dazu wurden die Verteilungen unimodaler Urteile denen multimodaler Urteile gegenüber gestellt. Es stellte sich heraus, dass die beobachteten Verteilungen multimodaler Urteile nachträglich aus einer probabilistischen Kombination beider unimodaler Verteilungen berechnet werden konnten. So wurde in einer Untersuchung, in der audiovisuelle Reize lokalisiert werden mussten, gezeigt, dass in Abhängigkeit von der Deutlichkeit (Zuverlässigkeit) visueller Reize sowohl die auditorische als auch die visuelle Modalität die aktuelle Wahrnehmung dominierte (Alais & Burr, 2004). Je nach Zuverlässigkeit, die eine Modalität kennzeichnete, wurden die beiden Reize unterschiedlich miteinander verknüpft. Ernst und Banks (2002) zeigten, dass multisensorische Ereignisse nach der Maximum-Likelihood-Schätzung miteinander kombiniert werden. Durch dieses Prinzip gelingt es dem Wahrnehmungsapparat offenbar, Informationen mehrerer Sinneskanäle zu nutzen, indem die jeweiligen Schätzungen unterschiedlich gewichtet werden, um die Gesamtreliabilität einer (multimodalen) Schätzung zu erhöhen. Durch eine statistisch optimierte Bewertung multimodaler Information gelingt es dem Wahrnehmungssystem demzufolge, in ihrer Zuverlässigkeit unimodalen Informationen überlegene Urteile zu fällen.

Das menschliche Wahrnehmungssystem integriert widersprüchliche Reize aus zwei Modalitäten jedoch nicht fortwährend zuverlässig. Vielmehr ist es für ein optimales Handeln notwendig, dass sich das Gehirn ständig flexibel anpasst und durch das Erkennen von Diskrepanzen zwischen Modalitäten an Änderungen der Umgebung adaptieren kann. Vorhersagen über den Nutzen einer Verbindung bzw. Trennung multimodaler Reize können anhand des Bayes'schen Theorems getroffen werden (Ernst, 2006; Kersten et al., 2004; Yuille & Bühlhoff, 1996)

Bayes'sche Schlussfolgerungen stellen einen Weg dar, Unsicherheiten in bezug auf Umweltereignisse zu modellieren, indem Vorwissen mit beobachteten, sensorischen Gegebenheiten kombiniert wird, um wiederum Schlüsse über die wahrscheinlichste Interpretation der Außenwelt zu ziehen. Das Bayes-Theorem nutzt bedingte Wahrscheinlichkeiten, um über Rückwärtsinduktion von Ereignissen auf Ursachen zu

schließen. So stellt dieses Konzept eine Möglichkeit dar, zu erklären, unter welchen Bedingungen beispielsweise ein auditorisches und ein visuelles Ereignis in Anbetracht eines bestimmten Vorwissens perzeptuell ein und derselben Quelle zugeordnet werden. Das Modell geht von einem idealen Beobachter aus, der versucht, bestmögliche Schätzungen über den Ursprung physikalischer Ereignisse durchzuführen. Um ein interpretierbares Perzept herzustellen, berücksichtigt dieser Beobachter verschiedene Auftretenswahrscheinlichkeiten: die sogenannte posteriore Wahrscheinlichkeitsverteilung eines jeden möglichen Zustands eines Ereignisses unter der Voraussetzung einer bestimmten sensorischen Stimulation; die sogenannte a-priori-Wahrscheinlichkeit eines jeden möglichen Zustands eines Ereignisses, bevor ein Stimulus wahrgenommen wurde; und eine Plausibilitätsannahme über das Auftreten eines Stimulus unter der Voraussetzung eines jeden möglichen Zustands der äußeren Umgebung (z. B. Ernst, 2006; Ma & Pouget, 2008; Shams et al., 2005). Dabei stellt die posteriore Wahrscheinlichkeitsverteilung eine aktuelle Szene dar, in der der Wahrnehmungsapparat stimuliert wird, die a-priori-Wahrscheinlichkeitsverteilung Wissen über Gesetzmäßigkeiten möglicher Stimulusbeschaffenheiten, aber auch von Zuordnungen zwischen Stimuli verschiedener Modalitäten, und die Verteilung der Plausibilitätsannahmen Wissen darüber, wie Szenen oder multisensorische Ereignisse entstehen können. Wenn Wahrnehmungen anhand des Bayes-Theorems beschrieben werden, wird folglich davon ausgegangen, dass (multisensorische) Perzepte wesentlich durch Vorerfahrungen und früheren Interaktionen mit der Umgebung oder genetischem Wissen geprägt sind. Dabei wird angenommen, dass die Verteilungen einer Gauss-Funktion folgen. Die Bayes'sche Regel besagt, dass sich die posteriore Wahrscheinlichkeit proportional zu dem Produkt aus Plausibilitätsfunktion und a-priori-Wahrscheinlichkeitsverteilung verhält (s. Formel 1).

$$p(U|S) \propto p(S|U)p(U)$$

**Formel 1.** Die posteriore Verteilung  $p(U|S)$  der subjektiven Wahrscheinlichkeit eines Ursprung (U) für das Auftreten eines Stimulus (S) verhält sich proportional zum Produkt aus der Verteilung der Plausibilitätsannahmen [ $p(S|U)$ ] über das Auftreten eines Stimulus unter der Voraussetzung eines Ursprungs und der a-priori-Verteilung  $p(U)$  von Vorannahmen über mögliche Ursprünge.

Diese Formel beschreibt, dass Plausibilitätsannahmen und Vorwissen auf das Entstehen eines gewissen Perzepts einwirken. Anhand dieses Bezugsrahmens wurde Verhalten und Wahrnehmung bereits in Studien mit visuellen (Adams & Mamassian, 2004; Saunders & Knill, 2001; Weiss et al., 2002), visuo-haptischen (Adams et al., 2004; Ernst & Banks, 2002; Hillis et al., 2002), audiovisuellen (Alais & Burr, 2004) und visuomotorischen Aufgaben (Kording & Wolpert, 2004) beschrieben. Andere Untersuchungen ergaben jedoch, dass durch zeitliche Adaptation beeinflusste Gleichzeitigkeitsurteile von audiovisuellen Stimuli nur mit Einschränkung durch das Bayes'sche Theorem erklärt werden können (Miyazaki et al., 2006), belegten aber gleichzeitig die Gültigkeit dieses Modells, um zeitliche Adaptation auf taktile Stimuli psychophysisch zu erklären. Die Autoren schließen, dass die verschiedenen Modalitäten über unterschiedliche Mechanismen zur Anpassung an externe Reize verfügen. So sei es beispielsweise vorteilhaft, dass audiovisuelle Wahrnehmung unterschiedliche Laufzeiten von Signalen aufgrund von beispielsweise variablen Entfernungen zwischen Reiz und Beobachter ausgleichen kann, während dies für das taktile System, das Reize nur aus unmittelbarer Nähe empfangen kann, nicht notwendig sei.

In bezug auf ein Modell multisensorischer Integration könnte die Verteilung von  $p(A|U_A)$  die Annahme über das Auftreten eines sensorischen (auditorischen) Signals  $A$  unter Voraussetzung einer physikalischen Ursache  $U_A$  darstellen und die Verteilung von  $p(V|U_V)$  analog dazu die Annahme über das Auftreten eines visuellen Signals  $V$  unter der Voraussetzung des physikalischen Ursprungs  $U_V$ . Das Vorwissen  $p(U_A, U_V)$  würde das perzeptuelle Wissen des Beobachters über audiovisuelle Ereignisse darstellen. Infolge des Bayes'schen Theorems würden multimodale Perzepte durch die Auftretenswahrscheinlichkeiten beider einzelner Modalitäten und dem Vorwissen über verknüpfte audiovisuelle Ereignisse determiniert (Shams et al., 2005; s. Formel 2).

$$p(U_A, U_V | A, V) = \frac{p(A|U_A) p(V|U_V) p(U_A, U_V)}{p(A, V)}$$

**Formel 2.** Bayes'sches Inferenzmodell audiovisueller Verarbeitung. Die subjektive Ursachenattribution eines audiovisuellen Ereignisses  $p(U_A, U_V | A, V)$  hängt vom Wissen über das Auftreten auditorischer [ $p(A|U_A)$ ] und visueller [ $p(V|U_V)$ ] Ereignisse und dem Vorwissen  $p(U_A, U_V)$  über Ursachen auditorischer und visueller Stimuli ab.

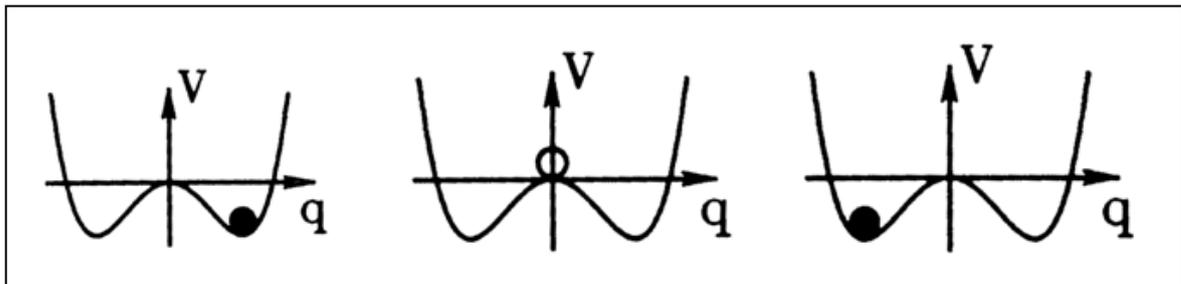
Diese Formulierung impliziert, dass sich die Modalitäten nicht direkt beeinflussen, sondern Einflussnahme einer Modalität auf eine andere durch die Ursachenattribution multimodaler Ereignisse erklärt wird. So würde ein auditorisches Ereignis die visuelle Wahrnehmung nur beeinflussen, wenn der Beobachter es für plausibel hält, dass beide Ereignisse denselben (z. B. räumlichen) Ursprung haben. Shams et al. (2005) konnten nachweisen, dass multisensorische Perzepte in bezug auf die oben vorgestellte Ton-induzierte Blitzillusion (Shams et al., 2000; 2002) durch das Bayes'sche Prinzip erklärt werden können.

Einige Autoren beschrieben, dass das Bayes'sche Theorem ein wirkungsvoller Ansatz sei, um Mechanismen der bimodalen Reizintegration auf Einzelzellebene (Anastasio et al., 2000; Patton & Anastasio, 2003) und auf der Ebene von Aktivität neuronaler Populationen (Deneve et al., 2001; Ma et al., 2006; Ma & Pouget, 2008) zu beschreiben. So gelingt es dem Wahrnehmungsapparat anhand der Antwortverteilung neuronaler Populationen optimale Schätzungen über die Stimulusbeschaffenheit zu berechnen bzw. anhand der sensorischen Information einer Zelle auf die Anwesenheit eines Reizes zu schließen.

### 1.2.3 Dynamische Systeme

Ein weiterer Ansatz, der menschliches Verhalten beschreibt, geht von einem System verschiedener Muster aus, die über die Zeit unterschiedliche Zustände annehmen können (z. B. Kelso, 1997). Der Organismus wird dabei als dynamisches System begriffen, das in einem Phasenraum existiert. Dabei repräsentieren Phasen, ähnlich wie unterschiedliche Aggregatzustände von Flüssigkeiten, gewisse Zustände, in dem sich das System befinden kann. Im Phasenraum befinden sich Stellen, die gegenüber anderen bevorzugt werden. Diese Stellen werden Attraktoren (s. Abb. 6) genannt. Das System strebt über den Verlauf der Zeit hin zu diesen Attraktoren. Diesen Prozess kann man beispielsweise anhand einer Feder erläutern, die, wenn sie auseinandergezogen wird, wieder in den zusammengezogenen Zustand zurückkehrt. Attraktoren können als Kurven dargestellt werden, die durch Täler und Anhöhen gekennzeichnet sind. Rollt eine Kugel, die einen Systemzustand darstellt in ein Tal, ist ein stabiler Zustand hergestellt. Attraktoren können sich durch Änderungen bestimmender Parameter verändern und zu

Instabilität führen (Abb. 6, mittlere Spalte) und damit einen Phasenübergang einleiten. Dieser Punkt stellt eine sogenannte Bifurkation dar, an der eine andere stabile Phase eingeleitet wird. Einen solchen Phasenübergang stellt bei Pferden beispielsweise der Übergang vom Gehen zum Trott dar. Dabei wählt ein Pferd keinen graduellen Übergang von einer Geschwindigkeit zu einer anderen, sondern vollzieht einen qualitativen Sprung zwischen den Geschwindigkeiten (Hoyt & Taylor, 1981).



**Abbildung 6.** Darstellung eines Attraktorraums und Bifurkation eines möglichen Bewegungsraums  $V$  für ein bestimmtes Verhalten.  $q$  stellt den Zustand eines Musters dar. Eine schwarze Kugel stellt stabile, eine weiße instabile Zustände dar. *Abbildung nach Kelso (1997).*

Das Konzept eines Phasenübergangs beobachtete Kelso (1984) anhand einer motorischen Aufgabe: Probanden wurden instruiert, die Zeigefinger ihrer beiden Hände mit einer bestimmten Frequenz entlang der horizontalen Ebene zu bewegen. Es wurde beobachtet, dass zwei bestimmte Muster stabil reproduziert werden konnten: entweder wurden homologe Muskelgruppen beider Hände zeitgleich kontrahiert (Bewegung ‚in Phase‘) oder abwechselnd (‚Anti-Phase‘). Als die Frequenz der Fingerbewegungen erhöht wurde, konnten Bewegungen, die anti-phasisch begannen, nicht aufrechterhalten werden, stabilisierten sich jedoch in solchen, die in Phase waren. Es wird vermutet, dass Organismen über unterschiedliche phasische Oszillationen verfügen, die miteinander abgestimmt werden. So konnte das Experiment von Kelso (1984) zeigen, dass die Bewegungen beider Finger zwei stabile Zustände annehmen konnten, die entweder in Phase oder anti-phasisch waren. Aufgrund eines Parameterwechsels (Erhöhung der Frequenz) wurde das Bewegungssystem instabil und ging (bei begonnenen anti-phasischen Bewegungen) in einen anderen Zustand über, der vom System offenbar präferiert wurde.

Die Bildung von Mustern (im obigen Beispiel eine bestimmte Bewegungsart) entsteht in Systemen spontan, und es wird daher davon ausgegangen, dass Systeme selbstorganisiert sind. Muster, die durch selbstorganisierte Prozesse entstehen, werden

durch interagierende Parameter beeinflusst (Kelso, 1997). Im Rahmen von dynamischen Systemen wird multimodale Integration als organisierendes Prinzip der Verknüpfung sensorischer Prozesse verstanden (Lagarde & Kelso, 2006). Dabei sind multisensorische Integration und Segregation koordinierte Zustände, zwischen denen das Gehirn wechseln kann. In der Literatur werden daher Parameter untersucht, die zum Entstehen oder Zusammenbrechen multimodaler Integration beitragen. So kombinierten Lagarde und Kelso (2006) Fingerbewegungen (jeweils Strecken und Beugen eines Fingers) mit auditorischen und taktilen Reizen. Dabei fanden sie, dass Teilnehmer die Kombination aus Strecken bei Tonsignal und Beugen bei Berührung gegenüber der umgekehrten Verknüpfung bevorzugten und dass die Integration der Bewegungen mit beiden Reizen bei hohen Bewegungsfrequenzen zusammenbrach. Die Dynamik multimodaler Koordination hängt demzufolge von Bewegungsart und Stimulusmodalität ab und wird durch die Bewegungsgeschwindigkeit determiniert.

Eine bildgebende Studie, die die neuronalen Korrelate von synchronen, asynchronen und von solchen audiovisuellen Stimuli, bei denen kein eindeutiges zeitliches bimodales Perzept ermöglicht wurde, untersuchte, zeigte, dass die erfolgreiche zeitliche Integration der Reize mit erhöhter neuronaler Aktivität in auditorischem, visuellem und Präfrontalkortex, aber auch in den Colliculi superiores (SC) einherging, während zeitliche Segregation den auditorischen, visuellen, Präfrontal- und Parietalkortex moduliert (Dhamala et al., 2007). Stimuli, bei denen kein eindeutiges Perzept hergestellt werden konnte, modulierten einen Teil der oben aufgeführten Netzwerke bestehend aus unisensorischen Kortexen und Präfrontalkortex (PFC). Die Autoren schließen aus diesen Befunden, dass die SC maßgeblich an der Herstellung eines synchronen Integrationszustands und der Parietalkortex an der Herstellung eines zeitlich segregierten Zustands beteiligt sind. Diese beiden Gebiete stellen laut Autoren wichtige Knotenpunkte dar, die durch ihre jeweilige Verknüpfung mit dem Basisnetzwerk stabile multimodale Zeitperzepte erzeugen.

#### **1.2.4 Neuronale Mechanismen multimodaler Integration**

Dieses Kapitel widmet sich der Frage, welche Strukturen an multisensorischer Integration beteiligt sind und welche Verarbeitungsmechanismen innerhalb dieser

---

Strukturen beobachtet werden können. Einzelzell-Untersuchungen an Katzen ergaben, dass die SC im Vergleich zu anderen Regionen des Gehirns viele Neurone enthalten, die auf mehr als eine Modalität antworten (z. B. Meredith & Stein, 1983; Stein & Meredith, 1993; Stein & Stanford, 2008). Dabei scheinen die Antwortmuster in diesen Zellen bestimmten Mechanismen bzw. Regeln zu folgen. Eine erste Regel folgt der Beobachtung, dass multisensorische Stimuli mit höherer Wahrscheinlichkeit integriert werden, wenn sie derselben räumlichen Quelle entspringen (Meredith & Stein, 1985). Dies wird anhand der Anordnung exzitatorischer und inhibitorischer Regionen erklärt, die die rezeptiven Felder multisensorischer Neurone definieren, in Folge der ein zentrales exzitatorisches Feld von einem inhibitorischen umgeben ist. Zudem überlappen auditorische und visuelle Felder multisensorischer Neurone (King & Palmer, 1985; Middlebrooks & Knudsen, 1984). Fallen beispielsweise räumlich kongruente audiovisuelle Stimuli in deren exzitatorische rezeptive Felder, verstärken sich die Effekte beider Modalitäten gegenseitig. Sind die beiden Ereignisse räumlich getrennt, würde ein Reiz in das inhibitorische Feld fallen und auf diese Weise die Antwort des anderen Ereignisses unterdrücken (Stein et al., 2004; Stein & Meredith, 1993).

Ein weiterer Mechanismus betrifft die zeitlichen Verhältnisse multisensorischer Ereignisse. Die Antwort eines bimodalen Neurons auf einen Reiz einer Sinnesmodalität kann verstärkt werden, wenn ein zweites Ereignis einer anderen Sinnesmodalität zeitgleich oder maximal mit 150 ms – 200 ms Verzögerung eintrifft (Meredith et al., 1987; Stein & Meredith, 1993). Entscheidend für die multisensorische Interaktion beider Modalitäten ist dabei die zeitliche Überlappung der maximalen Entladungsrate, die durch den jeweiligen Stimulus ausgelöst wird. Die zeitliche Verknüpfung bimodaler Reize auf neuronaler Ebene scheint demzufolge innerhalb eines gewissen Zeitfensters möglich zu sein. Dies deckt sich mit Beobachtungen psychophysischer Studien, die multimodale Integrationsmechanismen zeitlich verzögerter multimodaler Stimuli untersuchten: Das Zeitfenster, innerhalb dessen zwei Stimuli unterschiedlicher Modalität miteinander verbunden werden liegt zwischen 100 ms – 250 ms (z. B. Dixon & Spitz, 1980; Navarra et al., 2005; Spence et al., 2001; Stone et al., 2001; Vatakis et al., 2007; 2008; Zampini et al., 2005).

Wenn Reize multimodal integriert werden, ist auf neuronaler Ebene eine Veränderung des Antwortmusters im Vergleich mit Antworten auf einen unimodalen

---

Reiz zu erkennen. So können bimodale Stimuli die Antwortrate eines multisensorischen Neurons verstärken. Dabei werden signifikant mehr neuronale Impulse erzeugt als durch den effektivsten einer der beiden modalitätsspezifischen Stimuli (Meredith & Stein, 1983; Stein & Meredith, 1993; Stein & Stanford, 2008). Dieses Prinzip ist als Superadditivität bekannt und wurde im Rahmen von Tierstudien auch für Zellen im STS nachgewiesen (Barraclough et al., 2005). In analoger Weise können sich multimodale Antworten jedoch auch additiv und subadditiv verhalten. So kann ein auditorischer Reiz, der sich zeitlich oder räumlich inkongruent zu einem visuellen verhält, die Antwortrate eines multisensorischen Neurons auf den visuellen Stimulus unterdrücken, so dass ein segregiertes Perzept entsteht (Meredith & Stein 1983; Meredith et al., 1987).

Das Prinzip der inversen Effektivität vereint die drei oben aufgeführten Antwortmuster und besagt, dass die Verstärkung der neuronalen Antwortrate umso größer ist, je schwächer bzw. weniger effektiv die unimodalen Reize sind (Meredith & Stein, 1983). So würden ein schlecht sichtbarer visueller und ein schlecht hörbarer auditorischer Reiz stärkere neuronale bimodale Antworten hervorrufen als gut sichtbare bzw. hörbare unimodale Reize. Schwache unimodale Stimuli alleine würden nicht zu einer Antwort des Neurons führen, jedoch ihr gemeinsames Auftreten (z. B. Stein & Stanford, 2008). Diese Regel reflektiert den Tatbestand, dass durch einen zweiten Stimulus anderer Modalität keine bzw. wenig zusätzliche Information über einen vorliegenden, ohnehin sehr salienten Reiz gewonnen werden kann. Die hier aufgeführten Prinzipien wurden vor allem auf Einzelzellebene nachgewiesen. Einige Studien belegen deren Gültigkeit in bezug auf menschliches Verhalten. So zeigt eine Untersuchung, dass durch das kurzzeitig vorherige Darbieten eines Tonreizes visuelle Reize mit höherer Wahrscheinlichkeit entdeckt werden (McDonald et al., 2000). Andere Studien belegen, dass schwache visuelle Reize besser entdeckt werden, wenn sie mit auditorischen räumlich und zeitlich übereinstimmen, als wenn sie unimodal präsentiert werden (Bolognini et al., 2005; Frassinetti et al., 2002). Umgekehrt wurden Töne im Vergleich zur unimodalen Präsentation als lauter empfunden, wenn sie synchron mit einem Lichtreiz dargeboten wurden (Odgaard et al., 2004). Darüber hinaus konnte gezeigt werden, dass kurze visuelle Reize besser entdeckt werden, wenn sie mit auditorischen kombiniert werden, als wenn sie unimodal dargeboten werden (Noesselt et al., 2008). Diese Studien liefern Hinweise darauf, dass sich die Regeln der multimodalen Verstärkung, die auf

neuronaler Ebene an Tieren belegt wurden, bis auf die Ebene menschlichen Verhaltens durchzusetzen scheinen. Ross et al. (2007) zeigten, dass schwer verständliche audiovisuelle Sprachreize im Vergleich zu rein auditorischen leichter identifiziert wurden und belegten die Gültigkeit der inversen Effektivität auf Verhaltensebene. Sie fanden jedoch, dass die Effekte multimodaler Verstärkung am größten zwischen den beiden Extremen der Verständlichkeitsdimension war, und nicht, wie es die Regel der inversen Effektivität implizieren würde, an deren unteren Ende, bei sehr schwerer Verständlichkeit. Die Autoren schließen, dass das Gesetz der inversen Effektivität in bezug auf beobachtbares Verhalten bei Menschen nur innerhalb eines gewissen Bereichs Gültigkeit besitzt.

Obwohl die Praktikabilität im Bereich der bildgebenden Studien diskutiert wird (Laurienti et al., 2005), konnten bildgebende Studien Superadditivität und Subadditivität in bezug auf audiovisuelle Sprachstimuli im Vergleich zu unimodaler Sprache anhand von funktioneller Magnetresonanztomographie (fMRT) nachweisen (Calvert et al., 2001; Wright et al., 2003). Da dieses Verfahren (s. 2.1) nicht die Aktivität einzelner Neurone erfasst, sondern Signale anhand von Veränderungen in neuronalen Populationen gemessen werden, implizieren diese Befunde, dass die oben aufgeführten Prinzipien multimodaler Integration auch innerhalb neuronaler Populationen Gültigkeit besitzen. Regionen, die dieses neuronale Antwortmuster zeigten, waren die SC, der Sulcus temporalis superior (STS), die Insula, der Parietalkortex und der Präfrontalkortex (PFC). Andere fMRT-Studien, die audiovisuelle Sprachreize untersuchten, fanden jedoch keine superadditiven Effekte in diesen Regionen (Beauchamp et al., 2004a; van Atteveldt et al., 2004). Laurienti et al. (2002) zeigten, dass Regionen, die spezifisch auditorische Reize verarbeiten (auditorischer Kortex), durch Stimulation mit visuellen Reizen eine Verringerung des Antwortsignals aufweisen. Derselbe Mechanismus wurde in umgekehrter Weise für den visuellen Kortex gefunden. Bei bimodaler Stimulation wurde keine Deaktivierung der Antwortsignale entdeckt. Die Autoren schließen, dass dieser Befund den Mechanismus der multimodalen Verstärkung (Superadditivität) auf der Ebene neuronaler Populationen untermauert. Von Wassenhove et al. (2005) zeigten anhand einer Elektroenzephalographie (EEG)-Studie, dass audiovisuelle Sprachstimuli im Vergleich zu rein auditorischen Reizen mit einer Verringerung evozierter Potentiale

einhergehen. Giard und Perronnet (1999) konnten dagegen eine Verstärkung evozierter Potentiale durch audiovisuelle Stimulation im Vergleich zu unimodaler beobachten.

Weitere fMRT-Studien zeigten, dass Superadditivität bzw. inverse Effektivität in bezug auf die Änderungen des blood oxygen level dependent (BOLD)-Signals gültige Prinzipien darstellen, um Areale zu identifizieren, die maßgeblich an der multimodalen Integration von Ereignissen beteiligt sind (Stevenson et al., 2007; Stevenson et al., 2009; Stevenson & James, 2009). Dabei wurden nicht-sprachliche, komplexe audiovisuelle Stimuli verwendet, deren Wahrnehmungsschwellen variiert wurden (Stevenson et al., 2007), und anhand derer die Gültigkeit multimodaler Verstärkung im Sinne des Superadditivitätsprinzips im STS und in auditorischen bzw. visuellen Arealen untermauert werden konnte. Darüber hinaus wurde gezeigt, dass die BOLD-Muster im STS dem Prinzip der inversen Effektivität folgen, wenn sowohl sprachliche als auch semantisch nicht belegte audiovisuelle Stimuli verarbeitet wurden (Stevenson & James, 2009). In dieser Studie wurde das Signal-zu-Rausch-Verhältnis (SNR) audiovisueller Reize variiert: je schlechter die Reize dadurch wahrgenommen wurden, desto größer war der Zuwachs multimodaler BOLD-Antworten im STS im Vergleich zur unimodalen Stimulation. In einer weiteren fMRT-Untersuchung konnte das Antwortmuster der inversen Effektivität für die Verarbeitung audiovisueller Reize auch innerhalb präfrontaler Regionen, des Gyrus parahippocampalis, des Parietallappens und des posterioren Cingulums identifiziert werden (Stevenson et al., 2009). Zusätzlich wurden der Gyrus fusiformis (FG) und das Cerebellum anhand dieses Prinzips lokalisiert, wenn visuo-haptische Stimuli integriert wurden.

Der aktuelle Wissensstand zeigt, dass multimodale Reize auf Zellebene nach gewissen Regeln verarbeitet werden. Diese konnten anhand von Verhaltensstudien auch bei Menschen nachgewiesen werden, und es wurden Vorgehensweisen entwickelt, die es mit Hilfe dieser Regeln ermöglichen, multimodale Regionen in bildgebenden Verfahren zu identifizieren. Einige Autoren (Anastasio et al., 2000; Patton & Anastasio, 2003) zeigen, dass multimodale Verstärkung bzw. inverse Effektivität auf Zellebene anhand Bayes'scher Regeln nachzuvollziehen sind. Andere Autoren diskutieren jedoch, dass diese Prinzipien statistischen Artefakten unterliegen (Holmes, 2007; 2009).

Unabhängig von zugrundeliegenden Mechanismen multimodaler Integration konnte anhand dieses Abschnitts gezeigt werden, dass verschiedene Sinnesmodalitäten

schon auf der Ebene hierarchisch früher Prozesse stark miteinander interagieren und sich maßgeblich unter anderem anhand spezialisierter Neurone gegenseitig beeinflussen. Ein Konzept, in dem von eigenständigen Wahrnehmungsmodalitäten ausgegangen wird (s. 1.2), die über späte kognitive Prozesse miteinander in Verbindung gebracht werden, scheint daher nicht haltbar. Besondere Einblicke in das automatisierte, intensive Zusammenspiel verschiedener Sinnesmodalitäten gibt das Phänomen der Synästhesie.

### 1.2.5 Synästhesie

Synästheten sind Menschen, bei denen Wahrnehmungen einer Sinnesmodalität von illusorischen Perzepten in einer anderen begleitet werden. So berichteten Beeli et al. (2005) beispielsweise von einer Frau, die musikalische Intervalle unwillkürlich mit bestimmten Geschmäckern in Verbindung brachte. Die Präsentation einer großen Sekunde löste bei ihr beispielsweise kontinuierlich ein saures, eine kleine Septime ein bitteres Geschmackserzept aus. Es existieren unterschiedliche Formen von Synästhesie, wobei jeweils andere Eindrücke verschiedener Modalitäten automatisch miteinander verknüpft werden. Die am meisten berichtete Form ist jedoch die Graphem-Farb-Synästhesie, bei der visuell dargebotene Zeichen zu Farbassoziationen führen (Rich et al., 2005).

Eine Positronen-Emissionstomographie (PET)-Untersuchung verglich kortikale Aktivierungen von Graphem-Farb-Synästheten mit denen normaler Probanden und zeigte, dass Wörter bei Synästheten neben sprachspezifischen Gebieten visuelle Areale modulierten, nicht aber bei normalen Probanden (Paulesu et al., 1995). Dies ist ein Hinweis darauf, dass bestimmte Stimuli bei Synästheten modalitätsfremde Aktivierungen hervorrufen. Einige Autoren (Baron-Cohen et al., 1993; Hubbard & Ramachandran, 2005) schlagen vor, dass Synästheten Stimuli stärker multimodal verarbeiten als normale Probanden. Eine Studie, die das Verhalten von Synästheten anhand einer an die Form der Synästhesie angepassten Interferenz Aufgabe untersuchte, zeigte, dass Synästheten keine spezifischen crossmodalen Farbassoziationen zu Buchstaben berichteten, wenn die Graphemreize maskiert wurden, d.h. so kurz dargeboten wurden, dass sie zwar verarbeitet wurden, aber bewussten, höheren kognitiven Prozessen nicht zugänglich

waren (Mattingley et al., 2001). Die Autoren schließen, dass Synästhesie von Prozessen abhängt, die rein perzeptuellen Prozessen in der Verarbeitungshierarchie nachfolgen und stützen damit die Hypothese, dass im Vergleich zu normalen Probanden beispielweise Unterschiede in höheren multimodalen Arealen maßgeblich zu synästhetischen Perzepten beitragen. Es wird vermutet, dass der Parietalkortex am verstärkten Verbinden zweier Modalitäten bei Synästheten beteiligt ist (Beeli et al., 2008; Esterman et al., 2006; Weiss et al., 2005). Hinweise auf die Bedeutung des Parietalkortex bei synästhetischen Perzepten gibt eine Untersuchung, bei der anhand transkranieller Magnetstimulation (TMS) dieses Gebiet bei Graphem-Farb-Synästheten vorübergehend gestört wurde. Diese Untersuchung zeigt, dass daraufhin die Interferenzanfälligkeit der Probanden sank, was auf verminderte automatische Assoziationen zwischen Buchstaben und Farben hinweist (Esterman et al., 2006). Weiss et al. (2005) zeigten anhand einer fMRT-Untersuchung, dass durch Grapheme erzeugte Farbwahrnehmungen im Vergleich zu Stimuli, die keine Farbassoziationen hervorriefen mit erhöhter Aktivität im linken Sulcus intraparietalis (IPS) einhergingen. Da dieses Gebiet in die Verarbeitung multimodaler Ereignisse involviert ist (z. B. Calvert et al., 2001; Driver & Noesselt, 2008; Macaluso et al., 2004; Miller & D'Esposito, 2005), wird vermutet, dass im Vergleich zu normalen Probanden verstärkte crossmodale Verbindungen zu synästhetischen Perzepten führen. Beeli et al. (2008) zeigten anhand einer EEG-Studie, dass der Parietalkortex, der FG, der inferiore Temporalkortex und orbitofrontale Gebiete bei Synästheten während der Darbietung von Buchstaben und Wörtern stärker aktiviert waren als bei normalen Probanden. Diese Gebiete scheinen demzufolge kritisch für die Bildung synästhetischer Perzepte zu sein. Zusätzlich fanden die Autoren gleichzeitige Aktivierungen in auditorischen und visuellen Gebieten der Synästheten, die bereits früh (122 ms), nachdem die Stimuli jeweils dargeboten wurden erschienen und schließen auf automatische, schnelle Verarbeitungsprozesse, die dem Phänomen der Synästhesie unterliegen, die allerdings auch durch nachfolgende Rückkopplungsmechanismen beeinflusst werden können (Grossenbacher & Lovelace, 2001; Smilek et al., 2001). Auf Grundlage von mehreren Studien kommen Hubbard und Ramachandran (2005) zu der Annahme, dass Synästheten anhand derselben neuronalen Mechanismen wie normale Probanden Stimuli verarbeiten. Die stimuluskorrelierte neuronale Aktivität ist jedoch im Vergleich zu normalen Probanden größer. Diese Erkenntnisse ermöglichen es, anhand von Befunden zur

Synästhesie Vermutungen über crossmodale Verarbeitungsmechanismen aufzustellen (Sagiv & Ward, 2006).

Aufschluss darüber, warum Synästheten bestimmte Stimuli unter höherem neuronalen Aufwand verarbeiten, konnten Studien geben, die die neuroanatomische Beschaffenheit des Kortex von Graphem-Farb-Synästheten untersuchen. Sie zeigten, dass visuelle Gebiete [FG, Gyrus lingualis (LG), calcariner Kortex] bei Synästheten vergrößert sind (Jancke et al., 2009). Unterschiede der kortikalen Struktur wurden vor allem im FG gefunden, der an der Verarbeitung von Farbe, Buchstaben und Worten beteiligt ist (Bartels & Zeki, 2000; Joseph et al., 2006; Pernet et al., 2005; Vinckier et al., 2007). Erhöhte Konnektivität dieses Gebietes mit anderen kortikalen Strukturen konnte jedoch hier nicht nachgewiesen werden. Die Autoren kommen daher zu dem Schluss, dass Verarbeitungsmechanismen bei Synästheten im Vergleich zu normalen Probanden unterschiedlich organisiert sein könnten. Hanggi et al. (2008) konnten jedoch zeigen, dass der Kortex der am Kapitelanfang erwähnten Synästhetin erhöhte Konnektivität in auditorischen und insulären Gebieten [letzteres ist in die Wahrnehmung von Geschmäckern involviert (Augustine, 1996)] aufweist. Die Autoren schließen, dass vermehrte anatomische Verbindungen eine Grundlage von synästhetischer Wahrnehmung sind. Dies schließen auch Rouw und Scholte (2007), die bei Graphem-Farb-Synästhesisten sogenannte Hyperkonnektivität im inferioren Temporalkortex, im Frontal- und Parietalkortex nachwiesen.

Synästhesie stellt einen Spezialfall multimodaler Integration dar. Die Erkenntnisse aus Untersuchungen dieses Phänomens weisen darauf hin, dass unisensorische Kortices bzw. verschiedene Sinnesmodalitäten eng miteinander verknüpft sind und eine schnelle Integration multimodaler Ereignisse ermöglichen. Grossenbacher und Lovelace (2001) gehen davon aus, dass Synästhesie durch Disinhibition von Rückkopplungsmechanismen zwischen den Modalitäten entsteht. Diese Theorie impliziert ausgeprägte Interaktionen zwischen den modalitätsspezifischen Gebieten. Psychophysische und neuronale Grundlagen multimodaler Integration sollen im Folgenden detailliert besprochen werden.

## 1.2.6 Studien zur multimodalen Integration

### 1.2.6.1 Psychophysik

Für die gelungene Integration multimodaler Ereignisse, wie z. B. bei einem Gespräch, während dessen visuelle und auditorische Informationen verarbeitet werden müssen, spielen zeitliche, räumliche und semantische Faktoren eine wichtige Rolle. Nur wenn in bezug auf diese Umgebungsparameter weitestgehend Kongruenz vorliegt, kann das menschliche Wahrnehmungssystem einheitliche Perzepte erstellen. Um das Zusammenspiel mehrerer Sinneskanäle zu untersuchen, wurden im Rahmen zahlreicher Studien Informationen experimentell so variiert, dass verschiedene Modalitäten widersprüchliche Informationen erhielten. Dabei entstanden sogenannte multisensorische Illusionen: das Konzept des auditorischen Antriebs zeigt auf, dass visuelle Reize dem abweichenden Zeitmuster simultan dargebotener auditorischer Reize zugeordnet werden. So wird die Frequenz kurz hintereinander aufleuchtender Lichtblitze als schneller oder langsamer wahrgenommen, wenn lediglich die Frequenz simultan präsentierter kurzer Tonsignale verändert wird (Gebhard & Mowbray, 1959; Shipley, 1964). Der umgekehrte Einfluss visueller Zeitstrukturen auf auditorische hatte einen weitaus geringeren Effekt (Welch et al., 1986). Dieser Befund wird im Sinne der Modalitätseignungshypothese erklärt (Welch & Warren, 1980; s. 1.2.1), die besagt, dass das auditorische System bei der Beurteilung zeitlicher Muster dem visuellen überlegen ist. Fendrich & Corballis (2001) schlagen vor, dass visuelle und auditorische Reize zeitlich in subjektiven Phasen organisiert sind (s. auch 1.2.3), die die Zeitmuster beider Modalitäten gruppieren. Einen weiteren Beleg für den Einfluss der auditorischen Modalität auf die visuelle in bezug auf die Beurteilung von zeitlichen Mustern liefert die Ton-induzierte Lichtillusion (Shams et al., 2000; 2002; s.1.2.1). Dabei wird ein Lichtblitz als zwei Lichtblitze wahrgenommen, wenn er mit zwei auditorischen Tonsignalen zusammen dargeboten wird. Umgekehrt konnte dieser Einfluss nicht beobachtet werden.

Beim Bauchrednereffekt wurden Licht- und Tonsignale räumlich voneinander getrennt, aber gleichzeitig präsentiert. Dabei wurde die subjektive Tonquelle systematisch in die Richtung des Lichtreizes verzerrt (Bertelson & Radeau, 1981; Vroomen & De Gelder, 2004). Der umgekehrte Effekt konnte nicht beobachtet werden (Pick et al., 1969). Dieses Prinzip bewirkt, dass die Sprachreize eines Bauchredners im Allgemeinen den

Mundbewegungen der Puppe, die er simuliert zugeordnet werden. Darüber hinaus konnte auch gezeigt werden, dass sich eine wahrgenommene Tonquelle in Richtung eines taktilen Reizes verschiebt (Caclin et al., 2002). Im Sinne der Modalitätseignungshypothese (s. 1.2.1) tritt dieser Effekt auf, weil das visuelle System in bezug auf die Analyse räumlicher Muster dem auditorischen überlegen ist bzw. das taktile dem auditorischen.

Beim McGurk-Effekt werden der auditorischen und visuellen Modalität semantisch widersprüchliche Informationen dargeboten, die zu multisensorischen Perzepten führen, die weder der auditorischen, noch der visuellen Modalität entsprechen (McGurk & MacDonald, 1976; s. 1.2.1.). Das unwillkürliche Entstehen von Perzepten während der McGurk-Illusion, die durch alle beteiligten Modalitäten beeinflusst werden, weist darauf hin, dass multimodale Prozesse automatischen, eng miteinander verknüpften Mechanismen unterliegen.

Im Alltag werden jedoch meistens zusammengehörige Informationen über Objekte durch verschiedene Sinnesmodalitäten zusammengefasst. Dabei entsteht die Frage, welche Informationen einer Modalität in welcher Art und Weise mit Informationsanteilen eines anderen Sinneskanals verknüpft werden, um einheitliche Perzepte zu entwickeln. Die oben aufgeführten Illusionsstudien weisen darauf hin, dass zeitliche, räumliche und semantische Aspekte kritisch für die multimodalen Integrationsmechanismen sein können. Darüber hinaus scheint es inhärente Mechanismen zu geben, anhand derer Beiträge einzelner Modalitäten statistisch optimal miteinander verknüpft werden (s. 1.2.2).

Wenn audiovisuelle Signale verarbeitet werden, ist der Informationszuwachs über ein multimodales Ereignis, den eine Modalität durch eine andere bekommen kann, am größten, wenn beide Ereignisse räumlich und zeitlich beieinander liegen (Bolognini et al., 2005; Frassinetti et al., 2002). Räumliche Abweichungen von 16° und zeitliche von 100 ms vermindern die Interaktion beider Sinneskanäle bereits. Andererseits führt die klare räumliche Trennung bimodaler Ereignisse zur Verbesserung der Einschätzung des zeitlichen Abstands zwischen Ereignissen unterschiedlicher Modalität (Zampini et al., 2003). Laurienti et al. (2005) zeigten, dass semantische Kongruenz zwischen zwei Modalitäten die Integration eines Ereignisses fördert.

In mehreren Untersuchungen wurde der zeitliche Abstand zwischen Ereignissen zweier Modalitäten systematisch variiert, um Anhaltspunkte darüber zu gewinnen, wie

weit beide zeitlich disparat erscheinen können, um immer noch als gleichzeitig wahrgenommen zu werden. Diese Zeitspanne bewegt sich zwischen ca. 114 ms und 300 ms. (Dixon & Spitz, 1980; Fujisaki et al., 2004; McGrath & Summerfield, 1985; Spence et al., 2001; Stone et al., 2001; Vatakis & Spence, 2006; von Wassenhove et al., 2007; Zampini et al., 2005). Dabei lassen Sprachreize größere Zeitabstände zwischen beiden Modalitäten zu als semantisch nicht belegte (z. B. Dixon & Spitz, 1980). Des Weiteren zeigten die Studien, dass die Toleranz von audiovisueller Asynchronizität im menschlichen Wahrnehmungssystem größer ist, wenn das Bild dem Ton vorausgeht als umgekehrt. Diese Asymmetrie wurde als ökologisch plausibel bewertet, da es in der Natur aufgrund der langsameren Übertragung von Schall im Vergleich zu Licht oft vorkomme, dass die akustische Information später am Ohr eintrifft als die visuelle am Auge. Andere Studien (Grant et al., 2003; Munhall et al., 1996), in denen Wörter oder Silben untersucht wurden, konnten diese Asymmetrie bestätigen und fanden hinsichtlich der Verzögerungstoleranz vergleichbare Ergebnisse: Die Tonspur kann zwischen ca. 130 und 50 ms vor der Bildspur erscheinen, ohne beim Menschen ein inkongruentes Perzept zu erzeugen, jedoch kann sie von 180 bis zu ca. 250 ms hinter der Bildspur verlaufen, um mit der Bildspur gebunden werden zu können. Kurze, einfache audiovisuelle Reize werden innerhalb eines Zeitfensters von ca. 100-180 ms integriert (Fujisaki et al., 2004; Stone et al., 2001; Vatakis et al., 2007; Zampini et al., 2005).

In der multimodalen Forschung geht man aufgrund der oben dargestellten Befunde davon aus, dass es ein bewegliches Zeitfenster gibt, in dem die Verbindung zweier Ereignisse möglich ist (Navarra et al., 2005; Spence & Squire, 2003; Sugita & Suzuki, 2003; Vatakis et al., 2008). Studien zum McGurk-Effekt, in denen auditorische und visuelle Stimuli zeitlich gegeneinander verschoben wurden, belegen die oben aufgeführten Befunde: semantisch inkongruente auditorische und visuelle Reize werden selbst bei einer zeitlich verzögerten Präsentation von ca. 200 ms miteinander verbunden (Munhall et al., 1996; von Wassenhove et al., 2005; 2007), so dass der McGurk-Effekt aufrecht erhalten wird.

Neben räumlichen, zeitlichen und semantischen Faktoren scheinen kognitive Prozesse, wie beispielsweise die explizite Annahme darüber, dass zwei Ereignisse unterschiedlicher Modalität zusammengehören ‚müssen‘ (Welch & Warren, 1980) multimodale Integrationsmechanismen zu beeinflussen. So konnten Vatakis und Spence

(2007) zeigen, dass audiovisuelle Sprachreize präziser bewertet werden konnten, sobald diese Annahme vorlag. Sugita & Suzuki (2003) demonstrierten außerdem, dass implizite Annahmen über die Distanz audiovisueller Ereignisse die Integrationsmechanismen beeinflussen.

Es konnte gezeigt werden, dass das menschliche Wahrnehmungssystem an räumliche und zeitliche Diskrepanzen bimodaler Reize adaptiert. So schätzten Probanden, die ca. 20 Minuten Stimuli ausgesetzt wurden, die räumlich um 8° gegeneinander verschoben waren Ton- und Lichtreize um etwa dieselbe Distanz gegeneinander verschoben ein (Recanzone, 1998). Analog dazu konnte die audiovisuelle Gleichzeitigkeitswahrnehmung um etwa 10% der zuvor ausgesetzten Zeitdifferenzen zwischen Licht und Ton vorübergehend verschoben werden (Fujisaki et al., 2004; Navarra et al., 2005; Vatakis et al., 2008; 2007; Vroomen et al., 2004). Diese Befunde konnten sowohl anhand von Sprachreizen (Navarra et al., 2005; Vatakis et al., 2007; 2008), als auch anhand einfacher audiovisueller (Fujisaki et al., 2004; Vroomen et al., 2004), aber auch anhand audiotaktile (Navarra et al., 2007) Reize nachgewiesen werden. Darüber hinaus zeigten Navarra et al. (2005) und Vatakis et al., (2008), dass sich die Adaptation an Zeitabstände bei audiovisuellen Sprachreizen auf die nachträgliche Beurteilung einfacher audiovisueller Stimuli systematisch auswirkte. Dies weist darauf hin, dass audiovisuelle, zeitliche Adaptationsmechanismen nicht stimulusspezifisch sind, sondern in gewisser Weise über Stimuluseigenschaften generalisieren.

Des Weiteren zeigten Studien, dass unterschiedliche Sinneskanäle nicht nur Informationen über die gleichen Eigenschaften eines Objekts oder Ereignisses liefern, sondern dass eine Modalität die Beurteilung von Eigenschaften eines Objekts oder Ereignisses einer anderen Modalität beeinflussen kann. So zeigten McDonald et al. (2000), dass die Genauigkeit des visuellen Systems, das Auftauchen eines Reizes zu entdecken, unabhängig von Aufmerksamkeitseffekten zunimmt, wenn dieses visuelle Ereignis mit einem Ton gepaart wird. Des Weiteren konnte man zeigen, dass die gleichzeitige Präsentation von auditorischen und visuellen Stimuli die zeitliche Auflösung des visuellen Systems unabhängig von Aufmerksamkeitsprozessen verbessern kann (Noesselt et al., 2008).

Die oben aufgeführten Befunde legen dar, dass das menschliche Wahrnehmungssystem multimodale Reize effizient verarbeiten kann und durch die

adäquate Auswertung von Informationen mehrerer Sinneskanäle sogar Sensibilitätssteigerungen bewirken kann. Dabei besitzt es hohe Flexibilität und passt sich beispielsweise räumlichen und zeitlichen Veränderungen in bezug auf die Konstellation multisensorischer Reize an. Um ein einheitliches, multisensorisches Perzept herzustellen, werden kognitive Einflüsse, aber auch stimulusgetriebene Informationen (s. Adaptation) einbezogen. Im Folgenden sollen neuronale Grundlagen multisensorischer Mechanismen anhand bisher ermittelter Befunde vorgestellt werden.

### **1.2.6.2 Neuronale Grundlagen**

#### **1.2.6.3 Tierstudien**

Ergebnisse aus bisher durchgeführten Einzelzellstudien und fMRT-Untersuchungen am Menschen ergaben, dass das Gehirn über zahlreiche multisensorische Gebiete verfügt, in denen Neurone über Afferenzen Informationen aus mehreren Sinneskanälen empfangen können. Diese Gebiete wurden sowohl innerhalb subkortikaler als auch kortikaler Regionen identifiziert (z. B. Mesulam, 1998; Wallace et al., 2004). So wurde gezeigt, dass unter anderem die Basalganglien (z. B. Nagy et al., 2006) und die SC Informationen aus somatosensorischen, auditorischen und visuellen Gebieten empfangen (z. B. Meredith & Stein, 1983; Stein et al., 2004, vgl. 1.2.4). Die Aktivität der Zellen in den SC hängt dabei von der räumlichen und zeitlichen Konstellation der Informationen der verschiedenen Modalitäten ab. Anhand von Tierstudien konnte gezeigt werden, dass multisensorische Mechanismen in den SC von höheren multisensorischen Arealen moduliert werden. Durch reversible Läsionen des rostral lateralen Temporalkortex von Katzen wurden multisensorische Effekte der SC unterdrückt (Jiang et al., 2001). Darüber hinaus verschwanden die multisensorischen Eigenschaften der SC in Katzen, wenn der Teile des Temporalkortex in Jungtieren entfernt wurden (Jiang et al., 2007; Wallace et al., 2006).

Tierstudien an Primaten, die multimodale Eigenschaften kortikaler Areale untersuchten, zeigten, dass der rostrale Teil des STS über Verbindungen zu unisensorischen auditorischen, visuellen und somatosensorischen Kortices verfügt (Cusick, 1997; Padberg et al., 2003) und multisensorische Zellen enthält (Barraclough et al., 2005; Bruce et al., 1981).

Darüber hinaus wurde berichtet, dass Regionen innerhalb des Parietalkortex Informationen aus unisensorischen Kortices verarbeiten können (Duhamel et al., 1998; Ghazanfar & Schroeder, 2006; Schroeder & Foxe, 2004) und dabei räumliche multisensorische Prozesse abbilden können (Maravita et al., 2003). Des Weiteren wurde bei Affen gezeigt, dass prämotorische und präfrontale Regionen an der Verarbeitung multisensorischer Prozesse beteiligt sind (Barbas et al., 2005; Fuster, 1985; 1987; Fuster et al., 2000; Sugihara et al., 2006). Budinger et al. (2007) zeigten, dass Wüstenrennmäuse über direkte Verbindungen zwischen PFC und primären sensorischen Kortices verfügen.

Die Befunde aus Tierstudien weisen drauf hin, dass ein Netzwerk aus SC, STS, Parietalkortex, PFC und unisensorischen Kortices an multimodalen Prozessen beteiligt sind. Befunde bildgebender Verfahren bei Menschen sollen im folgenden Abschnitt besprochen werden.

#### **1.2.6.4 Humanstudien**

Neuronale Korrelate der psychophysischen Befunde zur multimodalen Integration wurden anhand von fMRT-, PET-, Magnetenzephalographie (MEG)-, und EEG-Studien untersucht. So wurden beispielsweise neuronale Korrelate der McGurk-Illusion (s.1.2.4) im Rahmen einer PET-Studie in Insula, STS und Claustrum identifiziert (Olson et al., 2002). Diese Gebiete wurden moduliert, wenn durch synchrone im Vergleich zu asynchroner Präsentation audiovisueller Silben der McGurk-Effekt hervorgerufen wurde. Mottonen et al. (2002) zeigten in einer MEG-Studie, dass die Darbietung inkongruenter visueller Silben in bezug zu auditorischen den posterioren Temporalkortex moduliert und dass ein Wechsel kongruenter audiovisueller Information hin zu inkongruenter mit schnelleren Veränderungen im auditorischen Kortex einherging, als dies bei unimodalen Reizen der Fall war. Die Autoren schließen, dass schnelle Interaktionsmechanismen zwischen unisensorischen Kortices multimodale Integrationsprozesse beeinflussen. Im Rahmen einer EEG-Studie, in der McGurk-Stimuli verwendet wurden, konnte gezeigt werden, dass visuelle Sprachreize die Verarbeitung audiovisueller Sprachreize im auditorischen Kortex beschleunigen (von Wassenhove et al., 2005). Eine weitere EEG-Studie demonstrierte, dass McGurk-Illusionen nacheinander Parietalkortex, PFC und schließlich visuelle Areale modulieren (Kaiser et al., 2005). Skipper et al. (2007) zeigten in einem fMRT-Experiment, dass während der McGurk-

Illusion Insula, visueller Kortex, PFC und motorische Areale moduliert werden. Sie beschrieben darüber hinaus, dass unterschiedliche Perzepte der Silbenreize mit systematischen Veränderungen der Aktivierungsmuster in motorischen Arealen einhergehen und schlossen, dass die McGurk Illusion durch die Einbindung bereits vorhandener motorischer Repräsentationen von Artikulationen in die Integration auditorischer und visueller Informationen entsteht. Jones und Callan (2003) fanden Effekte in Parietal-, Präfrontal- und superioren Temporalkortex wenn McGurk-Illusionen hervorgerufen wurden.

Neurophysiologische Untersuchungen des Bauchrednereffekts (s. 1.2.4) zeigten, dass STS und visueller Kortex durch die räumliche Integration audiovisueller Reize unterschiedliche Muster der zugrundeliegenden ereigniskorrelierten Potentiale (EKP) aufwiesen. Eine EEG-Studie demonstrierte, dass die tatsächliche Verschiebung einer Tonquelle dieselben Muster der EKP hervorruft wie die wahrgenommene Illusion, die ein Tonsignal hin zu einem Lichtreiz verschiebt (Stekelenburg et al., 2004). Busse et al. (2005) zeigten anhand einer kombinierten EEG-/fMRT-Studie, dass die Raumillusion mit höherer Aktivität im auditorischen Kortex einhergeht, und dass die gleichzeitige Darbietung eines räumlich disparaten Lichtreizes mit einem Tonreiz zu Veränderungen der EKP in frontalen Gebieten führt. Die Autoren schließen, dass eine Modalität Aufmerksamkeitsprozesse derart steuern kann, dass gleichzeitige Reize (trotz räumlicher Trennung) integriert werden können. Eine fMRT-Studie beschrieb, dass die räumliche Illusion mit erhöhter Aktivität in Insula und STS einhergeht (Bischoff et al., 2007). Schließlich zeigten Bonath et al. (2007) anhand einer kombinierten EEG/fMRT-Studie, dass der Ort einer entstandenen Raumillusion (links vs. rechts) kontralateral das Planum Temporale moduliert.

Anhand einer EEG-Studie zeigten Mishra et al. (2007), dass die Ton-induzierte Lichtblitz-Illusion mit frühen Modulationen visueller Areale und des STS einhergeht. Jedoch unterschieden sich Modulationen von Aktivitäten, die mit einem tatsächlichen (zweiten) Lichtblitz einhergingen. Die Autoren schließen, dass ein zweiter Ton durch schnelle Interaktionen unisensorischer Areale und dem STS eine visuelle Illusion auslöst (s. auch Mishra et al., 2008).

Leistungsgewinne in bezug auf zeitliche Auflösung, die das visuelle System durch gepaarte auditorische Reize erfahren kann (s. 1.2.4; McDonald et al., 2000), konnten im

Rahmen einer EEG-Studie mit frühen Modulationen des STS und kurz danach auftretenden Modulationen visueller Gebiete assoziiert werden. Die Autoren vermuten, dass Rückkopplungsmechanismen aus dem STS zu perzeptuellen Verbesserungen des visuellen Kortex führen können (s. auch Noesselt et al., 2007).

In Abschnitt 1.2.6.1 wurden neben multimodalen Illusionen räumliche, zeitliche und semantische Faktoren integrativer Prozesse beschrieben, deren neuronale Grundlagen im Folgenden besprochen werden. Im Rahmen einer PET-Studie wurden auditorische und visuelle Sprachreize räumlich voneinander getrennt (Macaluso et al. 2004). Die Autoren berichteten über Modulationen unisensorischer visueller Areale und des STS. Dabei variierte Aktivität im inferioren Parietallappen systematisch mit der räumlichen Trennung bei Sinnesinformationen. In einer weiteren PET-Studie konnte gezeigt werden, dass audio-taktile Reize prämotorische Areale und STS modulieren, aber auch, dass der IPS systematisch durch die räumliche Variation der Stimuli beeinflusst wird (Macaluso & Driver, 2001). Darüber hinaus wurden unisensorische visuelle Areale durch räumliche Veränderungen audio-taktile Reize moduliert (Macaluso & Driver, 2005).

Calvert et al. (2000) fanden in einer fMRT-Studie zu zeitlich asynchronen Sprachstimuli Modulationen in primär auditorischen und visuellen und in frontalen Arealen, aber auch im Bereich des STS im Vergleich zu synchronen audiovisuellen Sprachreizen. Ihr Augenmerk lag dabei auf Regionen, die sogenannte Supra- bzw. Subadditivität aufwiesen (s. 1.2.1). Auch Macaluso et al. (2004) untersuchten die Verarbeitung zeitlich und räumlich (in)kongruenter Wortstimuli in einer PET-Studie. Sie berichten eine Modulation des visuellen Kortex und des linken STS während der Darbietung asynchroner im Vergleich zu synchronen Stimuli. In einer fMRT-Studie, in der die Verarbeitung asynchroner einsilbiger Stimuli getrennt nach sensorischen und perzeptuellen Bedingungen analysiert wurde (Miller & D'Esposito, 2005), wurden sensorische Reize in den SC, im linken IPS, im rechten Parietallappen und in frontalen Arealen moduliert und perzeptuelle Modulationen im auditorischen Kortex und im STS gefunden. Van Atteveldt et al. (2007) zeigten anhand einer fMRT-Studie, dass zeitlich unterschiedlich gegeneinander verschobene audiovisuelle Silben verschiedene, aber überlappenden Bereiche des mittleren STS (mSTS) modulierten. Auditorische Areale wurden dagegen fast ausschließlich durch synchrone Reize moduliert. Bushara et al.

(2001) untersuchten in einer PET-Studie die Verarbeitung (a)synchroner einfacher audiovisueller Stimuli und fanden Modulierungen im rechten anterioren insulären Kortex, in posterior parietalen Bereichen, und im PFC. In einer fMRT-Studie wurde die automatische Verarbeitung einfacher, zeitlich nicht eindeutiger audiovisueller Reize mit synchronen verglichen. Synchronizität modulierte den STS in Kombination mit auditorischem und visuellem Kortex (Noesselt et al., 2007). Anhand von Konnektivitätsanalysen wiesen die Autoren nach, dass unisensorische Kortices über Rückkopplungsmechanismen mit dem STS interagieren, wenn physikalische Synchronizität vorliegt. Diese Studie zeigt außerdem, dass das menschliche Wahrnehmungssystem unter Inbezugnahme unisensorischer und heteromodaler Areale systematische zeitliche Relationen zwischen zwei Modalitäten präferentiell zu verarbeiten scheint. Dhamala et al. (2007) verwendeten im Rahmen einer fMRT-Studie einfache, kurze audiovisuelle Reize, die asynchron, synchron oder aufgrund hoher zeitlicher Frequenz der Abfolge der Ereignisse nicht eindeutig klassifizierbar waren. Dabei bestand das Netz perzeptueller Asynchronizität aus unisensorischen Kortices, dem Parietalkortex und präfrontalen Regionen und das Netzwerk perzeptueller Synchronizität aus unisensorischen und präfrontalen Kortices inklusive der SC. Das Netzwerk ambiguer Zeitperzepte bestand aus PFC und unisensorischen Kortices. Dhamala et al. (2007) schlossen, dass der Parietalkortex eine wichtige Rolle beim Erzeugen stabiler asynchroner Perzepte innehat und die SC beim Erzeugen stabiler synchroner Perzepte (s. 1.2.3). Die Verarbeitung der zeitlichen Verhältnisse einfacher audiovisueller Reize involviert laut den aufgeführten Befunden folglich ein kortikales Netzwerk aus unisensorischen Kortices und heteromodalen Gebieten in SC, STS, PFC und Parietallappen.

Zahlreiche bildgebende Verfahren untersuchten die Verarbeitung audiovisueller Sprachreize. Dabei wurden meist inhaltlich kongruente mit inkongruenten Inhalten in Bild und Ton verglichen. Die Stimuluskomplexität variierte dabei von einfachen Buchstaben, über Silben und Wörter. Viele dieser Studien fanden den STS in audiovisuelle Sprachwahrnehmung involviert (Callan et al., 2001; 2004; Calvert et al., 1997; 2000; 2001; Kawashima et al., 1999; Mottonen et al., 2006; Pekkola et al., 2006; Raij et al., 2000; Szycik et al., 2008a; 2008b; Tuomainen et al., 2005; van Atteveldt et al., 2004; 2007; Wright et al., 2003; s. auch Calvert & Lewis, 2004; Doehrmann & Naumer, 2008). Darüber hinaus modulierte audiovisuelle Sprache meist unisensorische Kortices (Calvert et al., 1999;

Kawashima et al., 1999; Miller & D'Esposito, 2005; Stevenson et al., 2009; Stevenson & James, 2009; van Atteveldt et al., 2007) und parietale Gebiete (Calvert et al. 2000; Miller & D'Esposito, 2005; van Atteveldt et al., 2007). Mittels ihrer PET-Untersuchung zeigten Kawashima et al. (1999), dass das Netz aus unisensorischen Arealen und STS nur moduliert wird, wenn keine der beiden Modalitäten selektiv attendiert wird. Dies legt die Vermutung nahe, dass dieses kortikale Netzwerk durch die Integration beider Modalitäten moduliert wird. Calvert (2001) weist darauf hin, dass unisensorische Areale über multisensorische Gebiete wie den STS und den Parietalkortex oder direkt miteinander kommunizieren, um eine effiziente Integration multimodaler Stimuli zu gewährleisten. EEG-Studien konnten zeigen, dass die Stimulation einer Modalität neuronale Aktivität in einer zweiten Modalität hervorrufen kann. So wurde gezeigt, dass einfache audiovisuelle Reize frühe Veränderungen der Aktivität in unisensorischen Kortices hervorrufen (Giard und Peronnet, 1999; Senkowski et al., 2007). Andere Autoren demonstrierten dies anhand audio-taktiler Reize (Fuxe et al., 2000; Murray et al., 2005). Des Weiteren konnte im Rahmen von fMRT-Studien gezeigt werden, dass auditorische Reize visuelle Kortices und vice versa modulieren (Laurienti et al., 2002; Martuzzi et al., 2007). FMRT- und Einzelzellstudien an Tieren konnten zeigen, dass primäre und sekundäre auditorische Areale durch taktile Stimuli moduliert wurden (Kayser et al., 2005; 2007; Lakatos et al., 2007). Darüber hinaus konnte bei Tieren gezeigt werden, dass auditorischer und visueller Kortex über direkte Bahnen miteinander kommunizieren können (Falchier et al., 2002; Rockland & Ojima, 2003).

Einige Autoren verglichen anhand von fMRT-Studien bei Menschen direkt bedeutungsvolle mit semantisch nicht belegten Reizen (Beauchamp et al., 2004b; Stevenson & James, 2009) und zeigten, dass die Antwortmuster im STS auf beide Stimulusarten identisch sind, die lokalen Maxima der Aktivierungen jedoch nicht überlappen, so dass vermutet werden kann, dass der STS als Integrationsareal anatomisch in stimulusabhängige Sektionen unterteilt ist. Des Weiteren wurde in einer fMRT-Studie am Menschen gezeigt, dass der STS in unisensorisch auditorische, visuelle und multisensorische Bereiche eingeteilt werden kann (Beauchamp et al., 2004a). Tierstudien belegen, dass der STS in getrennte, neurochemische Module aufgeteilt ist, die Verbindungen zu anderen kortikalen Gebieten aufweisen. Diese Bereiche innerhalb des STS könnten funktionelle Relevanz besitzen (Padberg et al., 2003).

Studien, die sich mit den neuronalen Grundlagen der Bedeutungsanalyse von Artgenossenlauten bei Affen beschäftigten, fanden Modulationen in visuellen, limbischen und präfrontalen Gebieten, aber vor allem im Bereich der beiden STS (Gilda-Costa et al., 2004; Rauschecker et al., 1995).

Zusammenfassend ist anhand der aufgeführten Befunde festzustellen, dass die Integration sprachlicher und nichtsprachlicher audiovisueller Stimuli scheinbar in bestimmten kortikalen Arealen stattfindet, die in mehreren Studien beschrieben wurden und miteinander vergleichbar sind. Der STS nimmt dabei neben Parietalkortex und PFC eine zentrale Position ein. Er wurde sowohl bei der Integration sprachlicher als auch nichtsprachlicher Reize moduliert. Dieses Gebiet scheint auch bei Tieren während der audiovisuellen Integration sowohl einfacher als auch semantisch belegter Reize von Bedeutung zu sein. In der Literatur gilt dieses Gebiet als multimodales Integrationsareal (Allison et al., 2000; Beauchamp et al., 2004a; Calvert, 2001; Hein & Knight, 2008). Darüber hinaus zeigen Studien an Tieren und Menschen, dass multisensorische Prozesse nicht an bestimmte kortikale Regionen geknüpft sind, sondern vielmehr aus einem Zusammenspiel unisensorischer Kortices und höherer, integrativer Areale entstehen. Anhand psychophysischer Befunde konnte demonstriert werden, dass multimodale Prozesse von Umgebungsparametern wie räumlicher, zeitlicher und semantischer Kongruenz, kognitiven Einflüssen (top-down-Prozessen), aber auch direkt von der Reizbeschaffenheit (bottom-up-Prozesse) abhängig sind (Doehrmann & Naumer, 2008). Die neuronalen Grundlagen dieser Prozesse wurden bisher nur teilweise untersucht. Die sich aus diesen Befunden ergebenden Fragestellungen für die vorliegende Arbeit werden im Folgenden besprochen.

### 1.3. Fragestellungen

Diese Arbeit fokussiert thematisch auf zeitliche und semantische Aspekte audiovisueller Reize. In einem ersten Experiment sollen im Rahmen einer fMRT-Untersuchung unisensorische und multisensorische Areale identifiziert werden, die in die perzeptuelle und wahrnehmungsunabhängige (physikalische) Verarbeitung synchroner und asynchroner semantisch nicht belegter Stimuli involviert sind. In einem zweiten fMRT-Experiment soll innerhalb dieser zuvor lokalisierten perzeptuellen Regionen

überprüft werden, inwieweit diese an adaptiven Prozessen audiovisueller Zeitlichkeit beteiligt sind. Schließlich soll anhand langer Sprachsequenzen in einem dritten fMRT-Experiment überprüft werden, ob multimodale Areale in Abhängigkeit unterschiedlicher Zeitperzepte funktionell unterteilt werden können.

### **1.3.1 Neuronale Grundlagen audiovisueller Zeitwahrnehmung einfacher Reize**

Aufgrund psychophysischer Befunde (Fujisaki & Nishida, 2005) sollen aufeinanderfolgende, audiovisuelle Reize mit einer Frequenz von 4 Hz präsentiert werden. Dabei wird der zeitliche Abstand zwischen Licht- und Tonreiz systematisch variiert. Da sich die Konfiguration kurzer aufeinanderfolgender audiovisueller Stimuli, die in ihrer Gleichzeitigkeit eingeschätzt werden sollen, auf die subjektive Beurteilung dieser Reize auswirkt (Vatakis & Spence, 2006), soll in Rahmen eines Verhaltensexperiments die Anzahl aufeinanderfolgender, zeitlich verschobener audiovisueller Ereignisse bestimmt werden, bei der diese Reize optimal bewertet werden können. Diese Reize werden für die fMRT-Analyse genutzt und sollen hinsichtlich ihrer Gleichzeitigkeit beurteilt und im Rahmen eines ereigniskorrelierten Modells analysiert werden. Darüber hinaus soll überprüft werden, welche Areale in ihren neuronalen Antwortmustern beobachtetes Verhalten in bezug auf die Beurteilung audiovisueller Gleichzeitigkeit nachvollziehen. Andere fMRT-Studien, die audiovisuelle Zeitprozesse anhand einfacher Reize untersuchten, nutzten Stimuli, die mit zu hoher Frequenz präsentiert wurden (Calvert et al., 2001) und daher nicht adäquat bewertet werden konnten (s. Fujisaki & Nishida, 2005), stellten keine experimentelle Aufgabe (Bushara et al., 2001) oder konfundierten Frequenzvariation und die Variation zeitlicher Abstände zwischen beiden Modalitäten (Dhamala et al., 2007). Keine dieser Studien identifizierte anhand von Lokalisierungsdurchgängen unisensorische und multimodale Gebiete, trennte die Effekte nach perzeptuellen und physikalischen Prozessen oder implementierte Verhaltenseffekte in die fMRT-Analyse, wie es in Experiment 1 dieser Arbeit geschieht. Experiment 1 lokalisiert unisensorische und bimodale Areale, die einerseits durch physikalische Prozesse und andererseits durch perzeptuelle Prozesse audiovisueller Zeitwahrnehmung moduliert werden. In der vorliegenden Arbeit wurden audiovisuelle

Stimulussequenzen verwandt, die sukzessive phasisch gegeneinander verschoben wurden. Dabei soll die experimentelle Fragestellung beantwortet werden, inwiefern unisensorische und multimodale Hirnareale audiovisueller Synchronizität bzw. Asynchronizität zugeordnet werden können. Da dieser Aufbau ein Experimentaldesign impliziert, das zuvor nicht verwandt wurde, hat diese Untersuchung explorativen Charakter und verzichtet auf die Formulierung expliziter Hypothesen. Da außerdem bisher keine expliziten Modelle audiovisueller Zeitmechanismen vorliegen, soll vielmehr der Versuch unternommen werden, einen konzeptuellen Rahmen zu generieren, der alle hier vorgestellten Resultate integrieren kann.

### **1.3.2 Neuronale Grundlagen audiovisueller Zeitlichkeitsadaptation**

In mehreren bisherigen Studien wurde nachgewiesen, dass sich das menschliche Wahrnehmungssystem an unterschiedliche zeitliche Relationen multimodaler Reize anpassen kann (s. 1.2.6.1). Die neuronalen Grundlagen dieser Prozesse wurden jedoch bisher nicht untersucht. Anhand mehrerer psychophysischer Vorexperimente soll bestimmt werden, welche Parameter adaptive Prozesse audiovisueller Zeitwahrnehmung beeinflussen. Die anhand dieser Befunde optimierten Reize werden im Rahmen eines fMRT-Experiments verwandt. Dieses Experiment wendet sich der Fragestellung zu, welche Areale durch bzw. auf welcher Verarbeitungsebene stimulusgetriebene und zeitliche Veränderungen einfacher audiovisueller Reize adaptiv moduliert werden. Da bisher keine Befunde neuronaler Grundlagen dieser Prozesse vorliegen, wird diese Untersuchung in einem explorativen Rahmen durchgeführt und keine Annahmen über korrelierte Aktivität bzw. Beteiligung von bestimmten Arealen formuliert.

### **1.3.3 Funktionelle Korrelate von Zeitprozessen audiovisueller**

#### **Sprache**

Anhand langer, ökologisch valider audiovisueller Sprachreize sollen anhand eines fMRT-Experiments die neuronalen Grundlagen zeitlicher Wahrnehmung audiovisueller Sprache identifiziert werden. Dabei werden zeitlich schwellennahe Reize verwandt, die einen Perzeptwechsel zulassen. In einem Vorexperiment sollen individuelle Schwellen für sprachliche Synchronizität bzw. Asynchronizität determiniert werden und daraufhin

während des fMRT-Experiments verwandt werden. Dieses Design ermöglicht eine Trennung entscheidungsabhängiger, physikalischer und stabiler perzeptueller Prozesse und kann daher die funktionellen Parameter stabiler Zeitperzepte aufdecken. Neben einem kortikalen Netz audiovisueller Zeitlichkeitswahrnehmung soll der STS als multimodales Integrationsareal auf funktionelle Repräsentationen zeitlicher Relationen externer Stimuli hin untersucht werden. Darüber hinaus soll anhand von Interaktionsanalysen überprüft werden, welche Gebiete durch funktionelle Muster audiovisueller Zeitperzepte komoduliert werden.

Dies ist insbesondere relevant, da keine der oben aufgeführten Studien zur semantischen, audiovisuellen Sprachwahrnehmung Reize untersuchte, die länger als ca. 1 Sekunde waren. Die hier verwendeten Sequenzen repräsentieren im Vergleich zu bisherigen Untersuchungen Stimuli, die alltäglichen Spracherlebnissen ähneln. Darüber hinaus bestimmte keine der Studien individuelle (A)Synchronitätsschwellen, trennte nicht unisensorische von bimodalen Regionen und untersuchte nicht die Konnektivität der durch ihre Effekte modulierten Areale. Dies führt dazu, dass keine Hypothesen über die zugrundeliegenden neuronalen Mechanismen formuliert werden können, sondern die Fragestellung explorativ untersucht wird.

Die Ergebnisse dieser Untersuchungen werden die Identifizierung eines kortikalen Netzwerks audiovisueller Zeitlichkeitsverarbeitung ermöglichen. Dabei können anhand der vorliegenden Stimuluskonfiguration, die bisher nicht verwandt wurde, präzise Rückschlüsse auf die Art der zugrundeliegenden neuronalen Mechanismen gezogen werden. So können unisensorische, multimodale, perzeptuelle, verhaltenskorrelierte, physikalische und stimulusgetriebene Prozesse voneinander differenziert werden. Darüber hinaus können semantische mit semantisch nicht determinierten neuronalen Korrelaten verglichen werden. Des Weiteren kann überprüft werden, wie multimodale Areale unterschiedliche Zeitrelationen funktionell abbilden.

## 2 Allgemeine Methodik

Im Folgenden werden in einer kurzen Übersicht die dieser Arbeit zugrundeliegenden Methoden vorgestellt und die Wahl der jeweiligen Verfahren begründet. Zunächst werden psychophysische Maße vorgestellt, die es ermöglichen, Schlüsse über die Zeitlichkeitswahrnehmung audiovisueller Reize zu ziehen. Daraufhin werden die Grundlagen neurophysiologischer Maße beschrieben, die Prinzipien ihrer Erfassung und die Grundlagen ihrer Analyse.

### 2.1. Psychophysische Maße

#### 2.1.1 JND und PSS

Zur Quantifizierung subjektiver Zeiteindrücke sollen zwei Maße benutzt werden: der PSS (point of subjective simultaneity) kann als Maß für das Zeitintervall, durch das ein sensorisches Ereignis dem anderen vorangehen muss, um als synchron wahrgenommen zu werden beschrieben werden, und die JND (just noticeable difference) als standardisiertes Sensitivitätsmaß, mit dem Probanden die zeitliche Reihenfolge zweier Ereignisse unterschiedlicher Wahrnehmungsmodalität erkennen können (Spence et al., 2001; Stone et al., 2001). Die JND kann als Zeitfenster verstanden werden, innerhalb dessen die Integration von Ton- und Lichtreiz gelingt und als gleichzeitig wahrgenommen wird. Ein enges Zeitfenster bzw. eine kleine JND spiegelt erhöhte Sensitivität für Asynchronizität wider. Die beiden hier besprochenen Parameter sollen in den folgenden Untersuchungen als Maße für die zeitliche Integrationsleistung von unterschiedlichen Stimuli bei Probanden herangezogen werden.

Um die Maße für JNDs (just noticeable differences) und PsSS (points of subjective simultaneity) zu schätzen, wurde für jede Verzögerungsbedingung der Anteil synchroner Urteile für jeden einzelnen Probanden ermittelt. Diese Werte wurden nach dem Ausmaß der Verschiebungsbedingung geordnet und einer abgeschnittenen Gauss-Funktion per Maximum-Likelihood-Schätzung (Myung, 2003) angepasst (s. Formel 3). Dies ist möglich, da der Anteil synchroner Urteile für extreme Verzögerungen (SOAs, Stimulus Onset Asynchrony) typischerweise gering ist, für Stimuli, die sich physikalischer Gleichzeitigkeit nähern größer wird, bei 0 ms Verzögerung maximal wird und am

entgegengesetzten Extrem wieder abnimmt. Dieser Werteverlauf spiegelt eine (abgeschnittene) Glockenkurve ohne asymptotische Annäherung an die x-Gerade wider.

$$a \times \exp - \frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2$$

**Formel 3.** Dichtefunktion, die an die Werteverteilung der Anteile synchroner Urteile angepasst wurde.

*a*: y-Koordinate des Kurvenmaximums; *x*: Messwert (SOA); *μ*: Mittelwert (Schätzwert für subjektive Gleichzeitigkeit); *σ*: Standardabweichung (Schätzwert für JND)

Die Ausdehnung (Standardabweichung) der geschätzten Funktion ist ein indirektes Maß für die JNDs, da sie einen umgekehrten/negativen Zusammenhang mit der zugrundeliegenden Steigung einer psychometrischen Funktion aufweist, und daher ein Maß dafür ist, wie schwer Probanden die Gleichzeitigkeitsurteile für die getesteten SOAs fanden (Zampini et al., 2005). Der Mittelwert der Funktion gibt den PSS wieder, also den optimalen zeitlichen Abstand, den eine Modalität gegenüber der anderen haben soll, damit Probanden einen bimodalen Stimulus als gleichzeitig bewerten. Zampini et al. (2005) variierten die Spanne der präsentierten SOAs audiovisueller Stimuli, um anhand der hier vorgestellten Methode JNDs und PsSS experimentell zu erfassen und zu überprüfen, ob Probanden einer gewissen Antworttendenz folgen und Antworten gleichmäßig über alle Experimentalbedingungen (SOA) verteilen, oder ob die abgeleiteten Maße von JND und PSS perzeptuelle Effekte repräsentieren. Die Autoren erhielten stabile Schätzungen unabhängig von der Variation der Spanne der SOAs. Demzufolge ist davon auszugehen, dass die hier verwendeten Maße als Indikatoren perzeptueller Effekte im Gegensatz zu Antworttendenzen zu betrachten sind. Offensichtliche Antwortverzerrungen wurden nicht analysiert und anhand etablierter statistischer Kriterien (Spence et al., 2001; Vatakis et al., 2007) Fälle eliminiert, die bestimmte Antwortmuster aufwiesen (s. jeweilige Abschnitte zur Methodik der Experimente).

Aus den angepassten Glockenkurven lässt sich ein dritter Wert ableiten: das Kurvenmaximum. Dies ist der Punkt, an dem die Dichtefunktion auf der Y-Koordinate ihren Höhepunkt erreicht. Dieser Wert kann als Zuverlässigkeit interpretiert werden, mit der Probanden ihre jeweiligen Beurteilungen machen. Da dieser Schätzer psychometrisch

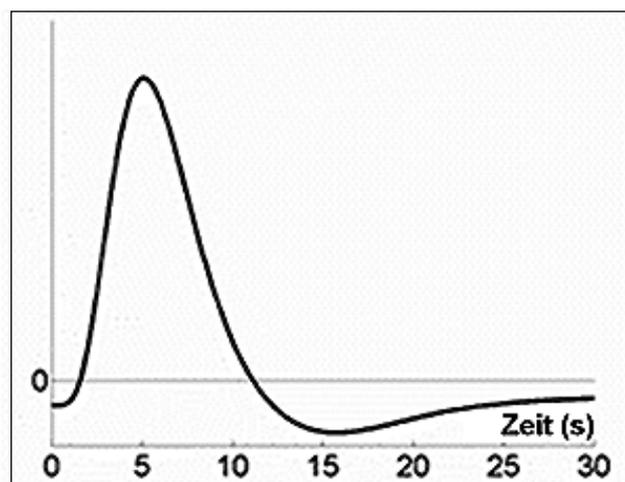
nicht zu deuten ist, bzw. keine relevanten psychophysikalischen Schlussfolgerungen zulässt, wird er daher im Folgenden nicht analysiert.

## 2.2. Funktionelle Magnetresonanztomographie

In diesem Kapitel soll fMRT, die eine hochauflösende Abbildung neuronaler Prozesse im Menschen ermöglicht und in dieser Arbeit benutzt wurde, erläutert werden. FMRT ist eine nicht-invasive bildgebende Technik, die magnetische Eigenschaften von Blut ausnutzt bzw. das Verhältnis von Oxy- zu Deoxyhämoglobin, was als BOLD-Kontrast (Blood Oxygen Level Dependent) in den Blutgefäßen bezeichnet wird und als ein Maß für neuronale Aktivität herangezogen wird (Ogawa et al., 1990; 1993). Es konnte gezeigt werden, dass Signale, die mit Hilfe der fMRT entdeckt werden, direkt mit dem Ausmaß neuronaler Antworten zusammenhängen, wobei die fMRT-Signale wahrscheinlich den synaptischen Informationsfluss reflektieren, der entsteht, bevor Neurone selbst ihre Feuerrate erhöhen (Logothetis et al., 2001; Logothetis, 2003). Die Grundlage der Magnetresonanz ist der sogenannte ‚Spin‘ von Atomkernen (z. B. Wasserstoff), der sich mit dem oder gegen das umgebende(n) Magnetfeld ausrichten kann und daher entsprechend geringe oder hohe energetische Zustände annehmen kann. Während der Aufnahme von fMRT-Bildern sorgt ein statisches Magnetfeld für eine Ausrichtung der Wasserstoff-Spins entlang des Feldes, bevor transversale Radiofrequenzen induziert werden, die ein Auslenken von Achsen und Phasen der Kernspins bewirken. Die gemessenen Parameter sind das Ergebnis der Zeit, die die Spins benötigen, um in den durch das statische Magnetfeld determinierten Zustand zurückzukehren, nachdem die transversale Radiofrequenz wieder abgeschaltet wurde. Diese Parameter verändern sich in Abhängigkeit der aktuellen Oxygenierung des Blutes. Ein großer Vorteil der MRT besteht darin, dass Bilder von neuronalen Prozessen mit hoher räumlicher Auflösung nicht-invasiv gewonnen werden können. Dabei können anatomische Gebiete präzise voneinander differenziert werden, und die funktionellen Aufnahmen ermöglichen Teststatistiken, anhand derer experimentelle Bedingungen miteinander verglichen werden können. Dies wiederum ermöglicht Rückschlüsse auf kognitive Prozesse. Da jedoch eine BOLD-Antwort bis mehr als 12 Sekunden nach Stimuluspräsentation anhalten kann (s. Abb. 7), ist die zeitliche

Auflösung der MRT im Vergleich zu anderen Messverfahren wie EEG oder MEG vergleichsweise gering.

Um zu explorieren, welche Hirnareale in audiovisuelle Integrationsleistungen involviert sind, scheint die Untersuchung anhand fMRT geeignet, da diese Methodik es erlaubt, wie oben erwähnt, neuronalen Blutfluss räumlich sehr genau (2 mm) abzubilden. Andere bildgebende Verfahren (z. B. MEG, EEG) haben zwar den Vorteil, genaue Informationen über den exakten Zeitverlauf neuronaler Antworten zu liefern, viel weniger genau jedoch über den Ort der Aktivitäten. Da im Rahmen dieser Arbeit kortikale Gebiete, die in die audiovisuelle Synchronizitätsverarbeitung involviert sind, exploriert werden sollen und an verschiedenen, über den Kortex verteilten Regionen vermutet werden (z. B. Calvert, 2001, Driver & Noesselt, 2008) scheint die fMRT das bestgeeignete Instrument für die hier durchgeführten Untersuchungen zu sein.



**Abbildung 7.** Zeitlicher Verlauf einer BOLD-Antwort (hämodynamische Antwortfunktion, HRF).

Darüber hinaus ermöglicht die spätere Datenanalyse der fMRT-Daten, unisensorische von potentiell multisensorischen (hier: audiovisuellen) Arealen durch das Ermitteln von sogenannten auditorischen bzw. visuellen Masken, und deren Überlappung zu trennen. Da in den hier durchgeführten Studien unter anderem abgeklärt werden soll, ob ermittelte multisensorische Wahrnehmung in typischerweise unisensorischen Arealen oder in Gebieten, die auf die Verarbeitung mehrerer Wahrnehmungsmodalitäten spezialisiert sind, stattfindet, bietet die fMRT in dieser Hinsicht für die Auswertung der vorliegenden Untersuchungen einen Vorteil verglichen mit anderen bildgebenden Verfahren.

Anhand psychophysikalischer Interaktionsparameter (Friston et al., 1997) können mit fMRT-Daten zudem Konnektivitätsanalysen durchgeführt werden, die Rückschlüsse darüber zulassen, inwiefern Areale, die durch multisensorische Wahrnehmung moduliert werden, mit anderen Gebieten interagieren bzw. verknüpft sind. Anhand dieses Vorgehens lassen sich demzufolge kortikale Verknüpfungen bzw. Netzwerke multisensorischer Verarbeitung beim Menschen ermitteln. Diese Art von Erkenntnisgewinn stellt in bezug auf die hier durchgeführten Untersuchungen einen weiteren Vorteil der fMRT dar, da es u.a. Ziel der Untersuchungen ist, Areale zu identifizieren, mit denen potentiell multisensorische Hirnregionen interagieren. Dies gewährt Einblick in kortikale Kommunikationsmuster, wenn Probanden Stimuli multisensorisch integrieren.

## 2.3. Analyse der fMRT-Daten

Die Auswertung der fMRT-Daten der hier vorgestellten Studie wurde mit SPM2 (Experiment 3) und SPM 5 (Experimente 1 und 2; Wellcome Department for Cognitive Neurology, London, UK) durchgeführt. Mit Hilfe dieser Programme können die notwendige Vorverarbeitung der fMRT-Daten und deren anschließende statistische Auswertung vorgenommen werden. Die Statistik beruht auf der Anwendung des allgemeinen linearen Modells (ALM) und wird durch Einführung spezieller Parameter, Filter oder Modulationen entsprechend angepasst (s. Jäncke, 2005).

Die abhängigen Variablen in einem fMRT-Datensatz sind Volumeneinheiten (sogenannte Voxel), deren Werte über die Zeit aufgezeichnet werden. Die Datenanalyse zielt darauf ab, den Verlauf dieser Voxelzeitreihen anhand des modellierten BOLD-Signals vorherzusagen und gleichzeitig Abweichungen von der Vorhersage zu bestimmen. Dabei wird davon ausgegangen, dass die hämodynamische Antwortfunktion (HRF) während der experimentellen Bedingungen steigt und in der Ruhebedingung fällt.

### 2.3.1 Das Allgemeine Lineare Modell

Dieser Algorithmus zielt darauf ab, gemessene Daten bzw. abhängige Variablen (AV) durch eine lineare Kombination von erklärenden bzw. unabhängigen Variablen (UV) vorherzusagen (s. Formel 4). Die Struktur des ALM ist immer:

$$Y = X_{i1} \cdot \beta_{i1} + X_{i2} \cdot \beta_{i2} + \dots + X_{ik} \cdot \beta_{ik} + E_i$$

**Formel 4.** Allgemeines lineares Modell.

$Y$ : gemessenes MR-Signal (AV);  $X$ : erklärende Variable (UV);  $\beta$ : Parametergewicht;  $E$ : Fehlerterm,  $i$ : Anzahl beobachteter Werte;  $k$ : Anzahl der erklärenden Variablen

Anhand des ALM wird versucht, einen hypothetischen Verlauf des BOLD-Signals auf die gemessenen Daten anzupassen. Da keine exakte Vorhersage der AV möglich ist, muss ein Fehlerterm in das Modell aufgenommen werden.

Die Bestimmung der fMRT-Daten als Linearkombination aus erklärenden und Fehlervariablen wird an jedem Voxel im Gehirn vorgenommen (Friston et al., 1995). Um die Daten möglichst gut vorherzusagen, werden für die hypothetischen Antwortfunktionen Gewichte ( $\beta_i$ ) nach der Kleinstquadratmethode (Hu et al., 2005) geschätzt, um die Fehlerterme möglichst gering zu halten. Diese Gewichte werden für verschiedene Experimentalbedingungen unterschiedlich hoch geschätzt und stellen damit die Grundlage für den Vergleich verschiedener Experimentalbedingungen und folglich der inferenzstatistischen Hypothesentestung dar (Hu et al., 2005).

Die tatsächlichen Daten eines fMRT-Datensatzes werden mittels der Design-Matrix mit den Parametern des ALM in Beziehung gesetzt. Dabei stellt jede Spalte eine Experimentalbedingung und jede Zeile eine Aufnahme (scan) dar.

## 2.3.2 Auswertung mit SPM

### 2.3.2.1 Vorverarbeitung

Zusätzlich zu den automatisch verworfenen ersten Messungen eines jeden Runs, blieben die ersten fünf akquirierten Bilder eines jeden Blocks (Lokalisationsdurchgang und Experimentaldurchgänge) der funktionellen Messungen bei der Analyse unberücksichtigt, um sicher zu stellen, dass die Spins vollständig saturiert waren, und damit der sogenannte ‚steady state‘ erreicht war (Frahm et al., 1994; Goerke et al., 2005).

Vor der statistischen Analyse der Daten wurden die verbliebenen funktionellen Volumen mittels SPM2 (Experiment 3) bzw. SPM5 (Experiment 1 und 2; Wellcome Department for Cognitive Neuroscience, London, UK) mit einer affinen rigid-body-

Transformation bewegungskorrigiert (realignment). Für das Realignment werden alle Bilder einer Zeitreihe angepasst, indem die Parameter der Relativbewegungen (drei Translationen und drei Rotationen) eines jeden Bildes bestimmt und anschließend angewendet werden. Danach wurden die Bilder in den Quasi-Talairach-Raum (MNI-Raum, SPM99, Cocosco et al., 2003) überführt (Normalisierung auf IR-EPI-Bilder), wobei sie interpolativ in 2 mm<sup>3</sup> große kubische Voxel umgewandelt wurden. Dazu wurde ein iterativer, nicht-linearer Algorithmus verwendet. Als letzter Schritt der Vorverarbeitung wurden die Daten räumlich geglättet (dreidimensionaler Gaußfilter, full width half maximum-Filter: 6mm). Zusätzlich wurde das normalisierte strukturelle T<sub>1</sub>-gewichtete Hirnvolumen auf die funktionellen Volumen koregistriert, um später fMRT-Ergebnisse der Einzelprobanden auf dem jeweiligen neuroanatomischen Substrat abbilden zu können.

### **2.3.2.2 Gruppenanalyse**

Die Auswertung der Gruppenstatistik erfolgte anhand einer Random-Effects-Analyse, die, im Gegensatz zur Fixed-Effects-Analyse, Probanden als unabhängig voneinander behandelt (Friston et al., 1999), und mit der die oben beschriebenen Kontraste bestimmt und auf die Gesamtpopulation bezogen interpretiert werden können.

## 3 Experimente

Anhand dreier fMRT-Experimente sollen die Fragestellungen 1.3.1, 1.3.2 und 1.3.3 nacheinander untersucht werden. Dementsprechend ist Experiment 1 dazu konzipiert, stimulus- und perzeptabhängige Prozesse einfacher phasenverschobener und phasengleicher audiovisueller Reize und deren neuronale Korrelate zu determinieren. Im Rahmen von Experiment 2 werden daraufhin kurzfristige Änderungen audiovisueller Zeitwahrnehmung und deren neuronale Grundlagen untersucht, die das Wahrnehmungssystem nutzt, um vorübergehende Änderungen der Umweltreize erfolgreich zu verarbeiten (Adaptation). Dies soll ermöglichen, neuronale Modulationen durch automatische Prozesse bzw. langfristige Einflüsse und durch kurzzeitige Einflüsse auf audiovisuelle Zeitwahrnehmung getrennt voneinander zu analysieren. In Experiment 3 werden schließlich audiovisuelle Sprachreize verwendet, um Zeitwahrnehmungsmechanismen anhand semantischer Prozesse zu untersuchen.

### 3.1. Experiment 1

#### 3.1.1 Einführung

Diese Untersuchung soll Aufschluss darüber geben, wie phasengleiche bzw. phasenverschobene einfache audiovisuelle Stimulussequenzen (hier: Sequenzen bestehend aus kurzen Tönen und Lichtblitzen) hinsichtlich ihrer Gleichzeitigkeit wahrgenommen werden, und welches neuronale Korrelate dieser Wahrnehmungsprozesse sind (s. 1.3.1). Im Folgenden werden phasengleiche Stimulussequenzen als synchrone und phasenverschobene Sequenzen als asynchron bezeichnet. Es sollen einzelne audiovisuelle Ereignisse so aneinandergereiht werden, dass die Stimuli sich etwa über eine Sekunde erstrecken und dass Licht und Ton phasisch in steigenden Abständen gegeneinander verschoben werden, ohne dass einzelne Ereignisse zeitlich überlappen. Fujisaki und Nishida (2005) konnten in einer Untersuchung zeigen, dass einfache Stimuli, die mit einer Frequenz von mehr als 4 Hz präsentiert werden, nicht mehr zuverlässig hinsichtlich ihrer Gleichzeitigkeit bewertet werden können. Die audiovisuellen Signale, die in der vorliegenden Studie benutzt werden, sollen daher mit einer Frequenz von 4 Hz gezeigt werden. Bisherige Studien, die die audiovisuelle

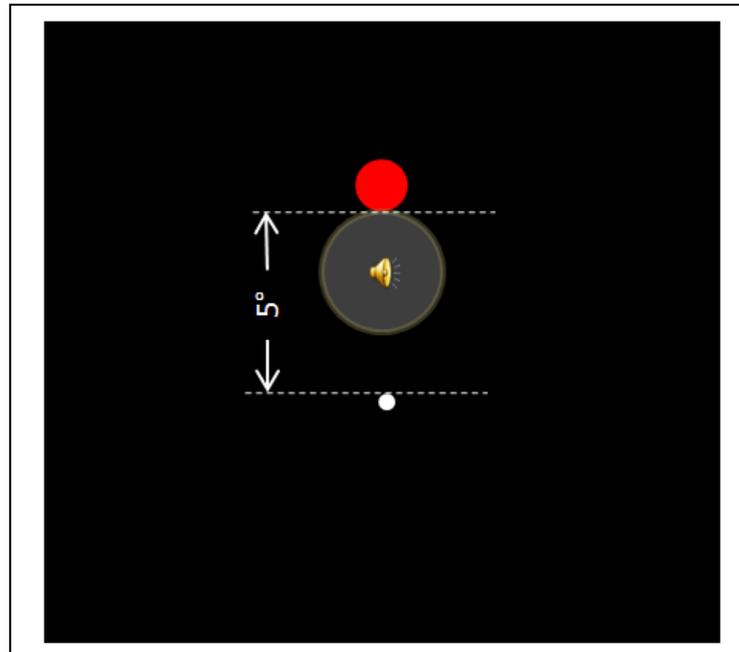
Integration einfacher Stimuli untersuchten (z. B. Fujisaki et al., 2004; Recanzone, 2003; Vatakis et al., 2007; 2008; Vroomen et al., 2004), verwendeten Lichtblitze und Tonsignale, die in unterschiedlicher Anzahl aneinander gereiht wurden. Der Einfluss der Stimulusanzahl (Anzahl aufeinanderfolgender audiovisueller Ereignisse) auf den Beurteilungsprozess wurde jedoch bisher nicht systematisch untersucht. Daher wurde hier eine Voruntersuchung durchgeführt, in der die Anzahl aufeinanderfolgender audiovisueller Ereignisse systematisch variiert wurde, um zu klären, bei welcher Stimuluslänge die Gleichzeitigkeitsbeurteilung audiovisueller Reize gegebenenfalls leichter oder schwerer bzw. am genauesten gelingt. Die Erkenntnisse aus bestehenden Studien und denjenigen der Voruntersuchung sollen Hinweise für den Aufbau einer fMRT-Studie geben, die geeignet ist, um neuronale Grundlagen audiovisueller Integration einfacher Reize zu ermitteln.

### 3.1.2 Vorexperiment 1

#### 3.1.2.1 Methodik

Sechzehn rechtshändige (Selbstbericht) Probanden (Durchschnittsalter 23, Spanne: 18-29, 8 weiblich) saßen in 57 cm Abstand vor einem Lautsprecher (Durchmesser: 7 cm), in dessen Zentrum ein weißes Fixationslicht (Leuchtdiode, LED) und ein rotes Ziellicht 5° Sehwinkel darüber angebracht waren (s. Abb. 8). Alle Teilnehmer berichteten, normales Hörvermögen und normales oder korrigiertes Sehvermögen und keine psychiatrischen oder neurologischen Störungen zu haben. Sie variierten hinsichtlich ihrer Vorerfahrung mit der Teilnahme an psychophysikalischen Studien und wurden über deren Zweck aufgeklärt. Das Experiment dauerte ca. 40 Minuten und wurde in einem vollständig dunklen und schallisolierten Raum durchgeführt.

Der auditorische Stimulus bestand aus einem reinen Sinuston (Frequenz: 1800 Hz, Samplingrate: 24420 Hz) und wurde durch den Lautsprecher direkt über der Leuchtdiode präsentiert, so dass auditorischer und visueller Reiz denselben räumlichen Ursprung zu haben schienen. Die Signaltöne wurden mit einer Lautstärke von 65 dB (SPL) ausgehend von der Ohrenposition der Probanden präsentiert. Der visuelle Reiz bestand aus einer roten Leuchtdiode, die pro Ereignis 30 ms bei einer Luminanz von 60 cd/m<sup>2</sup> leuchtete.



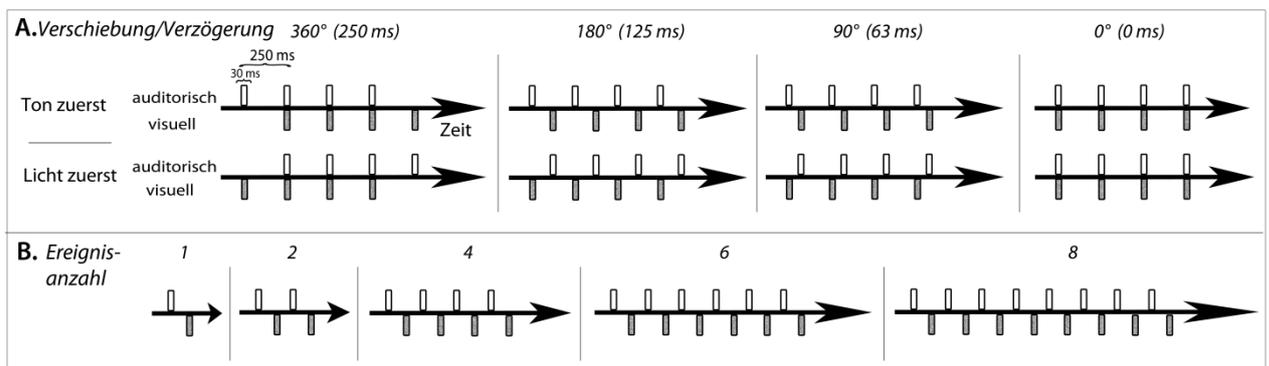
**Abbildung 8.** Experimentalaufbau für Experiment 1 und das dazugehörige Vorexperiment. Eine LED (Zielreiz) wurde 5° Sehwinkel über einem weißen Fixationspunkt angebracht. Sie konnte rote Lichtsignale senden. Töne wurden über einen Piezolausprecher direkt unter der LED präsentiert. Siehe Text für nähere Erläuterungen.

Die Stimuli, die den Probanden gezeigt wurden, bestanden aus einem, zwei, vier, sechs oder acht auditorischen Signaltönen und visuellen Lichtreizen, die miteinander kombiniert wurden und mit einer Frequenz von 4 Hz aufeinanderfolgend dargeboten wurden. D.h. dass der zeitliche Abstand zwischen audiovisuellen Ereignissen 250 ms betrug, falls Licht- und Tonreiz synchron dargeboten wurden. Die Anzahl der Ereignisse wurde gewählt, damit die Dauer einer audiovisuellen Sequenz ungefähr die Länge eines gesprochenen Wortes widerspiegelt.

Diese audiovisuellen Sequenzen waren entweder phasengleich oder phasenverschoben: asynchrone Sequenzen entstanden, indem die auditorischen Tonreihen phasisch um 90°, 180°, oder 360° gegen die Lichtreihen in beide Richtungen (Ton zuerst oder Licht zuerst) verschoben wurden. Es existierten demzufolge sieben Verzögerungsbedingungen (stimulus onset asynchronies, SOA): eine synchrone (keine Verzögerung) und sechs asynchrone. Für asynchrone Stimuli betrug die zeitliche Verzögerung zwischen Licht- und Tonreihe 63 ms bei 90°, 125 ms bei 180° und 250 ms bei 360° Phasenverschiebung (s. Abb. 9). Die sieben Verzögerungsbedingungen wurden mit

den fünf oben genannten Längenbedingungen kombiniert, so dass sich ein Experimentaldesign von  $7 \times 5$  Bedingungen ergab.

Probanden wurden instruiert, eine Taste auf einer Computertastatur zu drücken, wenn ein Stimulus ihnen den subjektiven Eindruck zeitlicher Synchronizität vermittelte und eine zweite Taste, um einem Reiz als asynchron zu bewerten. Sie wurden zudem gebeten, jeden Stimulus direkt nach seiner Darbietung zu bewerten, nicht aber während einer Sequenz und nicht, nachdem die nächste Sequenz bereits präsentiert würde. Vor jedem Experiment wurde ein Übungsdurchlauf durchgeführt, während dem alle vorkommenden Stimuluskombinationen dargeboten wurden. Der Versuchsleiter gab während des Übungsdurchgangs Rückmeldung über korrekte bzw. nicht korrekte Beurteilungen. Nachdem die Probanden phasengleiche (synchrone) und phasenverschobene (asynchrone) Reize jeweils mindestens zweimal korrekt detektiert hatten, wurde mit dem Hauptexperiment begonnen, andernfalls der Übungsdurchgang wiederholt.



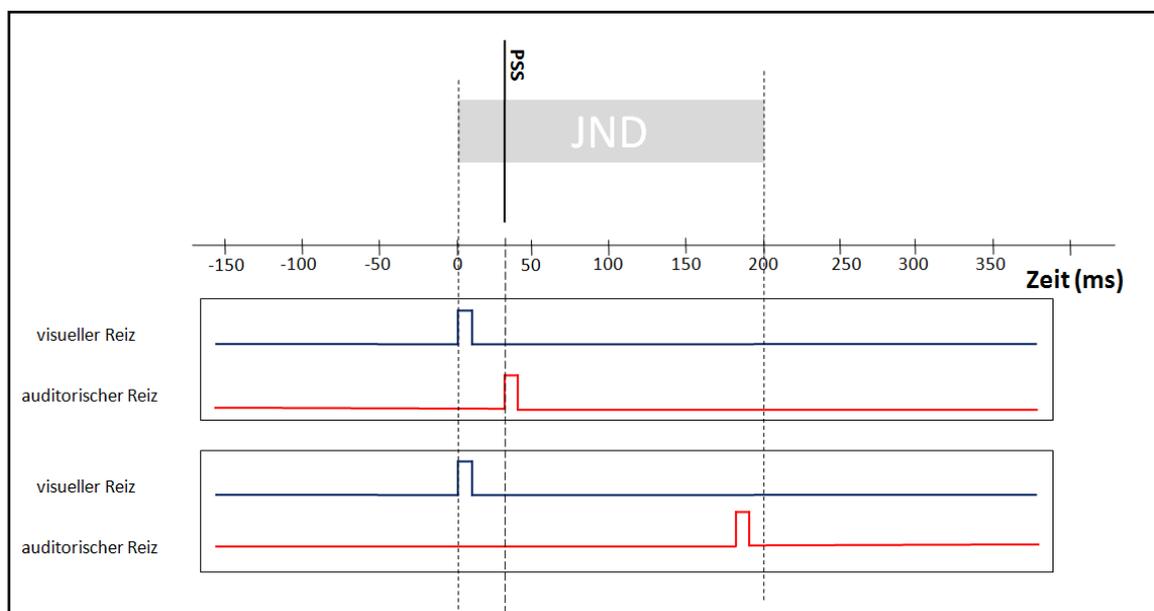
**Abbildung 9.** Schema der experimentellen Stimulation unter Vorexperiment 1 am Beispiel einer Sequenz mit vier Ereignissen. **A.** Licht- und Tonreize wurden phasisch in 7 (Un-)Gleichzeitigkeitsbedingungen (SOAs) gegeneinander verschoben (Faktor 1: SOA; A: 7 Stufen). **B.** Schema der experimentellen Stimulation illustriert am Beispiel einer Phasenverschiebung von Licht- und Tonreizen um  $180^\circ$ . Stimuli konnten eins, zwei, vier, sechs, oder acht Ereignisse enthalten (Faktor 2: Ereignisanzahl; B: 5 Stufen).

Ein Stimulus konnte mindestens 30 ms (Ereignislänge eins und keine Verzögerung) und maximal 2 Sekunden (Ereignislänge acht bei 250 ms Verzögerung) dauern (s. Abb. 9). Das Interstimulusintervall (ISI) betrug 1,5 Sekunden ab Stimulusende (Offset), so dass die Probanden mindestens 1,5 Sekunden Zeit hatten, auf einen Stimulus zu reagieren. Das ISI variierte demzufolge von 1530 ms bis 3500 ms. In zehn Durchgängen wurden je 70 Stimuli präsentiert. Pro SOA und Ereignislänge wurden 20 Stimuli

dargeboten, so dass eine Ereignislänge insgesamt 140 Mal mit sieben verschiedenen SOAs präsentiert wurde. Die Verzögerungsbedingungen der Stimuli wurden in randomisierter Reihenfolge dargeboten.

### 3.1.2.2 Fragestellung und Hypothesen

Zentrales Ziel dieses Vorexperimentes ist es, zu ermitteln, wie audiovisuelle Integration einfacher Reize von der Stimulusanzahl beeinflusst wird, um den Parameterraum so einzugrenzen, dass geeignete Stimuli identifiziert werden können, um in einem nachfolgenden fMRT-Experiment eine zuverlässige Erfassung der neuronalen Grundlagen audiovisueller Integration zu ermöglichen.



**Abbildung 10.** Stilisierte Darstellung von gleichzeitigkeitsabhängiger JND (Just Noticeable Difference) und PSS (Point of Subjective Simultaneity). Zwei Reize unterschiedlicher Modalität werden zeitlich zu einem multimodalen Ereignis integriert, solange der zweite Reiz innerhalb eines gewissen Zeitfensters nach dem ersten erscheint (JND). Obwohl dieses Zeitfenster ca. 200 ms breit sein kann, wird der Abstand zweier multimodaler Ereignisse von ca. 30 ms als optimal empfunden, damit Ereignisse als gleichzeitig wahrgenommen werden (PSS).

Um das Ausmaß der Integration der präsentierten audiovisuellen Reize zu überprüfen, wurden als abhängiges Maß PsSS (Punkte subjektiver Gleichzeitigkeit) und JNDs (gerade noch wahrnehmbare Unterschiede) berechnet. Der PSS kann als Maß für das Zeitintervall, durch das ein sensorisches Ereignis dem anderen vorangehen muss, um als synchron wahrgenommen zu werden, beschrieben werden, die JND als

standardisiertes Sensitivitätsmaß, mit dem Probanden die zeitliche Reihenfolge zweier unterschiedlicher Stimuli erkennen können. Die JND kann als Zeitfenster verstanden werden, innerhalb dessen die Integration von Ton- und Lichtreiz gelingt und als gleichzeitig wahrgenommen wird. Ein enges Zeitfenster bzw. eine kleine JND spiegelt erhöhte Sensitivität für Asynchronizität wider (s. Abb. 10).

Für das nachfolgende fMRT-Experiment können solche Stimuli als geeignet betrachtet werden, denen gegenüber Probanden sensitiver sind als weniger geeigneten. Bei weniger geeigneten Stimuli sind die JNDs demzufolge größer als bei geeigneten. Die Stimulusanzahl pro Sequenz (SAPS), bei der die Integration von Licht- und Tonreihe gut gelingt, wird folglich anhand kleiner Integrationszeitfenster (= JNDs) bei den Probanden zu identifizieren sein. Bei entsprechenden Stimuli fällt die Beurteilung hinsichtlich ihrer (Un-)Gleichzeitigkeit leichter, da Licht- und Tonreiz zeitlich näher beieinander liegen können als bei Stimuli mit großen JNDs, um schon als asynchron erkannt zu werden. Der optimale zeitliche Abstand zwischen Licht- und Tonreiz, den Probanden benötigen, um einen Stimulus als synchron wahrzunehmen (PSS) sollte, sofern die Stimuli unterschiedlicher SAPS untereinander vergleichbar sind, von der SAPS nicht beeinflusst werden. Aus dem Ausmaß des PSS kann keine entgeltliche Schlussfolgerung bezüglich der Eignung von Stimuli gezogen werden, da dieses Maß per se sehr variabel ist (z. B. Vatakis et al., 2007). Falls die SAPS den PSS systematisch beeinflusst, müsste SAPS als erklärender Faktor in Folgeexperimenten aufgeführt werden. Falls SAPS umgekehrt keinen Einfluss auf den PSS hat, können Sequenzen unterschiedlicher Anzahl als experimentell vergleichbar gewertet werden.

Aus dem oben Genannten ergeben sich zwei Fragstellungen:

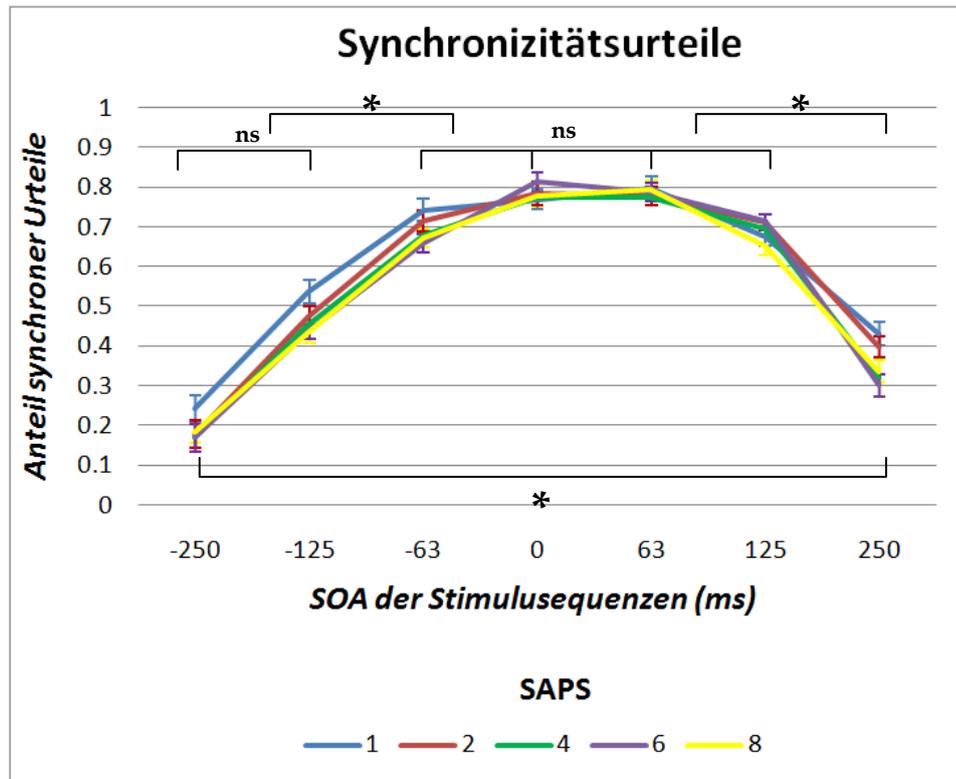
1. Beeinflusst die SAPS einfacher (a)synchroner Stimuli die JNDs? Falls ja, soll die SAPS für das Folgeexperiment gewählt werden, bei der JNDs gering sind.
2. Beeinflusst die SAPS einfacher (a)synchroner Stimuli die PsSS? Falls ja, soll SAPS als erklärender Faktor in Folgeexperimente einbezogen werden.

### 3.1.2.3 Ergebnisse des Vorexperiments 1

### 3.1.2.4 Synchronizitätsurteile

Die Anteile synchroner Urteile wurden anhand einer 7 x 5-faktoriellen Varianzanalyse mit Messwiederholung mit den Faktoren *Zeitverschiebung der Stimuli* und *Ereignisanzahl* (SAPS) analysiert (SPSS 15.0). Antworten, die während der Darbietung bis 200 ms nach Ende einer Sequenz gegeben wurde, flossen nicht in die Analyse ein. Probanden beurteilten einen Stimulus durchschnittlich in 19% (6%), 47% (4%), 70% (4%), 79% (5%), 79% (4%), 69% (4%) und 36% (5%) aller Fälle in den jeweiligen Verschiebungsbedingungen von -250 ms, -125 ms, -63 ms, 0 ms, 63 ms, 125 ms und 250 ms als synchron (Standardfehler in Klammern; s. Abb. 11). Bei den Ereignislängen eins, zwei, vier, sechs und acht wurden 60% (3%), 58% (2%), 56% (3%), 56% (2%) und 55% (2%) Synchronizitätsurteile getroffen. Die synchronen Urteile unterschieden sich auf dem Faktor *Zeitverschiebung der Stimuli* [ $F(6;90) = 25,7$ ;  $p < 0,001$ ;  $\eta^2 = 0,63$ ] und hinsichtlich der *Ereignisanzahl* [ $F(1,6;23,4) = 2,4$ ;  $p < 0,05$ ;  $\eta^2 = 0,14$ ]. In bezug auf die Interaktion aus beiden Faktoren [ $F(18,8;282) = 1,5$ ;  $p = 0,089$ ;  $\eta^2 = 0,09$ ] sind lediglich statistische Trends in Richtung Unterschiedlichkeit zu erkennen. Freiheitsgrade wurden korrigiert, falls keine Sphärizität nach Mauchly-Test vorlag. Nach Bonferroni korrigierte Post-Hoc T-Tests zeigten, dass signifikant mehr Synchronizitätsurteile für die mittleren SOA-Stufen (63 ms, 0 ms, 63 ms und 125 ms) abgegeben wurden als für die äußeren (-250 ms, -125 ms und 250 ms; alle  $p < 0,01$ ), und es signifikant weniger synchrone Urteile bei Stimuli gab, die um 250 ms (Ton zuerst) verschoben waren als bei solchen, die um -250 ms (Licht zuerst) verschoben waren ( $p < 0,05$ ). Sequenzen, die um 250 ms (Licht zuerst) und -125 ms (Ton zuerst) verschoben waren, wurden gleichhäufig als synchron bewertet ( $p = 0,85$ ). D.h. es gab mehr synchrone Urteile in den Bedingungen von -63 ms (Ton zuerst) bis 125 ms (Licht zuerst) als in den Bedingungen -250 ms (Ton zuerst), -125 ms (Ton zuerst) und 250 ms (Licht zuerst) und weniger synchrone Antworten bei 250 ms Verschiebung als bei -250 ms. Die Anteile synchroner Urteile waren bei -250 ms Verzögerungen vergleichbar hoch wie bei -125 ms. Das weist darauf hin, dass alle Verteilungen synchroner Antworten leicht asymmetrisch linksschief verteilt sind. Ein einfacher T-Test zeigte, dass sich Stimuli mit vier Ereignissen hinsichtlich ihrer Anteile synchroner Antworten von Stimuli der Ereignislänge eins ( $p < 0,05$ ) unterscheiden. Stimuli mit einem Ereignis wurden demzufolge häufiger mit synchron beurteilt als Stimuli mit vier. T-Tests, die hinsichtlich

der Unterschiede zwischen Stimuli anderer Ereignislängen (zwei, sechs, acht) durchgeführt wurden, waren nicht signifikant.



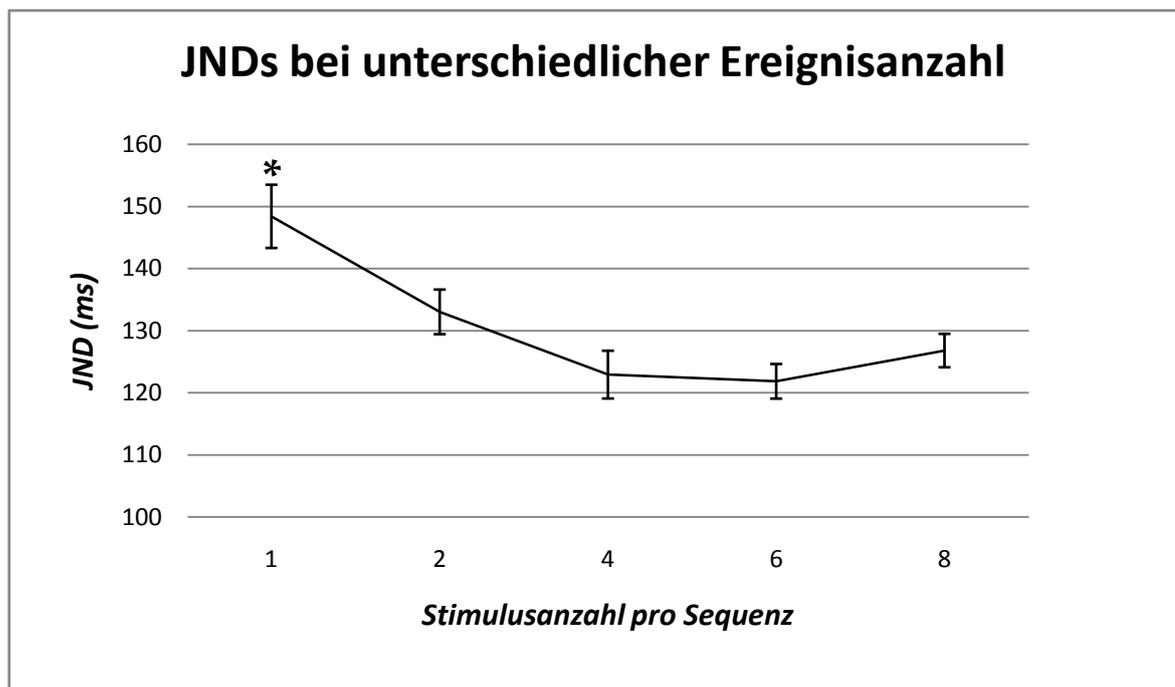
**Abbildung 11.** Anteil synchroner Antworten in Abhängigkeit des präsentierten SOAs und der Stimulusanzahl pro Sequenz (SAPS) in Vorexperiment 1. Signifikante Unterschiede der Urteile, die mittels einer 7 x 5 – stufigen Varianzanalyse ermittelt wurden, sind durch \* gekennzeichnet ( $p < 0,05$ ). S. Text zur Erläuterung. Die Fehlerbalken stellen Standardfehler dar. Dargestellte Unterschiede beziehen sich auf Datenpunkte, die über alle SAPS kollabiert wurden (s. Text). ns = nicht signifikant. Antworten unterschieden sich hinsichtlich der Faktoren SAPS, SOA und der Interaktion aus beiden Faktoren (s. Text). Post-Hoc T-Tests zeigten, dass Synchronizitätsurteile über alle SAPS kollabiert gleich häufig bei der Beurteilung von Stimuli vorkamen, die um -63 ms, 0 ms, 63, und 125 ms gegeneinander verschoben waren. Alle restlichen SOAs (-250 ms, -125 ms, und 250 ms) wurden seltener als synchron beurteilt.

### 3.1.2.5 JNDs und PsSS

Aus den Anteilen synchroner Urteile der sieben SOAs wurden je Proband und SAPS Kurvenfunktionen ermittelt, aus denen PSS und JND abgeleitet wurden (s. 2.4.1). Aus den geschätzten Werten für PsSS und JNDs wurden standardisierte Z-Werte berechnet, und alle Probanden, deren Datenpunkte zwei Standardabweichungen über- oder unterschritten eliminiert. Dies verhindert, dass unzuverlässige Schätzungen aufgrund willkürlichen Antwortverhaltens in die Datenanalyse einfließen (s. Spence et al.,

2001; Vatakis et al., 2007). PsSS und JNDs wurden jeweils anhand fünfstufiger Varianzanalysen (für den Faktor *Ereignisanzahl*) mit Messwiederholung ausgewertet, wobei die Stufen des definierten Faktors die fünf verschiedenen Stufen der SAPS widerspiegeln.

Die Datenanalyse ergab, dass der Faktor *Ereignisanzahl* 40,4% Varianz in den JND-Daten erklären kann [ $F(4;48) = 8,1$ ;  $p < 0,001$ ;  $\eta^2 = 0,40$ ]. Post-hoc-Tests (korrigiert nach Bonferroni) ergaben, dass JNDs bei Stimuli mit einem Ereignis größer waren als bei Stimuli mit den Ereignislängen vier ( $p < 0,01$ ), sechs ( $p < 0,05$ ) und acht ( $p < 0,05$ ). Stimuli mit zwei Ereignissen unterschieden sich weder von Stimuli mit einem, noch von Stimuli anderer Ereignislängen signifikant. Jedoch ergab ein einfacher T-Test, dass JNDs bei einer Ereignislänge von zwei im statistischen Trend kleiner sind als bei einer Ereignislänge von eins und größer als bei Stimuli mit vier Ereignissen ( $p < 0,1$ ).



**Abbildung 12.** JNDs in Abhängigkeit verwendeter SAPS bei Vorexperiment 1. JNDs bei einem Stimulus pro Sequenz waren signifikant größer (\*) als in allen anderen Experimentalbedingungen. Fehlerbalken stellen Standardfehler dar.

Im Durchschnitt betragen die JNDs für die Ereignisanzahl eins 148 ms (10,2 ms), 133 ms (7,2 ms) für zwei, 123 ms (7,7 ms) für vier, 122 ms (5,6 ms) für sechs und 127 ms (5,4 ms) für acht Ereignisse (Standardfehler in Klammern; s. Abb. 12).

Eine Varianzanalyse der PSS-Daten mit dem Faktor Ereignislänge ergab keine signifikanten Effekte [ $F(4;48) = 0,84$ ;  $p = 0,5$ ;  $\eta^2 = 0,07$ ]. Im Mittel mussten Lichter den Tönen um 22 ms (5 ms), 29 ms (6 ms), 26 ms (5 ms), 31 ms (4 ms) und 29 ms (5 ms) bei ein, zwei, vier, sechs und acht Stimulusereignissen vorangehen, um von den Probanden als synchron bewertet zu werden (durchschnittliche PsSS).

### 3.1.2.6 Bewertung der Ergebnisse des Vorexperiments 1

Die Schätzungen für JNDs und PSS stimmen mit Ergebnissen anderer Studien überein, die per Synchronizitätsurteil (asynchron vs. synchron) (a)synchrone einfache audiovisuelle Paare bewerten ließen und die Datenanalyse in vergleichbarer Weise durchführten (Fujisaki et al., 2004; Spence et al., 2001; Stone et al., 2001; Vatakis & Spence, 2006; Zampini et al., 2005; JNDs von 114 ms – 192 ms). Schätzwerte für PsSS fielen in diesen Studien in den Bereich 1 ms bis 60 ms. Generell ist hierzu anzumerken, dass multisensorische Integration durch zunehmende zeitliche Synchronizität zweier sensorischer Ereignisse begünstigt wird (z. B. Calvert & Lewis, 2004; De Gelder and Bertelson, 2003), genaue physikalische Gleichzeitigkeit jedoch keine zwingende Voraussetzung für das menschliche Wahrnehmungssystem ist, um eine einheitliche perzeptuelle Repräsentation eines multisensorischen Ereignisses zu erzeugen. Vielmehr scheint eine Verzögerung des Tonreizes gegenüber dem Lichtreiz von ca. 30 ms für ein synchrones multisensorisches Perzept optimal zu sein (z. B. King & Palmer, 1985; Stone et al. 2001; Vatakis et al. 2007). King & Palmer untersuchten die SC mittels Einzelzelleableitungen an Meerschweinchen. Wie eingangs erwähnt ist diese Struktur maßgeblich an multisensorischer Integration beteiligt (z. B. Stein & Meredith, 1993). Sie fanden heraus, dass die stärksten neuronalen Interaktionen auftraten, wenn bei audiovisuellen Stimuli der Lichtreiz 20 – 80 ms vor dem Tonreiz präsentiert wurde (s. auch Kayser et al., 2008). Die vorliegenden PsSS liegen exakt in diesem Bereich und lassen vermuten, dass diese zeitliche Differenz zwischen Licht und Ton bei Menschen optimal für eine synchrone Wahrnehmung beider Ereignisse ist (s. auch Stone et al., 2001).

Gemäß der statistischen Auswertung wirkt sich die Ereignislänge auf die Sensibilität der Probanden gegenüber Asynchronizität aus, nicht aber auf deren Punkt subjektiver Gleichzeitigkeit. Die Fähigkeit, synchrone von asynchronen Stimuli zu diskriminieren ist bei Stimuli mit einem Ereignis geringer als bei Stimuli mit mehr

Ereignissen, wobei die Diskriminationsfähigkeit bei Stimuli mit zwei Ereignissen tendenziell geringer ist als bei Stimuli der Ereignislänge vier, sechs und acht. Die Diskriminationsfähigkeit verändert sich hingegen nicht, wenn Probanden Stimuli beurteilen, die vier, sechs oder acht Ereignisse beinhalten. Mit steigender Ereignislänge steigt sie demzufolge bis zu vier Ereignissen an und hält dann ein Niveau. Probanden können folglich Stimuli mit mindestens vier Ereignissen entscheidend einfacher hinsichtlich ihrer (A)Synchronizität beurteilen als Stimuli, die weniger Ereignisse beinhalten, sie können jedoch keinen Leistungsgewinn verzeichnen, wenn Stimuli mit mehr als vier Ereignissen dargeboten werden. Vermutlich reflektiert diese Datenlage eine Unsicherheit des audiovisuellen Zeitsystems, wenn ein sehr kurzes audiovisuelles Ereignis zeitlich eingeordnet werden soll: Da ein kurzer Reiz dem perzeptuellen System eine geringe Zeitspanne lässt, die Zeitlichkeitseinschätzung zu überprüfen, vergrößert sich das audiovisuelle Zeitfenster, innerhalb dessen Ereignisse zweier Modalitäten immer noch als synchron wahrgenommen werden können. Das System benötigt folglich einen größeren Abstand zwischen beiden Ereignissen, um ein eindeutiges Perzept zu erzeugen.

Anhand der oben aufgeführten Befunde sollen für die folgenden Untersuchungen Stimuli mit vier Ereignissen verwendet werden. Ausgehend von der Datenlage des vorliegenden Experiments kann die initial aufgeführte Fragestellung 2 wie folgt beantwortet werden:

1. Die Sensitivität der Probanden gegenüber audiovisueller (A)Synchronizität bei einfachen Stimuli nimmt bis zu einer Ereignislänge von vier zu. Daher soll diese Ereignislänge für die Folgeexperimente verwendet werden.

### **3.1.3 Hauptexperiment**

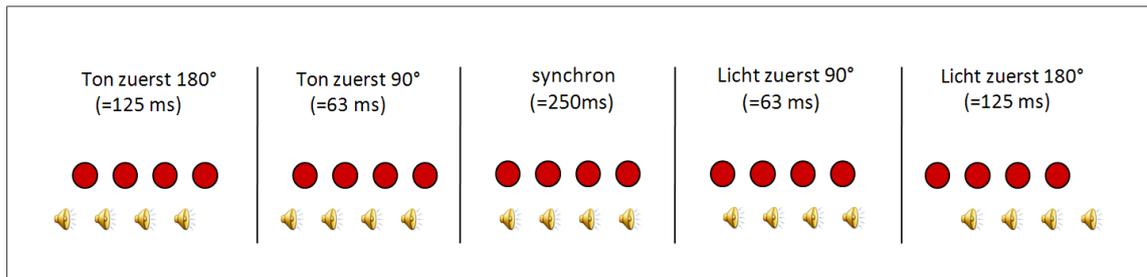
#### **3.1.3.1 Methodik**

Zweiundzwanzig rechtshändige (Selbstbericht) Probanden (Durchschnittsalter: 26, Spanne: 23-33, 11 weiblich) wurden in einen 3 Tesla Trio-Scanner (Siemens, Erlangen) gelegt, der mit einem piezoelektrischen Lautsprecher und Lichtleiterkabeln, die das Signal weißer und roter LEDs weiterleiteten, ausgestattet wurde. Alle Teilnehmer berichteten, normales Hörvermögen und normales oder korrigiertes Sehvermögen und keine

psychiatrischen oder neurologischen Störungen zu haben. Sie gaben schriftlich ihr Einverständnis, an der Untersuchung teilzunehmen und wurden über deren Zweck aufgeklärt. Das Experiment dauerte ca. 80 Minuten für funktionelle Messungen plus der Zeit, die für die Aufnahme struktureller Bilder benötigte wurde (ca. 15 min). Die visuellen Stimuli wurden per Lichtleiterkabel im Inneren der MR-Röhre dargeboten; ein weißer Fixationslichtpunkt wurde auf Augenhöhe angebracht und eine rotes Ziellicht 5° Sehwinkel darüber (60 cd/m<sup>2</sup>). Jedes auditorische Ereignis war ein deutlich hörbarer Ton, der durch den Lautsprecher direkt unter dem Fixationslicht bei 500 Hz und 65 dB (SPL, vom Ohr der Probanden aus gemessen) präsentiert wurde.

Die Reize bestanden aus vier aneinandergereihten akustischen und optischen Signalen, die miteinander kombiniert, mit einer Frequenz von 4 Hz präsentiert wurden und jeweils 30 ms dauerten (s. Abb. 13). D.h. der zeitliche Abstand zwischen einem audiovisuellen Ereignis und dem darauf folgenden innerhalb einer Sequenz betrug 250 ms, wenn es synchron dargeboten wurde. Die audiovisuellen Sequenzen wurden entweder synchron oder asynchron dargeboten. Bei nicht synchronen Sequenzen wurden auditorische gegen visuelle Ereignisse phasisch um 90° und 180° verschoben (Lichtblitze gehen auditorischen Signalen voran bzw. Lichtblitze folgten akustischen Signalen), so dass fünf verschiedene Stimulustypen bzw. Experimentalbedingungen existierten: ein(e) synchrone(r) und vier asynchrone. Für asynchrone Stimuli war der zeitliche Abstand zwischen Lichtblitzen und akustischen Signalen (oder umgekehrt) 63 ms bei 90° phasischer Verschiebung und 125 ms bei 180°. Die kurzen auditorischen und visuellen Ereignisse wurden so dargeboten, dass selbst bei der geringsten Verschiebungsbedingung (63 ms) keine Überlappung zwischen ihnen auftrat. Die Verzögerungsbedingungen der Stimuli wurden in randomisierter Reihenfolge präsentiert.

Die Probanden wurden instruiert, eine Taste auf einem MR-tauglichen Taster zu drücken, um einen subjektiv als synchron empfundenen Reiz anzuzeigen und eine zweite Taste, um einen subjektiv als asynchron empfundenen Reiz anzuzeigen, nachdem eine Sequenz gezeigt wurde. Während eines Übungsdurchgangs vor der Bildakquise wurden alle Bedingungen des nachfolgenden Experiments präsentiert. Die Probanden wurden dahingehend trainiert, ganze Sequenzen (aus vier audiovisuellen Ereignisse bestehend) zu beurteilen und nicht einzelne Ereignisse innerhalb von Sequenzen.



**Abbildung 13.** Stylisierte Darstellung der Stimulussequenzen, die für Experiment 1 genutzt wurden. Eine Sequenz bestand aus vier audiovisuellen Ereignissen (s. auch Abb. 8 und Kapitel 1.2.1 für nähere Spezifikationen). Lichtblitze und Tonsignale konnten in fünf Experimentalbedingungen um 125 ms und 63 ms in beide Richtungen phasisch verschoben oder phasengleich (0 ms) sein. Einzelne Ton- bzw. Lichtsignale wurden für die Dauer von 30 ms präsentiert.

Eine Pilotstudie zeigte, dass 20 Prozent aller synchronen Sequenzen nicht als synchron bewertet wurden. Um in der vorliegenden fMRT-Studie mindestens 40 Durchgänge zu erzielen, in denen Synchronizität nicht als synchron empfunden wird, wurde eine Mindestanzahl von 180 Durchgängen je Bedingung ermittelt. Diese Mindestanzahl ist notwendig, um robuste Modellparameter aus den fMRT-Daten schätzen zu können. Daher wurde ein Design mit sechs Durchläufen, die jeweils elf Minuten dauern und 150 Durchgänge beinhalten, gewählt. Die Stimulusabfolge wurde mit MATLAB 7.5 randomisiert und die Anfangszeiten jedes Durchgangs innerhalb eines Zeitfensters von 500 ms unterschiedlich verzögert oder vorgezogen (Jittern). Um eine zu starke Überlappung der BOLD-Antworten zu verhindern, wurden mittels eines Effizienz-Schätzalgorithmus (Hinrichs et al., 2000) sogenannte ‚Null-Ereignisse‘ eingefügt. Das Interstimulusintervall war demzufolge poisson-verteilt (Range: 3000 – 9000 ms) und betrug durchschnittlich 4,5 Sek. Die Stimuli wurden mit Presentation 9.11 (Neurobehavioral Systems Inc., Kalifornien, USA) dargeboten.

Die fMRT-Daten wurden mit einem 3 Tesla-Ganzkörperscanner von Siemens Trio (Erlangen) und einer 8-Kanal-Kopfspule (188 Volumen, die den ganzen Kopf abdeckten, TE 30 ms, flip Winkel 80°, Auflösung: 64 x 64 x 32 bei 3,5 x 3,5 x 4 mm) gemessen. Für die anatomische Struktur wurde ein T<sub>1</sub>-gewichtetes Volumen (räumliche Auflösung: 1 x 1 x 1 mm, 157 x 189 x 136 Schichten) aufgenommen (TR: 2500 ms, TE: 4,77 ms, flip angle: 7°). Für die funktionellen Messungen wurde eine T<sub>2</sub>\*-gewichtete Echoplanar-Sequenz (echo planar imaging, EPI) verwendet (räumliche Auflösung: 3,5 x 3,5 x 4 mm, 64 x 64 x 32 Schichten, TR: 2000 ms plus 2000 ms Pause, TE: 30 ms, flip angle: 80°, aufsteigend

aufgenommen). Zusätzlich wurde ein  $T_1$ -gewichtetes inversion recovery-EPI aufgenommen, das dieselben Parameter und damit räumlichen Verzerrungen wie die funktionellen EPI-Messungen aufwies (räumliche Auflösung:  $3,5 \times 3,5 \times 4$  mm,  $64 \times 64 \times 32$  Schichten, TR: 2500 ms, TE: 17 ms, TI: 1200 ms, flip angle:  $80^\circ$ ). Dieses IR-EPI wurde zur Bestimmung der Normalisierungsparameter verwendet. Die Stimuli wurden während Ruhephasen des Scanners (2 Sek.) gezeigt, die sich mit Aufnahmephasen abwechselten (2 Sek. Aufnahmezeit, was zu einer TR von 4 Sek. führt), um zu verhindern, dass die Scannergeräusche mit den auditorischen Stimuli interferieren („rapid sparse-sampling“, s. z. B. Noesselt et al., 2007). Die ersten fünf Volumen wurden aus der Analyse ausgeschlossen, um eine vergleichbare Saturierung für alle Volumen zu erreichen.

Augenbewegungen wurden anhand eines MRT-kompatiblen Aufnahmesystems (Kanowski et al., 2007) und einer Software (PupilTracker, HumanScan, Erlangen) überwacht. Die Probanden trugen Ohrstöpsel; die wahrgenommene Lautstärke und Balance wurden individuell angepasst, um sicher zu stellen, dass die Sequenzen klar verstanden werden konnten.

Vor dem Hauptexperiment wurde ein funktioneller Lokalisationsdurchgang durchgeführt, während dessen nur unimodal auditorische oder unimodal visuelle Stimuli in 4 Hz-Sequenzen gezeigt wurden (188 Volumen, TR 4 Sek, TE 30 ms, Flip-Winkel  $80^\circ$ , Auflösung:  $64 \times 64 \times 32$  bei  $3,5 \times 3,5 \times 4$  mm, aufsteigend aufgenommen). Die Sequenzen wurden jeweils in den Aufnahmepausen mit einer Länge von 1000 ms 20 Sekunden lang präsentiert (ISI: vier Sekunden). Auditorische und visuelle Blöcke wurden je drei Mal dargeboten. Die resultierenden visuellen, auditorischen und audio-visuell überlappenden Aktivierungskarten wurden genutzt, um spezifisch unisensorische und potentiell multisensorische Gebiete zu identifizieren.

Nachdem die aufgenommenen Bilder der Lokalisierungsdurchgänge vorverarbeitet wurden (s. 2.2.2.1) wurde ein Modell mit vier Box-Car-Funktionen erstellt (Blockdesign), das mit den hämodynamischen Antwortfunktionen für auditorische und visuelle Durchgänge und den jeweiligen stimulationsfreien Zeiträumen (Baseline) konvolviert wurde (s. auch Beauchamp, 2005; Noesselt et al., 2007; Szycik et al., 2008a; 2008b). Zeit- und Verteilungsableitungen, sowie Bewegungsparameter wurden in das Modell integriert. In einem Modell zweiter Ebene wurden die Daten anhand einer

Varianzanalyse (SPM5) ausgewertet. Auditorische und visuelle Masken wurden mit überlappenden, audiovisuellen Aktivierungen exklusive maskiert.

Die Bedingungen des Hauptexperiments wurden mit hämodynamischen Antwortfunktionen modelliert (ereigniskorreliertes Design, Vorverarbeitung: s. 2.2.2.1) und zusätzlich Zeit- und Verteilungsableitungen geschätzt. Anschließend wurden die Gruppenergebnisse des Hauptexperiments mit den Gruppenergebnissen des funktionellen Lokalisationsdurchganges maskiert. Eine audiovisuelle Maske (Überlappung auditorischer und visueller Aktivierungen) wurde zu diesem Zweck berechnet (Beauchamp, 2005). Für die Analyse des Lokalisationsdurchganges wurden auditorische bzw. visuelle Blöcke mit ihren Baselines verglichen, während der keine Stimuli gezeigt wurden ( $p < 0,001$ ;  $k > 10$ ). Die resultierenden Aktivierungskarten wurden mit  $6 \times 6 \times 6 \text{ mm}^3$  geglättet und als Masken für die Gruppenstatistik des Hauptexperiments verwendet. Für die Analyse der experimentellen Durchgänge wurden zehn Bedingungen definiert: zwei Antwortmöglichkeiten (synchron, asynchron) pro Bedingung [fünf SOAs bei -125 ms (Ton voran); -63 ms (Ton voran), synchron, 63 ms (Licht voran), 125 ms (Licht voran)]. Um Effekte für die Bedingungen statistisch zu testen, wurden lineare Kontraste für jeden Probanden erstellt (s. unten). Die resultierenden Kontrastbilder wurden für eine Analyse auf zweiter Ebene genutzt (Varianzanalyse). Die Schwellen für die parametrischen Karten der T-Statistiken für jeden Voxel wurden bei  $p < 0,05$  und  $k > 30$  Voxel gesetzt (s. Beauchamp, 2005). Dazu wurden zuerst unimodal auditorische und visuelle und potentiell multisensorische Gebiete identifiziert (s. oben), und anschließend Statistiken anhand linearer Kontraste innerhalb dieser Masken berechnet. Dies stellt einen statistisch konservativen Ansatz dar, da innerhalb bereits durch die verwendeten Stimuli modulierten Areale nach signifikanten Aktivierungen gesucht wird, die zusätzlich durch experimentelle Kontraste moduliert werden. Dies lässt vermuten, dass Voxel, die innerhalb der zuvor identifizierten Gebiete bereits verhältnismäßig geringe T-, bzw. p-Werte aufweisen, bedeutsame Aktivierungen darstellen. Derartige Aktivierungen sollen im Folgenden interpretiert werden, da durch die zuvor erstellten Masken sichergestellt wurde, dass die verwendeten Stimuli die entsprechenden Regionen modulieren.

Es wurden Kontraste für physikalische und perzeptuelle Synchronizität (Phasengleichheit) und Asynchronizität (Phasenungleichheit) erstellt: Physikalische

Synchronizität (die bloße Präsentation synchroner Stimuli) verglich das Mittel beider Antwortmöglichkeiten während der synchronen Bedingung mit dem Mittel der Antwortmöglichkeiten in allen anderen Bedingungen. Für physikalische Asynchronizität wurde dieser Vergleich umgekehrt. Um Areale zu identifizieren, deren Modulationen mit den psychophysischen Größen synchroner Wahrnehmung einhergehen, wurde die Anzahl synchroner Antworten für jede der fünf experimentellen Bedingungen vom Mittelwert aller synchronen Verhaltensantworten in allen Bedingungen subtrahiert und die resultierenden Werte als Kontrastwerte für die fünf modellierten Antwortverläufe für synchrone Urteile gesetzt [SOA -125 ms (Ton voran): -16,9; SOA -63 ms (Ton voran): 1,44; SOA 0 ms (synchron): 11,9; SOA 125 ms (Licht voran): 9,5; SOA 125 ms (Licht voran): -6,03]. Die resultierenden Aktivierungskarten würden demzufolge Modulationen in Arealen aufzeigen, deren neuronales Antwortmuster das mittlere Antwortmuster für synchrone Antworten in den Verhaltensdaten widerspiegelt. Um Areale für das Antwortverhalten asynchroner Wahrnehmung zu identifizieren, wurde analog dazu die Anzahl asynchroner Antworten für jede der fünf experimentellen Bedingungen vom Mittelwert aller asynchronen Antworten in allen Bedingungen subtrahiert, und die resultierenden Werte als Kontrastwerte für die fünf modellierten Antwortverläufe für asynchrone Beurteilungen gesetzt [SOA -125 ms (Ton voran): 16,9; SOA -63 ms (Ton voran): -1,44; SOA 0 ms (synchron): -11,9; SOA 125 ms (Licht voran): -9,5; SOA 125 ms (Licht voran): 6,03]. Darüber hinaus wurden Kontraste für synchrone und asynchrone Wahrnehmung aller Stimuli erstellt, die für jede Stimulusbedingung synchrone mit asynchronen Antworten (,synchrone > asynchrone Antworten') verglichen bzw. umgekehrt (,asynchrone > synchrone Antworten'). Dazu wurden für die jeweils beiden Versuchsbedingungen, unter denen Licht- und Tonreize um 125 ms und 63 ms (jeweils Licht voran und Ton voran) gegeneinander verschoben waren, asynchrone gegen synchrone Antworten verglichen (,asynchrone > synchrone Antworten') und umgekehrt (,synchrone > asynchrone Antworten'), um unter den oben aufgeführten Kontrasten für synchrone und asynchrone Perzeption aller Stimuli nach Stimulustyp zu differenzieren. Auf diese Weise kann exploriert werden, welche Modulationen synchroner und asynchroner Wahrnehmung auf Stimuli zurückgehen, bei denen der Ton dem Licht voranging oder umgekehrt. Für synchrone Stimuli wurden synchrone gegen asynchrone

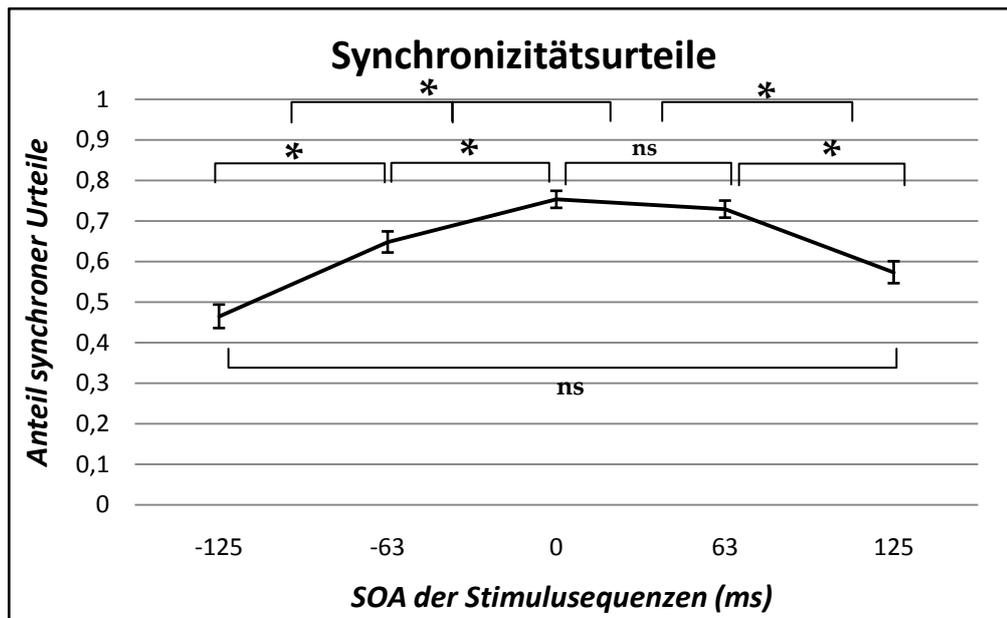
Antworten verglichen (,synchrone > asynchrone Antworten') und umgekehrt (,asynchrone > synchrone Antworten').

### 3.1.3.2 Ergebnisse

#### 3.1.3.2.1 Verhaltensdaten

Aufgrund von zu intensiven Augenbewegungen während der Stimuluspräsentation oder Ermüdungserscheinungen wurden sechs Probanden aus der Analyse sowohl der Verhaltens- als auch der fMRT-Daten ausgeschlossen. Im Durchschnitt wurden Stimuli, die um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) und 125 ms (Licht voran) gegeneinander verschoben waren in 47% (6%), 65% (5%), 75% (4%), 73% (4%) und 57% (5%) aller Fälle als synchron empfunden (s. Abb. 14). Eine einfaktorielle Varianzanalyse mit Messwiederholung mit den fünf Stufen der Stimulusbedingungen ergab signifikante Effekte für die Synchronizitätsurteile [ $F(2;30,4) = 16,1$ ;  $p < 0,001$ ;  $\eta^2 = 0,52$ ]. Die Post-hoc Analyse (nach Bonferroni korrigiert) zeigte, dass Stimuli der Stufe -125 ms (Ton voran) genauso oft als synchron bewertet wurden wie Stimuli der Stufe 125 ms (Licht voran;  $p = 0,74$ ). Dasselbe gilt für Stimuli der Stufen -63 ms (Ton voran) und 63 ms bzw. 125 ms (Licht voran;  $p = 0,14$  bzw.  $p = 0,48$ ) und für Stimuli der Stufe 0 ms (synchron) und 63 ms (Licht voran;  $p = 1$ ). Synchrone Stimuli wurden signifikant häufiger als synchron bewertet als asynchrone Stimuli mit Ausnahme der Stimulusbedingung 63 ms (Licht voran, alle  $p < 0,01$ ). Stimulusbedingung 63 ms (Licht voran) bekam signifikant mehr Synchronizitätsurteile als Stimuli, die um 125 ms verzögert waren (beide  $p < 0,01$ ), und Stimulusbedingung -63 ms (Ton voran) wurde signifikant häufiger als synchron empfunden als Stimulusbedingung -125 ms (Ton voran;  $p < 0,01$ ) und signifikant weniger als synchrone Stimuli ( $p < 0,01$ ). Diese Verteilung ist den Werten zufolge linksschief.

Aus den Synchronizitätsurteilen wurden JND und PSS abgeleitet: Die mittlere JND betrug 172 ms (9 ms), der mittlere PSS 10 ms (17 ms; Standardfehler in Klammern).



**Abbildung 14.** Anteil synchroner Antworten je SOA für fMRT-Experiment 1 über alle Teilnehmer und Durchgänge. Stimuli, die um 0 ms oder 63 ms (Licht voran) verschoben waren, wurden signifikant am häufigsten als synchron beurteilt. Fehlerbalken stellen Standardfehler dar. \* markiert signifikante Unterschiede, ‚ns‘ nicht signifikante Unterschiede.

### 3.1.3.3 fMRT-Ergebnisse

Zunächst werden die Ergebnisse des Lokalisationsdurchgangs besprochen und Areale aufgeführt, die durch unisensorische auditorische oder visuelle Stimuli moduliert wurden und solche, die durch beide Stimulustypen moduliert wurden. Daraufhin werden diejenigen Areale berichtet, die innerhalb dieser Gebiete durch experimentelle Kontraste (s. 3.1.3.1) angesprochen wurden.

#### 3.1.3.3.1.1 Lokalisationsdurchgang

Durch unimodale, rein auditorische Stimuli wurden temporale und präfrontale Areale beider Hemisphären moduliert (s. Anhang, Tab. 7 b), darunter die Heschl'schen Gyri (HG), der Gyrus temporalis superior (STG) und die Insula ( $p < 0,05$ ;  $k > 100$ ). Rein visuelle Aktivierungen wurden in okzipitalen, temporalen, parietalen und frontalen Arealen beider Hemisphären gefunden, u.a. im Gyrus okzipitalis inferior (IOG), im Gyrus fusiformis (FG), im Gyrus lingualis (LG), im Gyrus temporalis inferior (ITG), im Sulcus intraparietalis (IPS), im Gyrus frontalis medius (MFG), im Gyrus angularis und im Gyrus praecentralis ( $p < 0,05$ ;  $k > 100$ ; s. Anhang, Tab. 7 a). Bimodale Aktivierungen (Modulation durch auditorische und visuelle Stimuli) wurden in temporalen, (prä)frontalen und

parietalen Regionen beider Hemisphären gefunden. Darunter der STS, der STG, die Insula, der IPS und der Gyrus frontalis inferior (IFG;  $p < 0,05$ ;  $k > 100$ ; s. Anhang, Tab. 7 c). Kontraste des Hauptexperiments wurden innerhalb dieser unisensorischen und potentiell bimodalen Gebiete dargestellt.

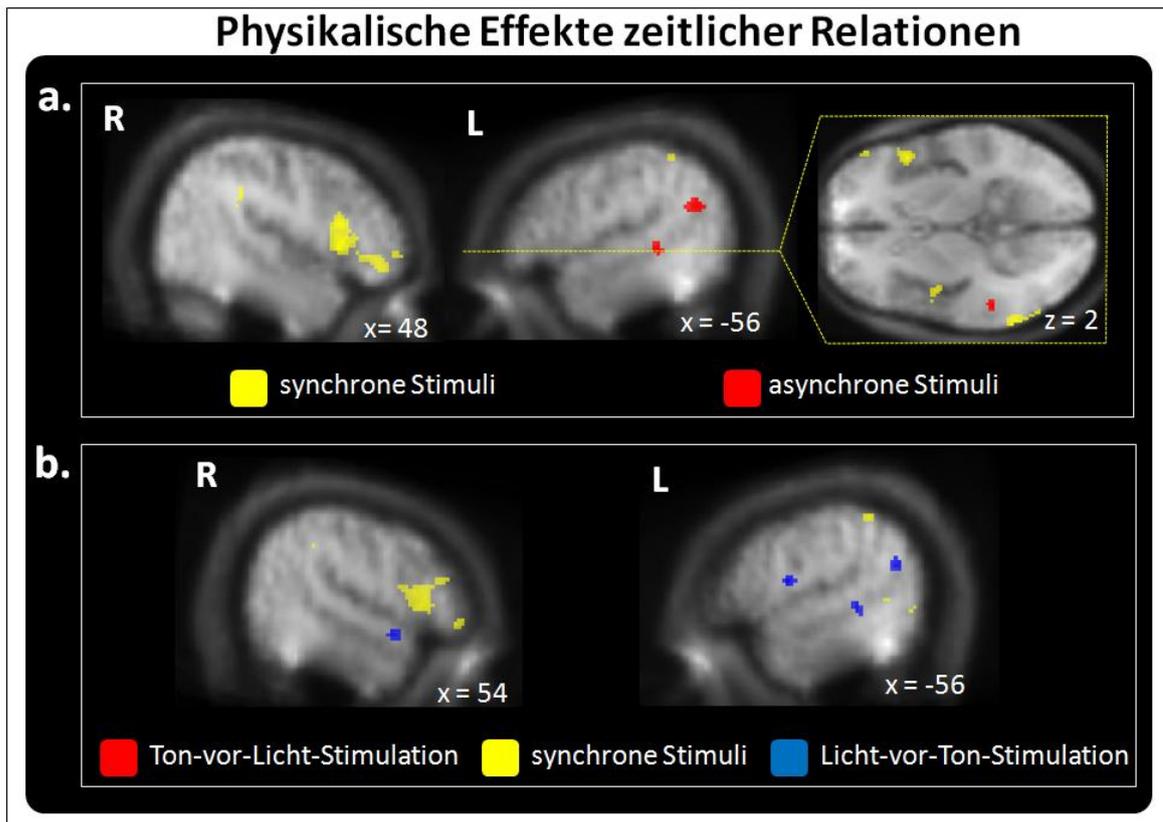
### 3.1.3.3.1.2 Experimentelle Kontraste

#### 3.1.3.3.1.2.1 Physikalische (Un-)Gleichzeitigkeit

Gebiete, die durch bloße Präsentation (physikalischer Effekt) synchroner Stimuli innerhalb bimodaler Areale mehr aktiviert wurden als durch Präsentation asynchroner (s. Anhang, Tab. 8 b und Abb. 15 a) Stimuli waren linkstemporal der MTG und STG, der posteriore Sulcus temporalis superior (pSTS), frontal die Insula und der MFG, linksseitig und rechtsseitig der IFG und parietal der rechte Gyrus supramarginalis (SMG). Innerhalb unisensorischer auditorischer Areale wurde der linke MFG durch diesen Kontrast moduliert. Visuelle Areale wurden nicht moduliert. Synchrone Stimuluspräsentation ging demzufolge mit vermehrter linkshemisphärischer Aktivität in temporalen, parietalen und prämotorisch frontalen Gebieten sowie mit beidseitiger Aktivität im dorso- und ventrolateralen Präfrontalkortex (DLPFC, VLPFC) einher (s. Anhang, Tab. 8 a). Dies geschah vor allem unter der Beteiligung bimodaler Areale und eines auditorischen.

Die bloße Präsentation aller gezeigten asynchronen Stimuli ergab in bimodalen Gebieten höhere Aktivität im Vergleich zu synchronen Stimuli im mittleren Sulcus temporalis superior (mSTS) und im pSTS linksseitig (s. Anhang, Tab. 8 b und Abb. 15 a). In unisensorisch lokalisierten Gebieten wurden das rechte Operculum frontale (auditorisch) und der rechte Gyrus praecentralis (visuell) durch diesen Vergleich moduliert. Asynchrone Stimuluspräsentation ging daher mit linkshemisphärischer Aktivität entlang des bimodalen STS und mit unisensorischer Aktivität in rechten prämotorischen Arealen einher.

Um die oben aufgeführten Ergebnisse näher zu untersuchen, wurden asynchrone Stimuli, bei denen der Ton dem Licht voranging und umgekehrt jeweils einzeln mit synchronen Stimuli verglichen. Auf diese Weise wurde präzisiert, welche Stimulusarten Effekte für physikalische Asynchronizität hervorrufen.



**Abbildung 15. a.** Effekte der Stimuluspräsentation unabhängig von ihrer Wahrnehmung innerhalb bimodaler Areale. Die Darbietung synchroner Reize (gelb) im Vergleich zu asynchronen ( $\pm 125/63$  ms) modulierte den PFC, die linke Insula und den MTG (s. Tab. 8a). Gestrichelte gelbe Linien stellen die Schnittführung dar, mit der der anatomische Bezug zweier Ebenen hergestellt wurde. Die Präsentation asynchroner Reize (rot) im Vergleich zu synchronen modulierte den mittleren und posterioren STS der linken Hemisphäre. **b.** Effekte der Stimuluspräsentation sortiert nach der Verschiebungsrichtung der Reize. Es wird sichtbar, dass im Vergleich zu *a.* die Modulationen asynchroner Reize auf Stimuli zurückgehen, bei denen das Licht dem Ton voranging (125/63 ms; blau; s. Tab. 8 c, d).

### 3.1.3.3.1.2.2 Licht vor Ton vs. Ton vor Licht (physikalisch)

Wenn Licht dem Ton voran ging, wurden linksseitig innerhalb bimodaler Gebiete des pSTS und mSTS, sowie der rechte anteriore STG und das rechte Cerebellum stärker moduliert als im Vergleich zu synchroner Stimuluspräsentation (s. Anhang, Tab. 8 d und Abb. 15 b). Auditorische Gebiete wurden durch diesen Kontrast im Bereich des rechten Operculum frontale aktiviert. Es wurden jedoch keine signifikanten Effekte in visuellen Gebieten beobachtet. Die Präsentation von Reizen, bei denen der Ton dem Licht voranging, erzeugte nur innerhalb visueller Areale mehr Aktivität im rechten Gyrus praecentralis (s. Anhang, Tab. 8 c). Insgesamt deckt sich dieses Aktivitätsmuster mit dem der physikalischen (A)synchronizität und zeigt gleichzeitig, dass die Effekte

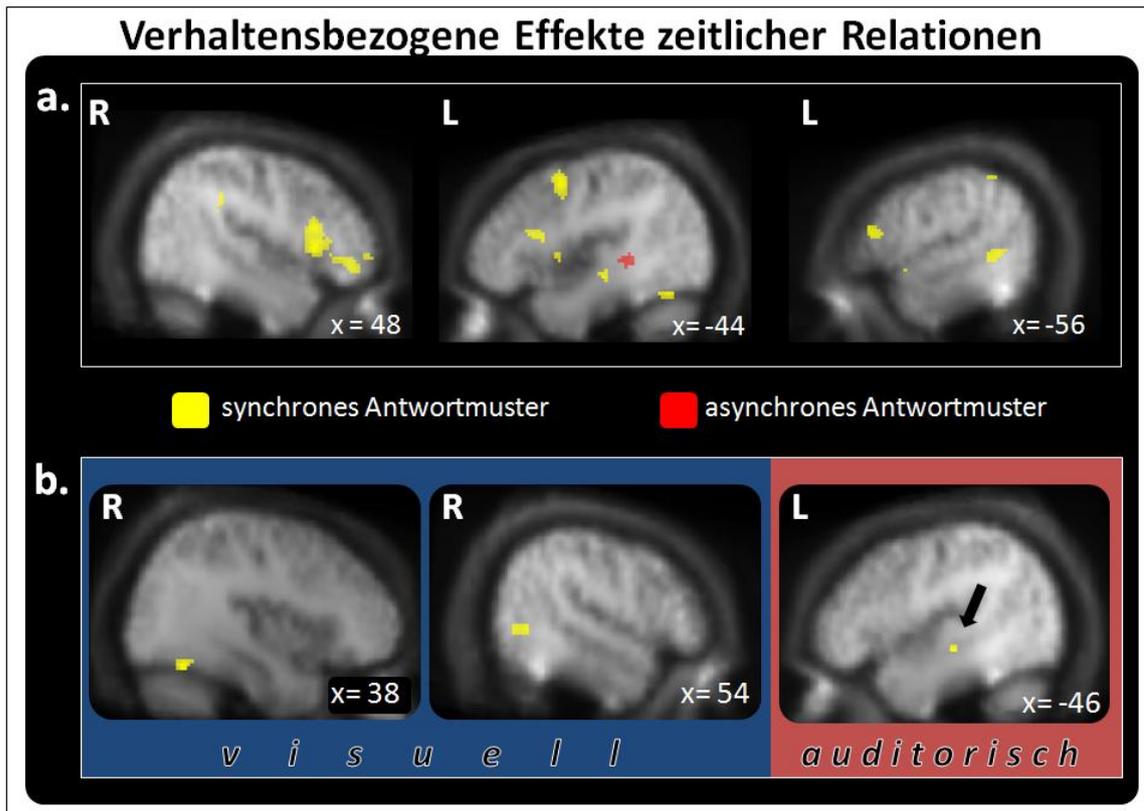
physikalischer Asynchronizität weitestgehend auf Stimuli zurückgehen, bei denen das Licht dem Ton voran geht.

#### 3.1.3.3.1.2.3 (A)synchrones Antwortverhalten

BOLD-Effekte, die das Antwortverhalten für Synchronizität nachvollziehen, wurden in bimodalen Arealen beidseits temporal in pSTS und mSTS, MTG, STG, und rechtsseitig in der Insula und dem PT gefunden. Frontal traten Effekte beidseitig in IFG und MFG und im linken Gyrus praecentralis und SFG auf (s. Anhang, Tab. 10 a und Abb. 16 a). Darüber hinaus wurden der Parietalkortex (IPS, SMG), sowie okzipitale Gebiete (FG, LG) beidseitig moduliert. Schließlich wurde das Kleinhirn rechtsseitig durch diesen Kontrast moduliert. Unisensorisch wurden Effekte beidseitig im primären auditorischen Kortex (HG, mSTS, STG, Insula) und im linken Frontal- (Gyrus praecentralis) bzw. Präfrontalkortex (MFG/IFG) gefunden. Innerhalb visueller Gebiete modulierte der Kontrast okzipitale Areale (FG, ITG) beidseits und linksseitig den IPS, Gyrus praecentralis und das linke Cerebellum (s. Abb. 16 b).

Asynchrones Antwortverhalten modulierte BOLD-Antworten linksseitig im mSTS und Gyrus praecentralis innerhalb ‚bimodaler‘ Gebiete (s. Anhang, Tab. 10 b und Abb. 16 a). Effekte innerhalb unisensorischer wurden nicht beobachtet (s. Abb. 16 b).

Das Antwortverhalten gegenüber synchronen Reizen ging demzufolge mit bilateralen Modulationen in Temporal-, Parietal-, in dorso- bzw. ventrolateralem Präfrontalkortex sowie Modulationen im prämotorischen Kortex und im Cerebellum einher, wobei sowohl Aktivierungen in unimodalen wie auch bimodalen Arealen gefunden wurden. Das Antwortverhalten, das mit asynchronen Wahrnehmungen einherging, ergab weniger Aktivierungen in ausschließlich bimodalen Gebieten des linken prämotorischen und Temporalkortex.



**Abbildung 16.** Aktivierungskarten neuronaler Modulationen, die Effekte synchroner (gelb) und asynchroner (rot) Verhaltensantworten auf alle Stimulustypen widerspiegeln. **a.** synchrones Antwortverhalten modulierte ein kortikales Netzwerk aus lateralem Temporalkortex, PFC und Parietalkortex innerhalb bimodaler Areale, während asynchrones Antwortverhalten ausschließlich den linken STS modulierte (s. Tab. 10 a). **b.** Innerhalb unisensorisch visueller (blau hinterlegt) und auditorischer (rot hinterlegt) Areale wurden rechter FG, lateraler Okzipitalkortex und mittlerer STS ausschließlich durch synchrones Antwortverhalten moduliert.

#### 3.1.3.3.1.2.4 (A) Synchroner Wahrnehmung

Synchrone Wahrnehmung (,synchrone > asynchrone Wahrnehmung aller Bedingungen') aller Stimuli führte innerhalb bimodaler Gebiete beidseits zu höherer neuronaler Aktivität als asynchrone Wahrnehmung in lateral (STS, MTG, STG) und medial temporalen Regionen (perirhinaler Kortex/Gyrus parahippocampalis), in beidseitig frontalen (MFG, Gyrus praecentralis) Regionen, in rechtsseitig okzipitalen Gebieten (IOG, LG) und im Kleinhirn (s. Anhang, Tab. 9 g und Abb. 17 a). In unisensorischen Gebieten wurden durch diesen Vergleich bilateral der primäre auditorische Kortex sowie linkslaterale präfrontale Gebiete (IFG) und innerhalb visueller Areale okzipitale (FG, IOG) und frontale (Gyrus praecentralis) Gebiete beidseitig moduliert. Asynchrone Wahrnehmung über alle Stimulustypen modulierte innerhalb bimodaler Gebiete frontale (Insula, anteriorer cingulärer Kortex, ACC) bzw. präfrontale

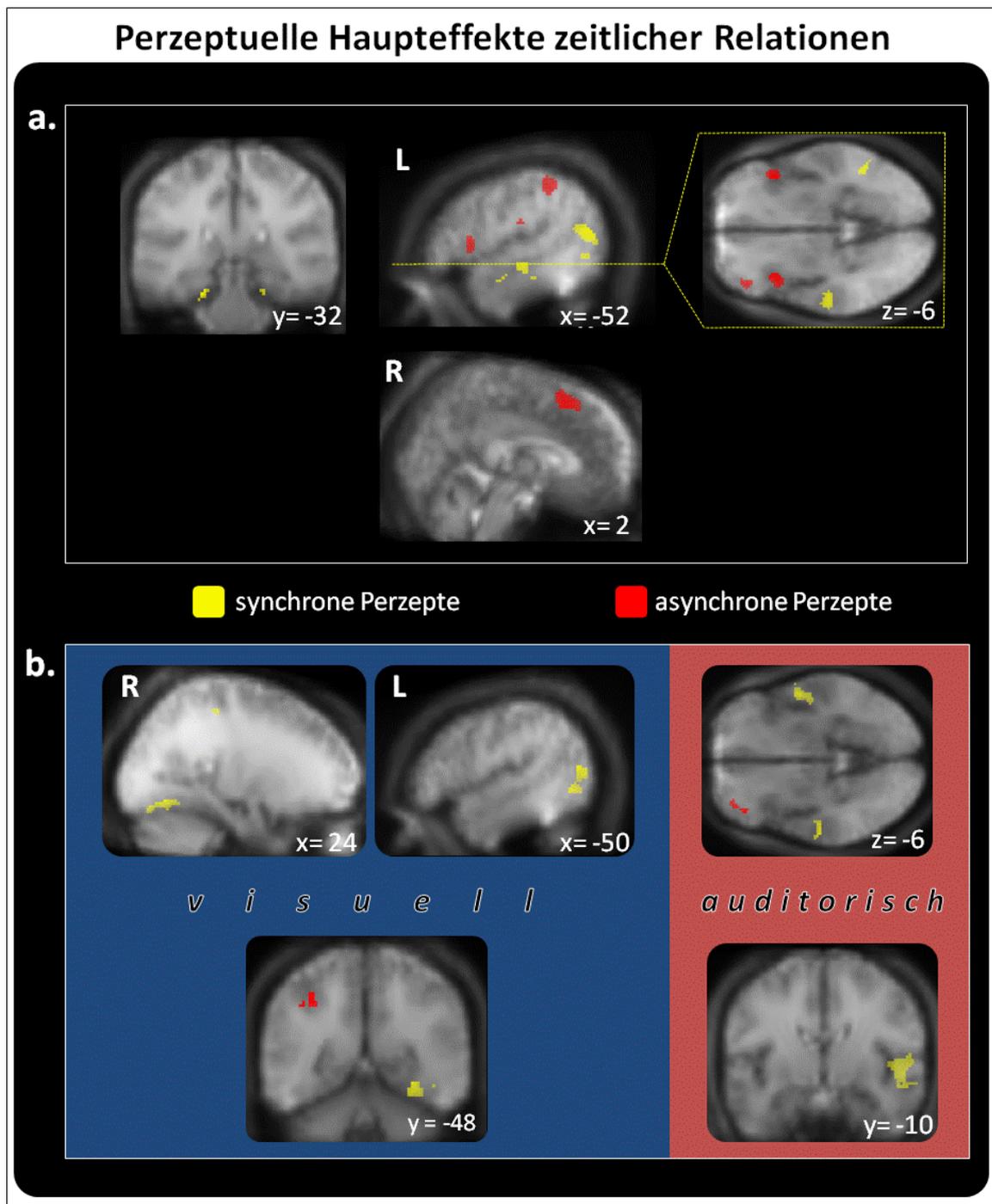
(IFG, MFG), linksseitig parietale (IPS, SMG) Areale und das Cerebellum (s. Abb. 17 a). Innerhalb unisensorischer Areale wurden linksseitig auditorische Gebiete des PFC (IFG) und der visuelle linke IPS moduliert (s. Anhang, Tab. 9 h und Abb. 17 b).

Wahrgenommene Synchronizität geht gemäß den aufgeführten Befunden mit Aktivität in einem bilateralen bimodalen kortikalen Netzwerk bestehend aus medial und lateral temporalen bzw. frontalen Aktivierungen unter Beteiligung präfrontaler prämotorischer, okzipitaler und cerebellärer Areale einher. Zudem korrelierte synchrone Wahrnehmung sowohl mit Aktivität in primären auditorischen und visuellen Kortices, als auch in auditorisch präfrontalen und visuell prämotorischen Bereichen. Asynchronizitätswahrnehmung ging vornehmlich mit linksseitig bimodalen Aktivierungen im Parietal-, im Präfrontalkortex, im anterioren Cingulum, in beiden insulären Kortices und im Cerebellum einher.

#### 3.1.3.3.1.2.5 *Synchrone Wahrnehmung von Licht vor Ton, Synchronizität und Ton vor Licht*

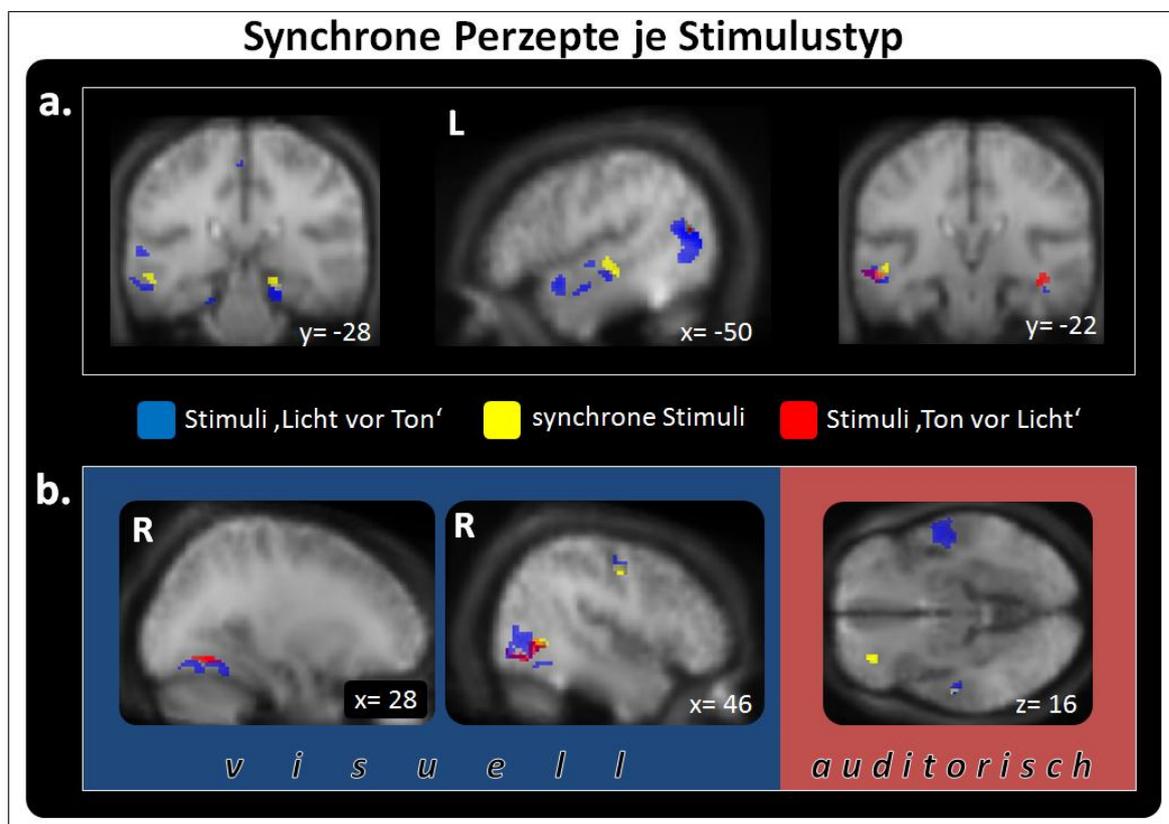
Um näher zu explorieren, auf welche Stimulustypen (Ton voran vs. Licht voran) die oben aufgeführten BOLD-Effekte für Asynchronizitätswahrnehmung zurückgehen, wurden asynchrone bzw. synchrone Antworten in bezug auf alle Stimuli, bei denen der Ton dem Licht voranging bzw. umgekehrt mit asynchronen bzw. synchronen Antworten auf synchrone Stimuli verglichen.

Synchronizitätswahrnehmung von Stimuli, bei denen der Ton dem Licht voranging (-125 und -63 ms), führten im Vergleich zu asynchroner Wahrnehmung dieser Stimuli zu höherer Aktivität im linken bimodalen Temporallappen, im rechten STS, in bimodalen Gebieten des PFC bilateral und in visuell maskierten Regionen des rechten FG (s. Anhang, Tab. 9 b und Abb. 18 a). Synchrone Wahrnehmung von Stimuli, bei denen Licht dem Ton voranging, führte zu höherer bimodaler bilateraler Aktivität in medial (perirhinaler Kortex, Gyrus parahippocampalis) und lateral temporalen (STS, STG, MTG), in frontalen (Gyrus praecentralis) bzw. präfrontalen (IFG, MFG) und in rechtsseitigen okzipitalen Gebieten (LG, FG; s. Anhang, Tab. 9 f und Abb. 18 a). Innerhalb auditorischer Gebiete ergab dieser Vergleich höhere Antwortsignale in beiden auditorischen Kortices im Vergleich zu asynchroner Wahrnehmung, innerhalb visueller Gebiete in beidseitig extrastriär okzipitalen [Gyrus okzipitalis superior (SOG), FG], prämotorischen Arealen (Gyrus praecentralis) und dem linken Parietalkortex (s. Abb. 18 b).



**Abbildung 17.** Gebiete, die durch synchrone (gelb) und asynchrone (rot) Wahrnehmung aller präsentierten Reize moduliert wurden. Gestrichelte gelbe Linien stellen die Schnittführung dar, mit der der anatomische Bezug zweier Ebenen hergestellt wurde. **a.** synchrone Perzepte modulierten den medialen (perirhinaler Kortex/Gyrus parahippocampalis) und den anterioren (aSTS) und posterioren Temporalkortex (pSTS, MTG; s. Tab. 9 g, h) innerhalb bimodaler Gebiete. Asynchrone Perzepte modulierten den linken Parietal- und Präfrontalkortex, beide anteriore Insulae und den ACC. **b.** Innerhalb unisensorisch visueller Gebiete (blau hinterlegt) modulierten synchrone Perzepte den rechten FG, LG und linken lateralen Okzipitalkortex und innerhalb auditorischer Gebiete die HG. Asynchrone Perzepte modulierten innerhalb visueller Regionen den linken IPS und innerhalb auditorischer den PFC.

Wenn synchrone Stimuli als synchron empfunden wurden, ging dies im Vergleich zu asynchroner Wahrnehmung synchroner Stimuli mit höherer bilateraler Aktivität in bimodalen Regionen des medial (perirhinaler Kortex, Gyrus parahippocampalis) und lateral (STS, STG, MTG) temporalen, des frontalen (Gyrus praecentralis) bzw. präfrontalen (IFG, MFG) und des rechten okzipitalen [Gyrus okzipitalis medius (MOG)] Kortex einher (s. Anhang, Tab. 9 c und Abb. 18 a). In auditorischen Gebieten modulierte dieser Vergleich primäre auditorische Kortices auf beiden Seiten und den linken PFC, in visuellen den rechten Okzipitallappen [Gyrus okzipitalis inferior (IOG)/FG] und beide Gyri praecentrales (s. Abb. 18 b).



**Abbildung 18.** Die Abbildung stellt Gebiete dar, die durch synchrone Wahrnehmung im Vergleich zu asynchroner Wahrnehmung von Stimuli, bei denen das Licht dem Ton (125 ms und 63 ms; blau), der Ton dem Licht (-125 ms und -63 ms; rot) voranging, und von solchen, die synchron dargeboten wurden (gelb), moduliert wurden. **a.** Synchrone Perzepte synchroner Reize und Reize, bei denen das Licht dem Ton voranging modulierten den medialen und anterioren und posterioren lateralen Temporalkortex und Reize, bei denen der Ton dem Licht voranging den lateralen Temporalkortex (s. Tab. 9 b,c,f). **b.** Synchrone Wahrnehmung beider asynchroner Reize moduliert in unisensorisch visuellen Gebieten (blau hinterlegt) den rechten FG und alle Stimulustypen den rechten ITG und lateralen Okzipitalkortex. Innerhalb auditorischer Gebiete (rot hinterlegt) modulierten Reize, bei denen das Licht dem Ton voranging HG, anterioren STS (aSTS) und Planum Polare. Synchrone Wahrnehmung synchroner Reize moduliert den linken aSTS und PFC.

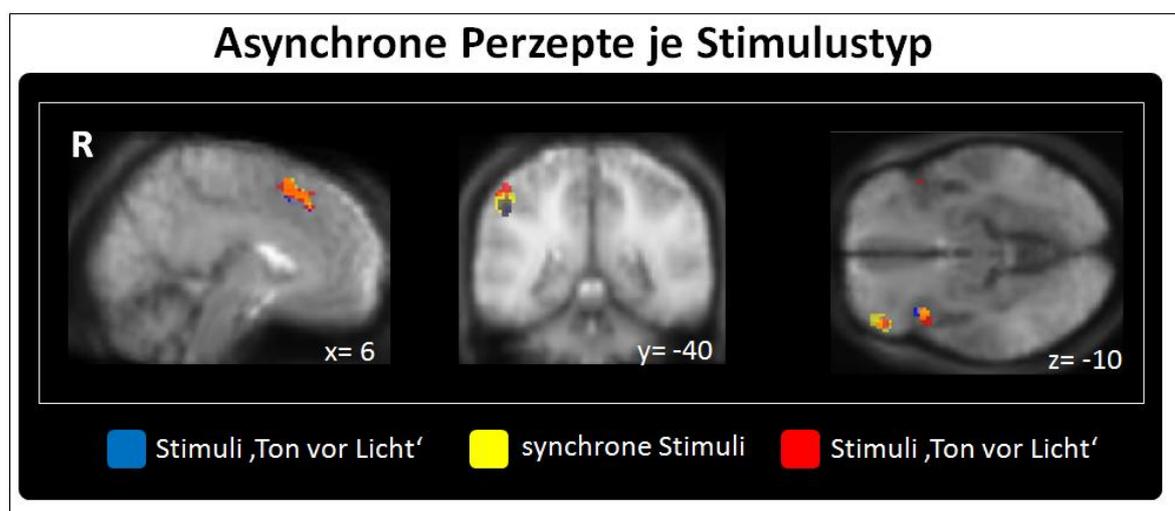
Die aufgeführten Befunde zur genaueren Untersuchung wahrgenommener Synchronizität replizieren die zuvor aufgeführten Ergebnisse zur Synchronizitätswahrnehmung aller Stimulustypen, geben jedoch darüber hinaus Aufschluss darüber, dass in bimodal temporalen Gebieten laterale Aktivität vornehmlich auf asynchrone Stimulustypen, in medial temporalen und prämotorischen Gebieten auf synchrone Stimuli und solche, bei denen das Licht dem Ton vorangeht zurückgeht. An präfrontalen Aktivierungen waren alle Stimulustypen beteiligt. Innerhalb visueller Gebiete gingen die Modulationen im Okzipitalkortex vorwiegend auf asynchrone Stimuli zurück, in prämotorischen Arealen auf synchrone und solche Stimuli, bei denen das Licht dem Ton voranging. Auditorische Gebiete im Temporalkortex wurden von Stimuli, bei denen das Licht dem Ton voranging, moduliert, aber auch von synchronen. Im PFC gingen Aktivierungen nur auf synchrone Stimuli zurück.

#### 3.1.3.3.1.2.6 *Asynchrone Wahrnehmung von Licht vor Ton, Synchronizität und Ton vor Licht*

Des Weiteren wurden, um den oben aufgeführten Kontrast zur Asynchronizitätswahrnehmung näher zu untersuchen, asynchrone Perzepte asynchroner (Licht voran und Ton voran) und synchroner Stimuli gegen synchrone Perzepte verglichen. Wenn Stimuli, bei denen der Ton dem Licht voranging, als asynchron empfunden wurden, waren der linke bimodale mittlere STS (mSTS), die Insula bilateral, der ACC, der IPS bilateral, präfrontale Gebiete (IFG, MFG) und das Cerebellum stärker aktiviert als bei synchroner Wahrnehmung dieser Stimuli (s. Anhang, Tab. 9 a und Abb. 19). In auditorischen Gebieten fanden sich für diesen Vergleich der linke auditorische Kortex, und innerhalb zuvor visuell lokalisierter Areale, der linke IPS. Wahrgenommene Asynchronizität von Stimuli, bei denen das Licht dem Ton voranging im Vergleich mit synchronen Perzepten in bezug auf diese Stimulusart, rekrutierte innerhalb bimodaler Regionen die Insula beidseits, den ACC, den linken PFC (IFG), den linken Parietalkortex und das rechte Cerebellum, jedoch keine unisensorischen Bereiche (s. Anhang, Tab. 9 e und Abb. 19). Die asynchrone Wahrnehmung synchroner Stimuli moduliert in bimodalen Gebieten beide insuläre Kortices, den ACC, den linken PFC (IFG, MFG), den Parietalkortex, das Operculum frontale beidseits und das Cerebellum beidseits stärker als

die synchrone Wahrnehmung dieser Stimuli (s. Anhang, Tab. 9 d und Abb. 19). Unisensorische Aktivierungen wurden nicht beobachtet.

Auch dieses BOLD-Antwortmuster deckt sich mit den Ergebnissen der asynchronen Wahrnehmung aller Stimulustypen (synchrone und asynchrone) und deckt auf, dass mit Ausnahme des parietalen Operculums alle Stimulustypen die genannten Hirnareale modulieren, wenn diese als asynchron wahrgenommen werden. Aktivierungen des parietalen Operculums gehen mit asynchroner Wahrnehmung ausschließlich synchroner Stimuli einher. Unisensorische Gebiete wurden lediglich von Stimuli, bei denen der Ton dem Licht voranging, moduliert.



**Abbildung 19.** Aktivierungskarten der Gebiete, die durch asynchrone Wahrnehmung im Vergleich zu synchroner moduliert wurden, wenn Stimuli präsentiert wurden, bei denen das Licht dem Ton (125 ms und 63 ms; blau), der Ton dem Licht (-125 und -63 ms; rot) vorrangig und solche, die synchron waren. Die asynchrone Beurteilung aller drei Stimulustypen moduliert innerhalb bimodaler Gebiete den ACC, den linken Parietalkortex, die anteriore Insel und PFC (s. Tab. 9 a, d, e).

### 3.1.3.4 Diskussion der Ergebnisse des fMRT-Experiments 1

Anhand der oben aufgeführten Studie wurden neuronale Korrelate physikalischer und perzeptueller Prozesse bei der Darbietung bzw. Verarbeitung einfacher audiovisueller Stimulussequenzen untersucht. Die Stimuli wurden auf Grundlage einer Voruntersuchung so gewählt, dass die zeitliche Anordnung der audiovisuellen Ereignisse innerhalb der Stimuli eine optimale Integration beider Modalitäten ermöglicht (s. 3.1.2.4). Im Vergleich zu anderen bildgebenden Studien, die audiovisuelle Integration anhand einfacher Stimuli untersuchten (Bushara et al., 2001; Calvert et al., 2001), wurden einsekündige Reize mit wiederkehrenden, kurzen auditorischen und visuellen

Ereignissen verwendet, die einer räumlich identischen Quelle entsprangen, sich weder in ihrer Dauer voneinander unterschieden, noch zeitlich überlappten und ausschließlich systematisch hinsichtlich ihrer zeitlichen Position zueinander variierten. Darüberhinaus wurden psychophysische Sensitivitätsmaße für die Stimuli bestimmt, für die Analyse der fMRT-Daten perzeptuelle, verhaltensbezogene und physikalische Effekte getrennt voneinander ausgewertet und schließlich neuronale Effekte unisensorischer und bimodaler Aktivität eindeutig zugeordnet. Die Anlage der Paradigma früherer, oben aufgeführter Studien (Bushara et al., 2001; Calvert et al., 2001) lassen dagegen aufgrund von Konfundierungen der Experimentalbedingungen und der Stimuluspräsentation in Blöcken keine eindeutigen Rückschlüsse auf neuronale Korrelate perzeptueller bimodaler Effekte der audiovisuellen Integration einfacher Stimuli zu. Im Folgenden sollen nun die Verhaltensdaten und fMRT-Daten mit Bezug auf andere Befunde diskutiert werden.

#### **3.1.3.4.1 Verhaltensdaten**

JND und PSS in bezug auf die im Scanner präsentierten Stimuli wichen von den Werten des Vorexperiments 1 ab. Die im Vergleich höhere JND und der niedrigere PSS wurden jedoch von höherer Variabilität zwischen den Probanden begleitet, auf die die zum Vorexperiment 1 verschiedenen Werte gegebenenfalls zurückgeführt werden können. Des Weiteren wurden im Vergleich zu Vorexperiment 1 keine extremen SOAs präsentiert (-250 ms bzw. 250 ms). Diese konnten im Vorexperiment 1 eventuell als interne Referenz für Asynchronizität dienen, die im fMRT-Experiment 1 nicht mehr gegeben war, so dass sich das Integrationsfenster (JND) vergrößerte, die Unsicherheit in bezug auf die (A)Synchronizitätsschätzung stieg und infolgedessen der PSS zugunsten der auditorischen Modalität verschoben wurde, bei der, sobald sie einen audiovisuellen Reiz anführt, die Beurteilung leichter gelang als wenn das Licht den Ton anführt. Diese Asymmetrie zugunsten der Tonspur wird durch die Synchronizitätsurteile sowohl des Vorexperiments 1, als auch des fMRT-Experiments 1 deutlich, wobei Probanden gegenüber Reizen, bei denen der Ton voranging während des fMRT-Experiments 1 sensibler waren als während des Vorexperiments. Diese Ergebnislage unterstützt die oben aufgeführte Vermutung hinsichtlich des abweichenden PSS in beiden Untersuchungen. Beide psychophysischen Größen liegen jedoch innerhalb von Bereichen, die sich mit anderen Studien decken (vgl. 3.1.2.4). Es kann daher davon ausgegangen werden, dass

asynchrone Stimuli perzeptuell hinreichend voneinander differenziert werden und synchrone zuverlässig miteinander verbunden werden konnten.

#### 3.1.3.4.2 fMRT-Daten

Die Analyse neuronaler Korrelate einfacher, nicht überlappender, phasisch unterschiedlich verschobener Stimuli trennte Stimuluspräsentation (unabhängig von der Wahrnehmung), Stimuluswahrnehmung und solche neuronalen Modulationen, die durch das mittlere Antwortverhalten der Probanden beeinflusst werden (unabhängig von der Stimuluspräsentation). Dabei wurden spezifische kortikale Modulationen für physikalische bzw. perzeptuelle Gleichzeitigkeit und Ungleichzeitigkeit identifiziert. Während physikalische Stimulusunterschiede (,synchrone Stimuli > asynchrone Stimuli' & ,asynchrone Stimuli > synchrone Stimuli', s. Anhang, Tab. 8 a, b) neuronale Aktivität lateralisiert im multimodalen STS und PFC hervorriefen, modulierten wahrnehmungsbezogene Effekte (,synchrone Perzeption aller Stimuli' > ,asynchrone Perzeption aller Stimuli' & ,asynchrone Perzeption aller Stimuli' > ,synchrone Perzeption aller Stimuli', s. Anhang, Tab. 9 g, h), neben einem bilateralen Netzwerk bekannter multimodaler Integrationsareale (STS, IPS, anteriore Insula, PFC), auditorisch und visuell unisensorische Kortices und das Cerebellum. Wahrgenommene Asynchronizität (,asynchrone Perzeption aller Stimuli' > ,synchrone Perzeption aller Stimuli', s. Anhang, Tab. 9 h) ging vorwiegend mit neuronalen Modulationen eines bimodalen frontoparietalen Netzwerkes einher, wahrgenommene Synchronizität mit vornehmlich frontotemporalen Modulationen. Physikalische Stimulation mit asynchronen und synchronen Reizen modulierte, wie oben erwähnt, überlappende linkstemporale Gebiete, physikalische Synchronizität zusätzlich frontale Regionen (s. Anhang, Tab. 8 a). Antwortverhaltensbezogene BOLD-Effekte von Synchronizität (s. Anhang, Tab. 10 a) waren durch ein breites Netzwerk kortikaler Modulationen in bimodalen und unisensorischen Kortices nachvollziehbar, während lediglich zwei BOLD-Modulationen, die das Antwortverhalten von Asynchronizität (s. Anhang, Tab. 10 b) abbildeten in ausschließlich bimodalen Regionen gefunden wurden.

Bisherige Studien, die zeitlichkeitsabhängige Prozesse bimodaler Stimuli anhand bildgebender Verfahren untersuchten, identifizierten neuronale Modulationen in IPS, anteriorer Insula, STS, Cerebellum, SC und präfrontalen Gebieten (Bushara et al., 2001,

Calvert et al., 2001). Diese Effekte konnten jedoch nicht zwischen wahrnehmungsbezogenen und aufgabenabhängigen Prozessen trennen. Weitere Studien analysierten die Konnektivität kortikaler Regionen bei audiovisueller Zeitverarbeitung per fMRT und lokalisierten unimodal visuelle und auditorische Areale, die mit dem bimodalen STS interagierten (Noesselt et al., 2007) und identifizierten darüber hinaus ein differentiell verknüpftes kortikales Netzwerk wahrgenommener Asynchronizität vs. Synchronizität (Dhamala et al., 2007). Dabei bestand das Netz perzeptueller Asynchronizität aus unisensorischen Kortices, dem Parietalkortex und präfrontalen Regionen und das Netzwerk perzeptueller Synchronizität aus unisensorischen und präfrontalen Kortices inklusive der SC. Dhamala et al. (2007) schlossen aus diesen Befunden, dass der Parietalkortex eine wichtige Rolle beim Erzeugen stabiler asynchroner Perzepte innehat und die SC beim Erzeugen stabiler synchroner Perzepte.

#### 3.1.3.4.2.1 *Physikalische (A)Synchronizität*

Sowohl physikalische Gleichzeitigkeit, als auch Ungleichzeitigkeit, die den Effekt der Stimuluspräsentation darstellt, und die unabhängig von subjektiver Wahrnehmung ist, modulierten den linken STS. Diesem Gebiet werden in der Literatur zahlreiche funktionelle Abhängigkeiten zugeschrieben: die Verarbeitung animierter Figuren (Pelphrey et al., 2003; Thompson et al., 2005), sprachrelevanter Merkmale (Binder et al., 2000; Jancke et al., 2002; Schmidt et al., 2008; Zatorre et al., 1992), empathischer Fähigkeiten (Gallagher et al., 2000; Vogeley et al., 2001), von Körper- (Bonda et al., 1996; Senior et al., 2000), Hand- (Grèzes et al., 1999; Rizzolatti et al., 1996) und Mundbewegungen (Bernstein et al., 2002; Puce et al., 1999), von Gesichtsmerkmalen (Haxby et al., 2000), aber auch integrative Funktionen auditorischer und visueller Signale, um sozial relevante Hinweise zu verarbeiten (Allison et al., 2000). Darüber hinaus wurde der STS in nahezu allen Untersuchungen zu multimodaler Integration moduliert (z. B. Calvert & Lewis, 2004; Fairhall & Macaluso, 2009; Macaluso et al., 2001; Meyer et al., 2007; Miller & D'Esposito, 2005; Noesselt et al., 2007; Skipper et al., 2007; Stevenson et al., 2009; Szyck et al., 2008a; 2008b; van Atteveldt et al., 2004; 2007), in denen die Kongruenz zweier bimodaler Informationsströme anhand bildgebender Verfahren untersucht wurde. Dabei können semantisch belegte (Calvert et al., 1997; Stevenson et al., 2009; Szyck et al., 2008a; Wright et al., 2003), nicht semantische, komplexe (Beauchamp et al., 2004b; Stevenson et

al., 2009) und einfache Stimuli (Bushara et al., 2001; Calvert et al., 2001) betroffen sein. Die Befunde der multimodalen Relevanz des STS wurden auch bei Tieren bestätigt (Barraclough et al., 2005; Baylis et al., 1987; Desimone & Gross, 1979; Hikosaka et al., 1988). Zudem kann der STS am Erlernen multimodaler Assoziationen beteiligt sein (Tanabe et al., 2005). Viele Befunde weisen darauf hin, dass funktionelle Unterschiede auf einer Achse von anterior nach posterior entlang des STS bestehen. So findet sprachassoziierte Verarbeitung vorwiegend anterior statt und Integrationsmechanismen, Gesichtsmerkmale und Bewegungen werden eher posterior verarbeitet (Hein & Knight, 2008). Von Kriegstein und Giraud (2004) differenzierten anteriore und posteriore Anteile des STS und fanden, dass im Bereich des anterioren STS Reaktionen auf bekannte Sprecherstimmen, im Bereich des posterioren STS vielmehr auf zeitliche Aspekte von Tönen hervorgerufen werden. Tierstudien belegen, dass Neurone im STS sowohl auf auditorische als auch auf visuelle Reize reagieren (Bruce et al., 1981; Cusick et al., 1995; Graziano et al., 1994; 1999; Hikosaka, 1993; Schroeder & Foxe, 2002) und Assoziationen zu unisensorischen auditorischen (Cappe & Barrone, 2005; Schroeder et al., 1998) und visuellen Arealen (Benevento et al., 1977; Desimone & Gross, 1979; Falchier et al., 2002; Jones & Powell, 1970; Rockland & Van Hoesen, 1994;) existieren.

Die Modulationen im STS können in Anbetracht der oben aufgeführten Untersuchungen durch die gleichzeitige Analyse auditorischer und visueller Informationsströme erklärt werden, denen Probanden ausgesetzt wurden. Dabei scheint die Art der Verschiebung unerheblich gewesen zu sein, da sowohl asynchrone als auch synchrone Reize den STS modulierten. Jedoch ist zu verzeichnen, dass beide Stimulustypen lokal distinkte, nicht überlappende Aktivierungen hervorriefen, wobei zwei lokale Maxima physikalischer Asynchronizität eines umgaben, dass durch physikalische Synchronizität moduliert wurde. Diese Modulationen befanden sich im mittleren bis posterioren Bereich des STS, die gemäß oben aufgeführter Befunde (Hein & Knight, 2008; von Kriegstein & Giraud, 2004) auf die Stimulusverarbeitung zweier Informationsströme im Gegensatz zu sprachrelevanten Aspekten schließen lassen. Des Weiteren fällt die Lateralsierung in bezug auf die temporalen Modulationen auf. Noesselt et al. (2007) boten einfache audiovisuelle Stimuli lateralisiert dar und beobachteten kontralaterale Aktivierungen im posterioren STS wenn Stimuli gleichzeitig präsentiert wurden. Miller und D'Esposito (2005) untersuchten

audiovisuelle Sprachstimuli im Scanner und zeigten, dass perzeptuelle Synchronizität während der Verarbeitung ihrer Stimuli linkslateral und perzeptuelle Asynchronizität rechtslateral verarbeitet wurde. Sekiyama et al. (2003) berichteten linkslaterale Modulationen des STS für verständliche im Vergleich zu schwer verständlicher audiovisueller Sprache. Da die Stimuli der vorliegenden Studie zentral präsentiert wurden und alle Stimulusbedingungen in derselben Hemisphäre neuronal moduliert wurden, kann die kortikale Lateralisierung der vorliegenden Untersuchung nicht, wie bei den oben aufgeführten Studien, auf Stimulusmerkmale attribuiert werden. Bushara et al. (2001), Calvert et al. (2001), und Dhamala et al. (2007) berichteten keine Lateralisierungseffekte, so dass die Vermutung geäußert werden kann, dass Lateralisierungseffekte vorwiegend bei sprachassoziierten audiovisuellen Reizmerkmalen auftreten (Campbell, 1990; Sestieri et al., 2006) und in der vorliegenden Studie nicht erklärt werden können. Die vorliegenden Ergebnisse auf Grundlage einfacher audiovisueller Stimuli deuten jedoch darauf hin, dass auch diese Reize lateralisiert verarbeitet werden. Möglicherweise ist der bestimmende Faktor die zeitliche Struktur der präsentierten Stimuli, wobei ‚langsame‘ Reize eher rechts- und ‚schnellere‘ eher linkshemisphärisch verarbeitet werden.

Neben Modulationen des mittleren und posterioren STS wurde ausschließlich bei synchron präsentierten Stimuli bilateral der PFC moduliert. Aktivierungen in diesem Bereich werden bei Untersuchungen zur audiovisuellen Integration oft beobachtet (Bushara et al., 2001; Calvert et al., 2001; Dhamala et al., 2007; Miller & D’Esposito, 2005; Ojanen et al., 2005; Sekiyama et al., 2003; Stevenson et al., 2009).

Der PFC wird mit funktionalen Korrelaten des Arbeitsgedächtnisses in Zusammenhang gebracht (Baddeley, 1992; Leung et al., 2002; Manoach et al., 2004; Smith & Jonides, 1998), mit Aktivität des Kurzzeitgedächtnisses, mit antizipatorischen, planenden Funktionen, mit Interferenzkontrolle (Fuster, 1985), aber auch mit integrativen Mechanismen. So wurde an Affen gezeigt, dass der PFC über zahlreiche Afferenzen und Efferenzen neuronale Signale mehrerer Modalitäten empfangen und senden kann (Chavis & Pandya, 1976; Fuster, 1985; Jones & Powell, 1970). Es wurden direkte Verbindungen zu auditorischen (Poremba et al., 2003; Romanski et al., 1999) und visuellen (Benevento et al., 1977) Regionen bei Tieren gefunden. Darüber hinaus wurde gezeigt, dass Zellen im PFC auf audiovisuell assoziierte Stimulusverbindungen

reagieren (Bodner et al., 1996; Fuster et al., 2000). Der PFC kommuniziert mit unisensorischen Kortices (Sugihara et al., 2006) und tut dies über direkte Verbindungen zu unimodalen Regionen, wie an Tierstudien gezeigt wurde (Budinger et al., 2007; Wang & Burkhalter, 2007). Zahlreiche weitere Tierstudien ergaben außerdem, dass sehr stabile reziproke neuronale Verbindungen zwischen präfrontalen Regionen und dem STS existieren (Barbas, 1992; Barraclough et al., 2005; Cavada & Goldman-Rakic, 1989; Cusick et al., 1995; Oram & Perrett, 1996; Petrides und Pandya, 1988; Romanski et al., 1999; Seltzer & Pandya, 1989). Dieser Tatbestand lässt darauf schließen, dass der PFC bereits integrierte Informationen aus anderen kortikalen Regionen erhält (Sugihara et al., 2006).

Caclin und Fonlupt (2006) präsentierten eine auditorische Zielaufgabe und ermittelten anhand von Konnektivitätsanalysen, dass der DLPFC stark mit auditorischen, visuellen, parietalen Arealen und dem STS interagiert und dabei die Rolle übernimmt, relevante sensorische Information von irrelevanter zu trennen und in Abhängigkeit davon per Rückkopplungsmechanismen unisensorische Kortices moduliert (Bishop & Miller, 2008). Andere Autoren schlagen vor, dass der PFC für das Aufrechterhalten sensorischer Informationen zuständig ist (Miller, 2000; Rizzuto, 2005; Tanabe et al., 2005) und dadurch zielgerichtete, aufgabenrelevante Prozesse koordiniert. Damit sei der PFC in der Lage, verhaltensbezogene Kontexte integrierter Informationen zu repräsentieren. Asaad et al. (2000) zeigten, dass neuronale Antworten auf Zielreize im PFC stark aufgabenspezifisch sind und unterstützen damit diese Vermutung. D.h. dass der PFC verhaltens- bzw. aufgabenabhängige Ziele aufrechterhält (Banich et al., 2001; Botvinick et al., 2001; Weissman et al., 2004), was wiederum zu erhöhter Aufmerksamkeit führt, wenn irrelevante Stimuli mit verhaltensbezogenen Zielen konkurrieren (Kerns et al., 2004). Beauchamp (2004b) untersuchte neuronale Aktivität, wenn semantisch belegte und nicht semantische Stimuli audiovisuell dargeboten wurden, und beschrieb, dass der DLPFC besonders stark bei schwierigen Stimuli (gemessen an längerer Reaktionszeit) moduliert wurde und brachte den PFC mit Aufgabenschwierigkeit in Verbindung. Weitere bildgebende Untersuchungen ergaben, dass erhöhte präfrontale Aktivität (Badre & Wagner, 2004; Durston et al., 2003) und erhöhte Aktivität im ACC (z. B. Kerns et al., 2004; s. unten) mit Aufgaben einhergeht, bei denen Antwortkonflikte induziert werden.

Loose et al. (2003) berichteten über erhöhte präfrontale Aktivität, wenn Probanden ihre Aufmerksamkeit zwischen visuellen und auditorischen Reizen teilen sollten und schlagen vor, dass der PFC durch potentiell zu verwechselnde Informationen moduliert wird (Johnson & Zatorre, 2006; Ojanen, 2005). Vohn et al. (2007) zeigten, dass der PFC bei crossmodal im Vergleich zu intramodal attendierten Aufgaben stärker moduliert wird und bekräftigen diese These. Giard, Peronnet (1999) und Calvert (2001) schlagen vor, dass präfrontale Beteiligung das Erlernen und Erinnern multimodaler Assoziationen reflektiert und mitunter eine Art Arbeitsgedächtnis multimodaler Integration darstellt (Yoo et al., 2004).

Fuster (1985) beschreibt, dass sich präfrontale Zellen bei Affen mit unterschiedlicher zeitlicher Verzögerung nach Stimuluspräsentation bzw. vor motorischer Antwort entladen und dass Affen mit Läsionen im PFC Leistungsnachteile erlitten, wenn bei Experimenten zeitliche separate Informationen integriert werden sollten. Onoe et al. (2001) zeigten bei Affen, dass ein experimentell herbeigeführter Ausfall neuronaler Funktionen im DLPFC zu Fehlleistungen bei Aufgaben führte, bei denen zeitliche Abstände zweier visueller Reize eingeschätzt werden sollten. Menschen mit präfrontalen Verletzungen schätzen die Dauer von Tonreizen schlechter ein als gesunde Probanden (Harrington et al., 1998a). Diese Befunde weisen darauf hin, dass dem PFC neben den zuvor aufgeführten Funktionen auch kritische Aufgaben hinsichtlich der Wahrnehmung von Zeitprozessen zukommen.

Insgesamt lässt sich aus den oben aufgeführten Befunden schließen, dass der PFC eine Region zu sein scheint, in der Informationen aus mehreren Modalitäten integriert werden können, die stärker moduliert wird, wenn aufmerksamkeitsbezogene Prozesse zwischen mehr als einer Wahrnehmungsmodalität kontrolliert werden müssen, wenn verhaltensabhängige Konsequenzen multimodaler Prozesse reguliert werden und die außerdem zeitkritische Prozesse reguliert. Dabei interagieren präfrontale Zellen mit posterioren unisensorischen und multimodalen.

Bildgebende Studien, die aufmerksamkeitsabhängige Prozesse untersuchten, zeigten, dass Parietallappen, Gyrus praecentralis, Okzipitalkortex, aber auch präfrontale Areale durch verdeckte Aufmerksamkeitslenkung moduliert werden (z. B. Hopfinger et al., 2000). Anhand von Läsionsstudien wurde gezeigt, dass neben posterior-parietalen Arealen (z. B. Driver & Mattingley, 1998) auch Bereiche des frontalen Kortex (Heilman

& Valenstein, 1972) und des Gyrus cinguli (Watson et al., 1973) an räumlicher Aufmerksamkeitslenkung beteiligt sind. Basierend auf Tierstudien postulierte Mesulam (1981; 1990) ein Modell der Lenkung visueller Aufmerksamkeit, an dem kortikale und subkortikale anatomische Strukturen beteiligt sind. Dazu gehören das frontale Augenfeld, der Gyrus cinguli und posterior-parietale Bereiche. Die Beteiligung präfrontaler Areale im Rahmen der vorliegenden Studie könnte in Anbetracht der Befunde zur Aufmerksamkeitsregulierung, wie oben erwähnt, aufmerksamkeitsabhängige Prozesse reflektieren.

Andere Untersuchungen zur audiovisuellen Kongruenz lokalisierten höhere Aktivität im PFC bei inkongruenten im Vergleich zu kongruenten komplexen (Belardinelli et al., 2004) und semantisch belegten Reizen (Ojanen et al., 2005; Saito et al., 2005; Sekiyama et al., 2003), aber auch während aufgabenabhängiger Verarbeitung im Vergleich zur Abwesenheit einer Aufgabe (van Atteveldt et al., 2007) und bei einfachen audiovisuellen Stimuli (Bushara et al., 2003).

Studien, die die Zeitlichkeit zwischen beiden Wahrnehmungsmodalitäten untersuchten, assoziierten rechtslaterale präfrontale Aktivität mit physikalisch asynchronen (im Vergleich zu synchronen) und linkslaterale mit synchronisierter Wahrnehmung (Miller & D'Esposito, 2005) bei semantisch belegten Reizen und mit der Präsentation zu bewertender asynchroner Stimuli (Bushara et al., 2001) bei einfachen Reizen. Des Weiteren entdeckten Calvert et al. (2001) im Sinne einer Superadditivität stärkere BOLD-Antworten auf einfache audiovisuelle Reize in dieser Region als auf unimodale. Dhamala et al. (2007) gehen davon aus, dass der PFC zu einem Netzwerk audiovisueller Integration gehört, das die Stabilität eines Zeitlichkeitsperzepts modulieren kann und daher sowohl synchronen, als auch asynchronen Reizen zugeordnet werden kann. Des Weiteren scheint der PFC Ressourcen zu verwalten, um Zeitinformationen zu erlangen und aufrechtzuerhalten (Casini & Ivry, 1999; Mangels et al., 1998).

Die hier beschriebene Modulation des PFC während der antwortunabhängigen Präsentation synchroner audiovisueller Reize kann durch das bimodale Paradigma erklärt werden, da anhand der oben aufgeführten Studien belegt wurde, dass diese Region kritisch für multimodale Reizverarbeitung zu sein scheint. Gemäß anderen erwähnten Studien könnte diese präfrontale Modulation erhöhte attentionale

Ressourcen anzeigen, die aktiviert wurden, sobald sich auditorische und visuelle Ereignisse überschneiden und daher nur bei synchronen Stimuli angesprochen wurden. Viele Studien erklärten Aktivität im PFC durch Konflikte, die durch inhaltliche Inkongruenz und der damit verbundenen Fehlervorhersage zwischen den Modalitäten entstehen. In Anbetracht der Tatsache, dass in der vorliegenden Studie der PFC nur während der Präsentation synchroner Stimuli neuronal moduliert wurde, kann diese Erklärung vermutlich nicht auf die zeitliche Inkongruenz (Asynchronizität) einfacher audiovisueller Stimuli übertragen werden. Auch konnte der Befund von Miller und D'Esposito (2005) nicht repliziert werden, bei denen physikalische Asynchronizität mit präfrontaler Aktivität einherging, die die Autoren auf die Sprachrelevanz ihrer Stimuli zurückführten.

Aufgrund der Relevanz des PFC für zeitliche Aspekte der Stimulusverarbeitung, könnte die erhöhte Aktivität mit höherer Salienz synchroner Stimuli einhergegangen sein und über Rückkopplungsmechanismen die gleichzeitige Stimulation auditorischer und visueller Kortices signalisiert haben.

Es konnte ermittelt werden, dass die temporalen Aktivierungen infolge physikalisch asynchroner Stimulation in der vorliegenden Untersuchung auf Reize zurückgingen, bei denen das Licht dem Ton voranging. Asynchronizität bei dieser Stimulusart wurde schwerer entdeckt als bei Stimuli, bei denen der Ton voranging (s. 3.1.3.3.1). Dies könnte ein Hinweis darauf sein, dass der bimodale STS besonders bei mehrdeutigen Stimuli die Informationen mehrerer Sinneskanäle moduliert. Stevenson et al. (2009) zeigten, dass neuronale Antworten im STS bei der Verarbeitung kontrastarmer bimodaler Stimuli stärker waren, als bei kontrastreichen (salienten). Das Konzept, dass eine Zellantwort mit abnehmender Stimulussalienz bimodaler Reize steigt, ist unter dem Begriff der inversen Effektivität bekannt (Stein & Meredith, 1993; s. 1.2.1). Analog dazu könnten zeitlich verschobene Stimuli des vorliegenden Experiments, bei denen das Licht dem Ton voranging, weniger salient gewesen sein als Stimuli, bei denen der Ton dem Licht voranging. Ähnliche Befunde wurden unter Verwendung auditorischer, aber auch audiovisueller Sprachstimuli entdeckt, die ergaben, dass anteriore Regionen des STS durch verständliche und posteriore Regionen durch schlecht verständliche Reize moduliert wurden (Buchsbaum & D'Esposito, 2009; Scott et al., 2000). Die Aktivität im mittleren und posterioren STS in der vorliegenden Studie könnte in Anbetracht dieser

---

Befunde auf eine zeitliche Ambiguität des betroffenen Stimulustyps (Licht voran) hinweisen, die nicht eindeutig zugeordnet werden konnte und daher höhere BOLD-Antworten in diesen Bereichen hervorrief.

#### 3.1.3.4.2.1.1 (A)synchrone Perzeption

Bimodale, wahrnehmungsbezogene Modulationen, unabhängig von der präsentierten Stimulusart im vorliegenden Experiment, wurden für Synchronizität in temporalen, frontalen und cerebellären Regionen gefunden. Asynchrone Perzeption modulierte innerhalb bimodaler Gebiete frontale, parietale, insuläre und cerebelläre Kortices.

Dabei modulierten synchrone Perzepte spezifisch den medialen Temporalkortex im Bereich des perirhinalen Kortex bzw. des Gyrus parahippocampalis. Diese Region wird mit dem Verknüpfen von Objektmerkmalen in Verbindung gebracht (Goulet & Murray, 2001). So wurde diese Region durch komplexe visuelle Diskriminierungsaufgaben moduliert (Devlin & Price, 2007; Gorno-Tempini et al., 1998; Xu et al., 2007). Affen (Buckley et al., 2001; Bussey et al., 2002;), aber auch Menschen (Barense et al., 2005) mit Läsionen in diesem Bereich konnten Merkmale komplexer visueller Reizen nicht voneinander diskriminieren. Darüber hinaus verfügt dieses kortikale Gebiet über direkte neuronale Projektionsbahnen zur Insula, zum Parietal- und Frontalkortex, zu visuellen Gebieten und zum multimodalen STS (Carmichael & Price, 1995; Friedman et al., 1986; Insausti et al., 1998; Suzuki & Amaral, 1994; Suzuki, 1996) und wird daher durch Informationen mehrerer Sinneskanäle moduliert. So wurde es durch erfolgreich integrierte, kongruente visuo-taktile (Holdstock et al., 2009) und audiovisuelle (Belardinelli et al., 2004), durch semantisch belegte, inkongruente audiovisuelle Reize (Taylor et al., 2006) und durch erlernte Assoziationen einfacher (Tanabe et al., 2005) und semantischer, audiovisueller Stimuli (Buchsbaum & D'Esposito, 2009) moduliert. Eine Affenstudie zeigte, dass Makaken mit lädierten perirhinalen Kortices nicht in der Lage waren, visuo-taktile Stimuli zu integrieren (Goulet & Murray, 2001). Hasson et al. (2008) demonstrierten im Rahmen einer fMRT-Studie, dass unter anderem der Gyrus parahippocampalis durch kurze, zeitlich zusammenhängende visuelle Sequenzen im Vergleich zu zeitlich ungeordneten Abschnitten moduliert wird. Die Autoren schließen, dass innerhalb dieses Areal

Neurone existieren, die zeitliche Abfolgen von Szenen innerhalb einer gewissen Zeitspanne verarbeiten können. Viele Autoren kommen zu dem Schluss, dass der mediale temporale Kortex essentiell für das Erstellen von Verbindungen zwischen Stimulusmerkmalen und sogar für das Einordnen dieser Merkmale in konzeptuelle Repräsentationen höherer Ordnung ist. In diesem Zusammenhang scheint dieses Gebiet im Rahmen des vorliegenden Experiments essentiell für die Verbindung auditorischer und visueller Ereignisse und für Einordnung zeitlich unterschiedlich verschobener audiovisueller Stimuli in ein synchrones Perzept gewesen zu sein. Damit werden bisherige Befunde durch die vorliegenden Ergebnisse erweitert und funktionelle Relevanz des medialen Temporalkortex auch für zeitliche Faktoren einfacher audiovisueller Stimuli festgestellt. Dies widerspricht Vermutungen anderer Autoren, die diesen kortikalen Bereich vor allem mit der Verarbeitung lebender und semantisch belegter multimodaler Informationen in Verbindung bringen (Taylor et al., 2006).

Die Datenanalyse zeigt auf, dass die Modulation des medialen Temporalkortex auf die synchrone Wahrnehmung synchroner Reize und solcher, bei denen das Licht dem Ton voranging, zurückzuführen ist. Stimuli, bei denen das Licht dem Ton voranging, waren laut Verhaltensdaten schwerer zu beurteilen (s. 3.1.3.2.1). Dies könnte ein Hinweis darauf sein, dass der mediale Temporalkortex vorwiegend an der Verarbeitung zeitlich ambiguer Reize beteiligt ist.

Im Vergleich zu asynchroner Wahrnehmung ist die Aktivität unisensorischer auditorischer, wie den Heschl'schen Gyri (HG), dem Planum Temporale (PT), dem anterioren STS (aSTS), und visueller Areale, wie dem Gyrus occipitalis medius bzw. inferior (MOG/IOG) und dem Gyrus Fusiformis (FG) auffallend, die durch synchrone Perzeption moduliert wurden.

Teilbereiche des FG werden in der Literatur als Gesichtsareal diskutiert (fusiform face area, FFA, Kanwisher et al., 1997). Studien belegen jedoch auch, dass diese Gegend Aktivierungen aufweist, wenn unbekannte multimodale Verbindungen erlernt werden (Gonzalo et al., 2000). IOG und MOG sind dem visuellen Kortex zugehörig und werden beispielsweise mit der Verarbeitung räumlicher Frequenzmuster (Rotshtein et al., 2007) und früher Erkennung von Gesichtsmerkmalen in Verbindung gebracht (Haxby et al., 2000). Das PT, das posterior der HG gelegen und möglicherweise Teil des Wernicke-Areals und sicherlich Teil des auditorischen Kortex ist (auditorische Belt-Region;

Formisano et al., 2003; Galaburda & Pandya, 1983; Pandya, 1995), ist funktionell und anatomisch stark lateralisiert. Während seine linkseitige Aktivierung oft mit der Verarbeitung sprachlicher Stimuli einhergeht, wird rechtseitige Aktivität mit der Verarbeitung nichtsprachlicher Stimuli wie z. B. Musik in Verbindung gebracht (s. Shapleske et al., 1999 für eine Zusammenfassung). Dahinter steckt die Annahme, dass im linken PT eher schnelle zeitliche Aspekte akustischer Stimuli verarbeitet werden und im rechten akustische Reize eher auf ihre Frequenzkomponenten hin analysiert werden (Jancke et al., 2002; Kumar et al., 2007; Liègeois-Chauvel et al., 1999; Zatorre et al., 2002). Zudem weisen Personen mit absolutem Gehör ein größeres linkes PT auf als normal Hörende (Schlaug et al. 1995). Neuere Arbeiten bringen Aktivität im PT mit mentaler Vorstellung von akustischen Reizen in Verbindung (Bunzeck et al., 2005; Jancke & Shah, 2004; Voisin et al., 2006). Calvert et al. (1997) brachten das rechte PT in Zusammenhang mit der Verarbeitung von Lippenbewegungen und Hasegawa et al. (2004) konnten zeigen, dass das PT bei der Verbindung von Handbewegungen mit Tönen während des Klavierspielens, d.h. bei audiovisueller Integration, aktiv ist. Obwohl das PT klassischerweise mit der Verarbeitung akustischer Reize in Verbindung gebracht wird, zeigen einige Studien, dass dieses Areal auch an der Integration audiovisueller Reize beteiligt ist und lassen daher die hier gefundene Aktivierung plausibel erscheinen (Bizley et al., 2007a; 2007b; Smith et al., 2004).

Interessanterweise waren in der vorliegenden Studie im Vergleich zu asynchroner Wahrnehmung primäre auditorische und visuelle Areale während synchroner Wahrnehmung zusammen mit dem bimodalen STS stärker aktiviert. Ein Befund, den auch Noesselt et al. (2007) für sensorische Verarbeitung audiovisueller Gleichzeitigkeit berichteten. Sie zeigten darüber hinaus, dass die Kommunikation zwischen unisensorischen Arealen und dem multimodalen STS bei zeitlicher audiovisueller Kongruenz höher war, als bei der Verarbeitung zeitlich auseinanderlaufender Informationen zwischen beiden Informationskanälen. Die Beteiligung primärer unisensorischer Areale an multisensorischer Verarbeitung wird in der Literatur auch anhand von Untersuchungen an Tieren diskutiert (Brosch et al., 2005; Foxe & Schroeder, 2005; Ghazanfar & Schroeder, 2006; Kayser & Logothetis, 2009; Lakatos et al., 2007). Ghazanfar et al. (2008) konnten im Rahmen einer Untersuchung an Affen zeigen, dass visuelle Reize dazu beitragen können, Oszillationen von Zellen in

auditorischem Kortex und multimodalem STS zu synchronisieren und so zu einer erfolgreichen Integration zwischen Bild und Ton beitragen können. Es wird vermutet, dass unisensorische Gebiete infolge multimodaler Stimulation über Rückkopplungsmechanismen mit dem multisensorischen STS kommunizieren (Bushara et al., 2003; Driver & Noesselt, 2008; Fairhall & Macaluso, 2009; Kayser et al., 2008). Hinweise auf direkte Verbindungen zwischen visuellem bzw. auditorischem Kortex und dem STS liefern Tierstudien (Falchier et al., 2002). Ghanzanfar et al. (2005) stellten fest, dass der auditorische Kortex bei Affen durch audiovisuell kongruente Stimuli moduliert wurde und schließen auf Rückkopplungsprozesse zwischen primärem auditorischen Kortex und multimodalem STS. Brosch et al. (2005) konnten auditorische Areale bei Affen durch visuelle Reize modulieren. Anhand EKP bei Menschen konnte Aktivität in unisensorischen Arealen durch multisensorische Stimulation nachgewiesen werden (Murray et al., 2005), aber auch anhand fMRT bei Affen (Kayser et al., 2007). Weitere fMRT-Studien am Menschen zeigten, dass klassische unisensorische Regionen multisensorisch moduliert wurden (Amedi et al., 2002; Calvert et al., 1999; Macaluso et al., 2000; Martuzzi et al., 2007; Pekkola et al., 2005; Watkins et al., 2006). Konnektivitätsanalysen ergaben darüber hinaus, dass frühe auditorische Gebiete stark mit dem STS interagieren (Noppeney et al., 2008; Sukhbinder et al., 2007). Tanabe et al. (2005) schlagen vor, dass multimodal modulierte Aktivität in auditorischen und visuellen Kortices ein sensorisches Gedächtnis reflektiert. Eine weitere Möglichkeit ist die Bindung bzw. Verstärkung multimodaler Objekteigenschaften (s. Noesselt et al., 2007; Schoenfeld et al., 2003).

Untersuchungen an Tieren ergaben des Weiteren, dass auditorischer und visueller Kortex über direkte Verbindungen unmittelbar miteinander kommunizieren können (Bizley et al., 2007a; Budinger et al., 2006; Falchier et al., 2002; Rockland & Ojima, 2003).

Die besprochenen Befunde sind kongruent mit den hier vorliegenden Befunden und weisen darauf hin, dass die Modulation multimodaler temporaler und unisensorischer auditorischer und visueller Areale durch perzeptuelle Synchronizität ein Netzwerk reflektiert, das multimodale relative Zeitinformationen über Rückkopplungsmechanismen verarbeitet und austauscht. Dabei könnte die gleichzeitige

Aktivierung unisensorischer Kortices ausschlaggebend für die Entstehung eines Synchronizitätsperzepts sein, die durch den multimodalen STS moduliert wird.

Es fiel zudem auf, dass im vorliegenden Experiment anteriore insuläre, parietale und cinguläre Gebiete spezifisch durch asynchrone Perzeption moduliert wurden. Der insuläre Kortex gilt als Gebiet, das für viszerale und limbische Funktionen zuständig ist (Buechel et al., 1998; Fischer et al., 1998; Ploghaus et al., 1999). Studien an Primaten und anderen Säugetieren konnten jedoch auch zeigen, dass die Insula über direkte Bahnen zu prämotorischen, parietalen, cingulären, perirhinalen, auditorischen und visuellen Arealen verfügt und Verbindungen zum frontalen Operculum und zum STS aufweist (Augustine, 1996; Graybiel, 1973; Guldin & Markowitsch, 1984; Hicks et al., 1988). Einzelzellableitungen ergaben, dass die neuronale Aktivität im insulären Kortex in Abhängigkeit vom Ausmaß an zeitlicher Verzögerung von kombinierten auditorischen und visuellen Stimuli moduliert wird (Benevento et al., 1977; Kayser & Logothetis, 2007; Loe & Benevento, 1969; Meredith et al., 1987). Andere Autoren vermuten, dass die Insula den Austausch zwischen unimodalen Kortices reguliert (Amedi et al., 2005) und bildgebende Studien zeigten, dass die anteriore Insula durch Relevanzverarbeitung von Reizen (Buechel et al., 1998), durch das Erlernen hemmenden Vermeidungsverhaltens (Bermudez-Rattoni et al., 1991) und durch Fehlerverarbeitung (Menon et al., 2001) angesprochen wird.

In einer Studie mit einfachen (a)synchronen audiovisuellen Reizen fanden Calvert et al. (2001) beidseitige Aktivierungen der Insula. Miller und D'Esposito (2005) wiesen unter anderem linkslaterale Aktivierungen der anterioren Insula während der Präsentation asynchroner Sprachstimuli nach. Bushara et al. (2001) konnten zeigen, dass die rechte Insula bei der Beurteilung gegeneinander verschobener einfacher auditorischer und visueller Stimuli angesprochen wird. Herdener et al. (2009) zeigten, dass diese Region durch zeitversetzte einfache auditorisch und visuelle Reize moduliert wurde. Geiser et al. (2008) konnten zeigen, dass unter anderem die anteriore Insula an der Wahrnehmung akustischer Zeitintervalle beteiligt ist. In weiteren bildgebenden Studien zur multimodalen Integration ging insuläre Aktivität mit der Präsentation semantisch belegter (Meyer et al., 2007; Stevenson et al. 2009), kongruenter (Callan et al., 2004; Holdstock et al., 2009; Sestieri et al., 2006) und inkongruenter (Szycik et al., 2008b) Stimuli einher. Preuschoff et al. (2008) diskutieren Modulationen der anterioren Insula durch

Vorhersagefehler. In diesem Zusammenhang könnte die in dieser Untersuchung vorliegende Modulation der anterioren Insula mit entscheidungsabhängigen Prozessen in Verbindung gebracht werden, die die Beurteilung asynchroner Sequenzen mit sich brachte. Die Beurteilung dieses Stimulustyps (im Gegensatz zu synchronen Reizen) würde dementsprechend mit erhöhter subjektiver Fehlerwahrscheinlichkeit einhergehen.

Die oben aufgeführten Befunde sprechen dafür, dass die Insula ein multisensorisches Gebiet sein könnte, das mit multi- und unimodalen Regionen des Kortex kommunizieren kann. Es wurde mit multimodaler Integration in Verbindung gebracht, insbesondere mit der Verarbeitung zeitlich verschobener Stimuli, wobei bislang unklar ist, inwieweit die Aufgabenstellung bzw. -verarbeitung diese Modulationen beeinflusst. Die vorliegende Studie bringt beide anteriore Insulae in Verbindung mit asynchroner Wahrnehmung einfacher audiovisueller asynchroner und synchroner Reize. Damit werden bisherige Befunde untermauert, die insuläre Modulationen mit asynchronen Reizen assoziierten und diese gleichzeitig erweitert, indem die Modulationen durch perzeptuelle Effekte erklärt wurden. Unter Einbeziehung der Bedeutung der Insula für kontrollabhängige und fehlerregulierende Verarbeitungsmechanismen könnte die insuläre Aktivität im vorliegenden Experiment durch ein Abschätzen der beiden Informationsströme hinsichtlich ihrer Zeitlichkeit erklärt werden, das vor allem bei asynchroner Wahrnehmung hervortrat, da die zeitliche Auswertung eventuell mehr Kontrolle verlangte als bei synchroner Wahrnehmung. Dies könnte insbesondere der Fall gewesen sein, weil die Wahrnehmung fortwährend gegen ein synchrones Perzept abgegrenzt werden musste und beide Informationsströme zeitlich voneinander dissoziiert werden mussten. Die Analyse der Interaktionseffekte [d.h. der Vergleich synchroner gegen asynchroner Perzepte in Abhängigkeit des Stimulustyps (‘Licht vor Ton’, ‘synchron’, ‘Ton vor Licht’)] zeigte, dass dieser Effekt bei allen Stimulustypen, synchronen und asynchronen, auftrat, was das Konzept einer Dissoziation unterstreicht, da auch synchrone Ereignisse zeitlich voneinander getrennt werden mussten. Die Ergebnisse passen in den Corpus der meisten bisherigen Befunde (Bushara et al., 2001; Calvert, 2001; Calvert & Lewis, 2004; Geiser et al., 2008; Herdener et al., 2009; Holdstock et al., 2009; Meyer et al., 2007; Stevenson et al., 2009; Szykik et al., 2008b), stehen jedoch teilweise in Widerspruch zu den Befunden von Miller & D’Esposito (2005), die insuläre Aktivität vor allem mit physikalischen Effekten asynchroner audiovisueller

Sprachstimuli in Verbindung brachten, wogegen im vorliegenden Experiment diese eindeutig auf wahrnehmungsabhängige Prozesse zurückgeführt werden konnten. Die Befunde der hier besprochenen Arbeit bekräftigen jedoch die Vermutung anderer Autoren, dass der insuläre Kortex eine zentrale Rolle beim zeitlichen Vergleich asynchroner audiovisueller Stimuli spielt. Weitere Hinweise auf die Rolle der Insula werden durch die Paradigma der Experimente 2 (s. 3.2.7.4.2) und 3 (s. 3.3.3.2.2) untersucht.

Neben insulären Gebieten wurde der linke Parietalkortex im Bereich des IPS und seiner Umgebung spezifisch durch asynchrone Wahrnehmung moduliert. Diese Region war in bisherigen bildgebenden Verfahren in Aufgaben der räumlichen Übereinstimmung während audiovisueller und visuo-taktiler Integration involviert (Eimer, 1999; Eimer & Schröger, 1998; Macaluso et al., 2000; Macaluso et al., 2004). An Affen und Menschen wurde gezeigt, dass der IPS ein multisensorisches Areal ist, das Bewegungsinformationen verschiedener Modalitäten integrieren und sie in einen körperzentrierten Bezugsrahmen überführen kann (Bremmer et al., 2001). Der IPS war in Studien zur Koordination multimodaler räumlicher Aufmerksamkeit aktiv (Bushara et al., 1999; Corbetta et al., 2002; Driver & Spence, 1998; Eimer, 1999; Macaluso et al., 2000), ebenso wie in Studien zur audiovisuellen (A)Synchronizität, in denen man neuronale Netzwerke fand, innerhalb derer der IPS eine Rolle spielt (Bushara et al., 2001; Calvert, 2001; Macaluso et al., 2004; Miller & D'Esposito, 2005). Des Weiteren wird der Parietalkortex mit dem Wechseln der Aufmerksamkeit von einer Modalität zu einer anderen (Lewis et al., 2000; Shomstein & Yantis, 2004; 2006), der Verarbeitung von multimodalen Objektmerkmalen (Grefkes et al., 2002; Murata et al., 2000) und mit Prozessen infolge von Stimulusänderungen (Downar et al., 2000) in Verbindung gebracht. Einige Studien (Saito et al., 2005; Sestieri et al., 2006; Szycik et al., 2008b; van Atteveldt et al., 2007) zeigten, dass der Parietalkortex durch die Verbindung audiovisueller Sprachinformation moduliert wird, besonders wenn durch inkongruente Stimulation zwischen den Modalitäten Antwortkonflikte hervorgerufen werden (Durstun et al., 2003; Ojanen et al., 2005) und zeitlich asynchrone Stimuli perzeptuell ‚synchronisiert‘ wurden (Miller & D'Esposito, 2005).

Patientenstudien ergaben, dass parietale Läsionen zu verminderter Leistung bei Aufgaben in bezug auf räumliche und zeitliche Anordnung führen (Hamilton et al., 2006). Zeitkritische Verarbeitung verstärkte auch bei Affen Antworten in parietalen Zellen (Leon

& Shadlen, 2003). Andere Studien an Affen konnten belegen, dass der Parietalkortex durch Intervalldiskriminationsaufgaben moduliert wird (Onoe et al., 2001) und Regionen innerhalb des Parietalkortex [ventrale intraparietale Region (VIP) und laterale intraparietale Region (LIP)] über multimodale Neurone (Andersen & Buneo, 2002; Duhamel et al., 1998; Ghazanfar & Schroeder, 2006; Lewis & van Essen, 2000; Xing & Andersen, 2000) und Projektionsbahnen zu auditorischen, vestibulären, cerebellären und frontalen Arealen verfügen (Cohen et al., 2002; Rizzolatti et al., 1997). VIP und LIP liegen beim Menschen wahrscheinlich innerhalb des IPS.

Es wird vermutet, dass unimodale Informationsströme hinsichtlich ihrer Zeitlichkeit im Parietalkortex verbunden und verarbeitet (Hamilton et al., 2006) und Stimuli mit adäquaten Antworten im Sinne einer Kategorisierung (Corbetta et al., 2002) bzw. dem Bilden einer perzeptuellen Entscheidung (Heekeren et al., 2008) verbunden werden können. Darüber hinaus weisen die oben aufgeführten Befunde darauf hin, dass der Parietalkortex in intra- und crossmodale Aufmerksamkeitsprozesse, in Zeitverarbeitung und in die Verarbeitung von Stimulusmerkmalen involviert ist. In bezug auf multimodale Verarbeitungsmechanismen formulieren Dhamala et al. (2007) die Hypothese, dass parietale Aktivität mit der Entstehung eines stabilen Asynchronizitätsperzepts in Zusammenhang steht. Dies ist hinsichtlich der Entstehung von kategorisierten Perzepten in Einklang mit anderen hier aufgeführten Befunden und spezifiziert sie gleichzeitig in bezug auf die Relevanz für asynchrone Perzepte. Die hier ermittelte Modulation des linken Parietalkortex durch asynchrone Perzepte untermauert bisherige Studien zur audiovisuellen Integration und könnte im oben formulierten Zusammenhang durch das Entstehen eines manifesten asynchronen Perzepts erklärt werden. Unklar ist jedoch noch, inwieweit diese Ergebnisse von Ergebnissen eines (multimodalen) Aufmerksamkeitsnetzwerks abgegrenzt werden können (Corbetta et al., 1993; 1998; 2000; Mesulam, 1981; 1999).

Ein weiteres Gebiet, das innerhalb der vorliegenden Untersuchung spezifisch durch wahrgenommene Asynchronizität moduliert wurde, war der ACC. Diese Region wird in der Literatur mit Aufgabenschwierigkeit (Corbetta et al., 1991) der Konfliktverarbeitung ablenkender Ereignisse (Botvinick et al., 2001; Kerns et al., 2004; MacDonald et al., 2000) und von aufmerksamkeitsregulierenden Prozessen in bezug auf die Aufgabenrelevanz eines Stimulus (Dreher & Berman, 2002; Weissman et al., 2004) in

Verbindung gebracht. Brown und Braver (2005; 2007) schlagen vor, dass der ACC in die Vorhersage von Fehlerwahrscheinlichkeiten in einem bestimmten Kontext involviert ist, auch wenn kein Konflikt in bezug auf Stimuli oder Antwort vorliegt. Studien zur audiovisuellen Sprachverarbeitung berichteten die Modulation des ACC durch widersprüchliche Informationen in beiden Sinneskanälen (Bishop & Miller, 2008; Miller & D'Esposito, 2005; Ojanen et al., 2005; Szycik et al., 2008b; van Atteveldt et al., 2007).

Es wird davon ausgegangen, dass der ACC den PFC moduliert und dadurch ein Netzwerk beeinflusst, das kognitiv kontrollierende Prozesse steuert (Kerns et al., 2004), und bei Stimuli, die im Widerspruch zu Zielaufgaben stehen andere neuronale Pfade dahingehend beeinflusst, relevante Stimuli, Assoziationen und Antworten bevorzugt zu verarbeiten (Weissman et al., 2004).

Unter Einbeziehung dieser Befunde könnte die vorliegende Modulation des ACC durch asynchrone Wahrnehmung auf das Erkennen von zeitlichen Abweichungen zwischen auditorischen und visuellen Ereignissen und einer Überprüfung des Perzepts gegenüber physikalischer Gleichzeitigkeit zurückgegangen sein. Dieser Vermutung folgend würden synchrone Perzepte einen Ausgangszustand darstellen, der durch erhöhte Fehlerkontrolle verlassen würde, wenn ein asynchrones Perzept entsteht.

Bemerkenswert ist, dass die Modulationen, die mit asynchronen Perzepten korrelierten laut Datenanalyse in fast allen ermittelten Gebieten einheitlich sowohl auf synchrone, als auch auf asynchrone Stimuli beider Ausprägungen zurückgingen. D.h. die Modulationen perzeptueller Asynchronizität überlappten für alle Stimulustypen. Dagegen zeigten die Modulationen synchroner Perzepte bei unterschiedlichen Stimulustypen ein räumlich heterogenes Muster, d.h. sie überlappten nicht im selben Maß wie asynchron wahrgenommene Reize.

Daher lässt sich vermuten, dass Areale, die asynchrone Wahrnehmung verarbeiten weniger auf bestimmte zeitliche Informationskonstellationen externer Reize spezialisiert sind als Areale, die synchrone Perzepte verarbeiten. Dies könnte reflektieren, dass asynchrone Perzepte in geringerem Maße von spezifischen physikalischen Stimulusmerkmalen abhängig sind als synchrone Perzepte. Gegebenenfalls wäre dies ein Hinweis darauf, dass synchrone Perzepte über vermehrten Abgleich mit den tatsächlichen Stimulusmerkmalen kritischer überprüft würden als asynchrone.

Des Weiteren fällt auf, dass an der Verarbeitung asynchroner Perzepte, im Gegensatz zu synchronen Perzepten, weder unisensorische Kortices noch der multimodale STS beteiligt waren. Diese Areale scheinen durch oben bereits erwähnte Verbindungen und Rückkopplungsmechanismen spezifisch an der Entstehung audiovisuell synchroner Perzepte beteiligt zu sein. Dabei ist entscheidend, dass anhand der vorliegenden Ergebnisse physikalische Gleichzeitigkeit nicht ausreichte, um unisensorische Kortices und den multimodalen STS zu modulieren. Dies geschah nur, wenn Probanden Stimuli als synchron empfanden, unabhängig davon, ob der Reiz tatsächlich synchron war oder nicht. Gegebenenfalls existieren Gebiete, die als entscheidende Instanz dieses Netzwerk aus unisensorischen Kortices und multimodalem STS modulieren und damit Impulse für synchrone oder asynchrone Perzepte geben.

Ein Areal, das sowohl durch synchrone, als auch durch asynchrone Perception moduliert wurde, war das Cerebellum. Aktivität in diesem Areal wurde während der Präsentation einfacher synchroner und asynchroner audiovisueller Stimuli beobachtet (Bushara et al., 2001; Calvert et al., 2001), aber auch während unimodaler und audiovisuell kombinierter Sprachstimulation (Belardinelli et al., 2004; Saito et al., 2005; Sekiyama et al., 2003) und visuo-haptischer Integration (Stevenson et al., 2009). Dieses Areal scheint daher für die Verarbeitung multimodaler Inhalte von gewisser Relevanz zu sein, obwohl seine Bedeutung nicht geklärt ist, da andere Studien dieses Gebiet nicht mit multimodaler Integration verbinden (z. B. Dhamala et al., 2007; Miller & D'Esposito, 2005; Noesselt et al., 2007; Szyck et al., 2008a; van Atteveldt et al., 2007). Weitere Studien lieferten jedoch Hinweise darauf, dass das Cerebellum kritisch in die Aufrechterhaltung und Ausführung zeitlich abgestimmter Bewegungen involviert ist (Ivry, 1993; Ivry et al., 1988; Kawashima et al., 2000; Timmann et al., 1999; Topka et al., 1993). Eine andere Untersuchung von Patienten mit cerebellären Läsionen legt nahe, dass das Kleinhirn entscheidend für zeitliche Abstimmungen von Bewegungen ist, wenn ein aufgabenabhängiges Ziel vorliegt (Spencer & Ivry, 2005) und wenn Reize in ihrer Dauer voneinander unterschieden werden sollen (Hetherington et al., 2000; Ivry, 1997). Des Weiteren war das Cerebellum in einer fMRT-Untersuchung an zeitlichen Diskriminationsaufgaben von Stimulusdauer beteiligt (Mathiak et al., 2004). Toplak et al. (2005) schließen, dass das Cerebellum spezifisch in Aufgaben involviert sein kann,

die eine interne Repräsentation feiner zeitlicher Abstimmungen über die Abstände zwischen Ereignissen voraussetzen, auch wenn motorische Antworten ausbleiben (Casini & Ivry, 1999). In Anbetracht dieser Befunde könnten die vorliegenden cerebellären Modulationen durch perzeptuelle Gleichzeitigkeit und Ungleichzeitigkeit audiovisueller Ereignisse interne zeitliche Vergleichsmechanismen darstellen, damit die Zielaufgabe bewerkstelligt werden konnte.

#### 3.1.3.4.2.2 (a) *synchrones Antwortverhalten*

Das Muster kortikaler Aktivität, das durch verhaltensbezogene Effekte von Synchronizität (, die das Antwortverhalten synchroner Wahrnehmung nachempfinden) moduliert wurde, ähnelt dem stimulusunabhängiger synchroner Perzepte und besteht aus temporalen, parietalen, präfrontalen, cerebellären und unisensorisch auditorischen und visuellen Arealen. Es unterscheidet sich hinsichtlich der parietalen Aktivierungen von der stimulusunabhängigen Synchronizitätsperzeption. Einen weiteren Unterschied stellen die oben aufgeführten Modulationen des medialen Temporalkortex dar, der spezifisch durch stimulusunabhängige Synchronizitätsperzeption moduliert wurde: dieses Areal wurde durch verhaltensbezogene BOLD-Muster nicht moduliert. Die durch asynchrones Antwortverhalten modulierte Wahrnehmung zeigt lediglich Effekte im linken STS. Die oben aufgeführten spezifischen Effekte der Asynchronizitätswahrnehmung in Insula und Parietallappen wurden nicht gefunden. Darüber hinaus wurden unisensorische Kortices nur durch verhaltensabhängige Synchronizitätswahrnehmung moduliert.

Die Inbezugnahme der verhaltensabhängigen Effekte in die Analyse der funktionellen Daten spiegelt die Kongruenz zwischen tatsächlichem Ausmaß der zeitlichen Verschiebung zwischen Licht und Ton eines Stimulus und der subjektiven Wahrnehmung dieses Reizes wieder. Areale, die durch diesen Kontrast moduliert wurden, sollten folglich umso mehr Aktivität aufweisen, je besser die korrekte Zeitwahrnehmung der Stimuli gelang. Demzufolge ging die adäquate Detektion synchroner Stimuli mit mehr kortikaler Aktivität in mehr Gebieten einher, als die korrekte Identifizierung asynchroner Reize. Dies könnte bedeuten, dass die korrekte Ablehnung asynchroner Reize leichter gelingt als die eindeutige Zuordnung synchroner Reize zu einem adäquaten Perzept. Dies wiederum könnte die bevorzugte ‚Auswertung‘

---

multimodal synchroner Ereignisse widerspiegeln, die unter Inbezugnahme unisensorischer und multimodaler, aber auch kontrollierender Areale, wie dem PFC, zur sorgfältigen Verarbeitung der Informationen aller Sinneskanäle führt. Korrekt wahrgenommene asynchrone Perzepte würden dagegen im Sinne eines Effizienzkriteriums unter wenig neuronalem Aufwand vom System ignoriert bzw. von der Weiterverarbeitung im Sinne einer Integration ausgeschlossen.

Eine fragliche Rolle spielt der linke Parietalkortex, der sowohl durch stimulusunabhängige Asynchronizitätswahrnehmung, als auch durch kongruent wahrgenommene Synchronizität moduliert wurde. In anderen Studien zur audiovisuellen Zeitwahrnehmung war dieses Gebiet am ‚Verschmelzen‘ asynchroner Information (Miller & D’Esposito, 2005) und am Entstehen stabiler asynchroner Perzepte beteiligt (Dhamala et al., 2007). In diesem Zusammenhang könnte dieses Areal sowohl über synchrone als auch asynchrone Referenzen verfügen und anhand eines Vergleichs der eintreffenden Informationen mit verfügbaren Wahrnehmungsoptionen über die Gültigkeit eines perzeptuellen Zustands entscheiden. Gegebenenfalls moduliert der Parietalkortex das Netzwerk aus unisensorischen Kortices und multimodalem STS, um synchrone oder asynchrone Perzepte zu fördern. Dass dieses Gebiet durch verhaltensabhängige Synchronizitätswahrnehmung moduliert wurde, spiegelt seine Relevanz in bezug auf eine Aufgabenorientierung wider. Dies unterstreicht, gemäß den oben aufgeführten Befunden, seine Rolle als Region, die durch aufmerksamsbezogene Prozesse, seine Sensibilität für Stimulusänderungen und seine Beteiligung an der Modulation von Antwortkonflikten kategorisierende Prozesse in Abhängigkeit einer aktuellen Aufgabe regulieren kann (Corbetta et al., 2002; Hamilton et al., 2006; Heekeren et al., 2008).

#### 3.1.3.4.2.3 Zusammenfassung

Insgesamt zeigte die vorliegende Studie differenziert die Beteiligung zuvor definierter unisensorischer und potentieller bimodaler Kortices und konnte zwischen rein stimulusbezogenen (physikalischen), wahrnehmungsbezogenen (perzeptuellen) und antwortverhaltensabhängigen BOLD-Mustern unterscheiden. Diese Kombination aus Experimentaldesign und Auswertungsverfahren ist aus bisher bestehenden Studien nicht bekannt.

Dabei wurden distinkte kortikale Aktivierungen für physikalische und perzeptuelle Effekte identifiziert. Sowohl synchrone, als auch asynchrone Reize modulierten unabhängig vom Perzept den linken bimodalen STS, wogegen sich, sobald ein stimulusunabhängiges Perzept entstanden war, spezifische Modulationen für Synchronizität bzw. Asynchronizität ergaben. Dabei schien ein Netz aus bimodalem STS und unisensorischen Kortices inklusive medial temporaler Areale synchrone Perzepte bevorzugt zu bearbeiten (s. auch Driver & Noesselt, 2008; Noesselt et al., 2007) und eines bestehend aus anteriorer Insula, ACC und linkem Parietalkortex asynchrone Perzepte. Der PFC ergänzte diese kortikalen Modulationen unabhängig von der wahrgenommenen Zeitlichkeit der Stimuli. Die Analyse der antwortverhaltensmodulierten Prozesse ergab, dass Synchronizität ein breiteres neuronales Netz aktivierte als verhaltensmodulierte Asynchronizität. Dabei wurde auf die spezifische Rolle des Parietalkortex hingewiesen, der an der Kategorisierung audiovisueller Informationsströme hinsichtlich ihrer zeitlichen Relationen zueinander beteiligt sein könnte.

Es wurde anhand bereits bestehender Befunde aufgezeigt, dass die hier modulierten Areale über direkte Verbindungen untereinander verfügen und teilweise über Rückkopplungsmechanismen miteinander kommunizieren können. Es kann daher vermutet werden, dass es sich um kortikale Netze handelt, die unter anderem multimodale Zeitinformationen verarbeiten können. Dabei werden synchrone Reize offenbar anders als asynchrone verarbeitet. Dies könnte auf eine Präferenz des Verarbeitungssystems für zeitliche Koinzidenz hindeuten. Asynchrone Perzepte wurden tendenziell in Arealen verarbeitet, die aus anderen Studien auch durch Stimulusabweichungen und Fehler moduliert wurden, während synchrone Perzepte durch gleichzeitige Aktivität unisensorischer Bereiche und multimodalem STS verarbeitet wurden. Dies könnte bedeuten, dass erkannte Asynchronizität als ungeeignet zur Weiterverarbeitung klassifiziert wurde und audiovisuelle Informationsströme über unisensorische Beteiligung und Rückkopplung zu bimodalen Arealen einander angeglichen wurden. Dabei könnte der mediale Temporalkortex für die Aufrechterhaltung eines stabilen Synchronizitätsperzepts, und die anteriore Insula bzw. der Parietalkortex für die eines stabilen Asynchronizitätsperzepts von Bedeutung sein. Dhamala et al. (2007) entwickelten die Vermutung, dass spezifische Areale

synchrone bzw. asynchrone Wahrnehmungszustände modulieren. Ihre Befunde weisen auf die Beteiligung des Parietalkortex an asynchronen und die der SC an synchronen Zuständen hin. Der insuläre und der mediale Temporalkortex wurden von ihnen nicht untersucht. Der Bereich der SC, der laut Dhamala et al. (2007) durch stabile synchrone Perzepte moduliert wird, lag in der vorliegenden Studie nicht innerhalb der verwendeten unisensorischen oder bimodalen Masken. Dies könnte die zu Dhamala et al. (2007) unterschiedlichen Befunde erklären. Des Weiteren wurde der PFC mit Ausnahme stimulusunabhängiger Asynchronizitätswahrnehmung vorwiegend durch physikalische oder perzeptuelle Synchronizität moduliert. Es wurde oben besprochen, dass dies dahingehend gedeutet werden könnte, dass synchrone Stimulation im Vergleich zu asynchroner erhöhte kognitive Regulation erfordert, die gegebenenfalls die Weiterverarbeitung audiovisuell synchroner Reize anzeigt, die beim Eintreffen bimodal asynchroner Informationen unterbunden würde.

Die vorliegenden Befunde untermauern existierende Befunde, die multimodale kortikale Netzwerke beschreiben, die die Zeitlichkeit zweier Informationsströme verarbeiten und dabei synchrone und asynchrone Reize voneinander differenzieren (s. oben). Diese Ergebnisse konnten durch die vorliegende Studie dahingehend erweitert werden, dass die bekannten kortikalen Netzwerke vorwiegend perzeptuellen, aber auch antwortverhaltensbezogenen Effekten zugeordnet werden konnten, dass Effekte eindeutig in unisensorische und bimodale Regionen unterteilt werden konnten und dass Areale, die zuvor oft durch audiovisuelle Sprachreize moduliert wurden auch anhand einfacher, semantisch nicht belegter Reize identifiziert werden konnten.

## 3.2. Experiment 2

### 3.2.1 Einführung

Nachdem mit Hilfe von Experiment 1 sowohl die optimale SAPS zur Untersuchung audiovisueller Gleichzeitigkeit determiniert, als auch kortikale Gebiete identifiziert wurden, die durch die Darbietung phasengleicher bzw. phasenverschobener audiovisueller Sequenzen und Gleichzeitigkeits- bzw. Ungleichzeitigkeitsperzepte moduliert werden, sollen in Experiment 2 die neuronalen Grundlagen adaptiver Prozesse in bezug auf phasengleiche bzw. ungleiche Reize untersucht werden. Diese Untersuchung

besteht aus drei psychophysischen Vorexperimenten und einem fMRT-Experiment. Das Hauptexperiment hat zum Ziel, die Probanden auf ein bestimmtes Verzögerungsintervall zwischen Licht und Ton zeitlich zu adaptieren. Im Anschluss an die Adaptationsphase sollen synchrone und asynchrone Sequenzen bewertet werden. Die Wahrnehmung der nachfolgenden Testsequenzen soll in Abhängigkeit der vorangegangenen Darbietung zeitverzögerter Stimuli beeinflusst werden. Bei ausreichender Adaptation auf zeitliche Verzögerungen zwischen Licht und Ton, sollte sich der Punkt der subjektiven Gleichzeitigkeit (PSS) und/oder der gerade noch wahrgenommene Unterschied (JND) je nach Verzögerungsintervall der durchgeführten Adaptation verändern (PSS: z. B. Fujisaki et al., 2004; Vatakis et al., 2008; Vroomen et al., 2004; JND: z. B. Navarra et al., 2005) und so Hinweise auf eine Rekalibrierung des audiovisuellen Zeitsystems geben. Wie bereits erwähnt, kann die JND als ein Zeitfenster verstanden werden, innerhalb dessen die Integration von Ton- und Lichtreiz gelingt und als gleichzeitig wahrgenommen wird. Ein enges Zeitfenster bzw. eine kleine JND spiegelt erhöhte Sensitivität für Asynchronizität wider. Eine Veränderung des PSS kann als Hinweis auf eine Verschiebung der subjektiven Zeitkalibrierung in bezug auf die crossmodale Integration gedeutet werden (z. B. Navarra et al., 2005; Spence et al., 2001; Spence & Squire, 2003; Stone et al., 2001; Vatakis et al., 2007; 2008; Vroomen et al., 2004).

Im Gegensatz zu Experiment 1, wo durch bloße Stimulusdarbietung und -wahrnehmung die Perzeption einfacher Stimuli analysiert wurde, soll eine Adaptationsstudie Aufschluss darüber geben, wie bzw. wo im Gehirn stimulusbezogene Verarbeitung zeitlich verzögerter audiovisueller Stimuli von wahrnehmungsbezogener zu trennen ist. Falls Areale identifiziert werden können, die über die Dauer der Exposition mit audiovisuell verzögerten Stimuli ihre Aktivierungsmuster verändern, deutet dies darauf hin, dass sie bestimmte Zeitverzögerungen kompensieren und damit das Beurteilungs- bzw. Integrationsverhalten von Probanden gegenüber audiovisuellen Reizen beeinflussen. Dies würde auf stimuluspezifische kortikale Mechanismen hinweisen, die audiovisuelle Integrationsprozesse verändern können. Diese Daten sollen mit den Befunden rein perzeptueller Mechanismen audiovisueller Verarbeitung aus Experiment 1 abgeglichen werden und so Aufschluss über hierarchische Prozesse audiovisueller Zeitverarbeitung geben.

### 3.2.2 Fragestellungen und Hypothesen zu den Voruntersuchungen

Um relevante Parameter für den Experimentalaufbau der fMRT-Untersuchung zu bestimmen, wurden drei Vorexperimente durchgeführt, anhand derer ermittelt wurde,

- (1) wie lange zeitverzögerte Stimuli dargeboten werden müssen, damit die Sensitivität von Probanden gegenüber audiovisueller Asynchronizität nachweisbar beeinflusst wird, bzw. Adaptation auf audiovisuelle Zeitintervalle gelingt,
- (2) ob das Beurteilungsverhalten durch erneute kurze „Re-Adaptation“ während der Testphase gegenüber bloßer Abfrage ohne erneute Adaptation verändert wird, bzw. Probanden größere oder geringere Sensitivität gegenüber Asynchronizität zeigen, oder Adaptation auf audiovisuelle Zeitintervalle länger aufrechterhalten wird, oder vielmehr gestört wird, und
- (3) wie sich eine Veränderung der Länge der Abfragestimuli auf das Beurteilungsverhalten auswirkt.

Aus dem oben Aufgeführten ergeben sich drei Fragestellungen, die anhand dreier Vorexperimente exploriert werden sollen:

1. Hat die Dauer, mit der Probanden zeitverzögerten einfachen audiovisuellen Stimuli ausgesetzt werden einen Einfluss auf die Synchronizitätsurteile nachfolgender Teststimuli der Probanden bzw. deren JNDs und PsSS?
2. Hat die „Re-Adaptation“ während des Abfragens mittels Teststimuli einen Einfluss auf die Wahrnehmung von Probanden bzw. deren JNDs und PsSS?
3. Hat die Länge der Teststimuli nach vorhergehender Exposition einen Einfluss auf das Beurteilungsverhalten von Probanden bzw. deren JNDs und PsSS?

Für das fMRT-Experiment soll diejenige Stimulationskonstellation gewählt werden, bei der die Beeinflussung der audiovisuellen Zeitwahrnehmung durch Adaptation auf bestimmte Zeitverzögerungen zwischen Licht und Ton nachweisbar gelingt. Das bedeutet, dass diejenige Darbietungsdauer gewählt werden soll, bei der sich entweder JNDs oder PsSS in Abhängigkeit der zuvor dargebotenen Zeitintervalle ändern, dass „Re-Adaptations-Ereignisse“ eingeführt werden, falls durch diese PSS oder JND über die Verzögerungsbedingungen der Exposition hinweg signifikant beeinflusst werden und dass diejenige Länge der Abfragestimuli gewählt werden soll, bei der eine systematische Veränderung von PSS und JND in Abhängigkeit der zeitlichen Verzögerungen während der Expositionsphase nachgewiesen werden kann.

### 3.2.3 Vorexperiment 2.1

Das erste Vorexperiment 2.1 zielt darauf ab, zu erfassen, mit welcher Dauer audiovisuelle Verzögerungen zwischen Licht und Ton anhand einfacher Stimuli dargeboten werden müssen, um bei späteren Abfragestimuli Veränderungen im Beurteilungsverhalten gegenüber Synchronizität von audiovisuellen einfachen Stimuli bei Probanden hervorzurufen.

#### 3.2.3.1 Methodik

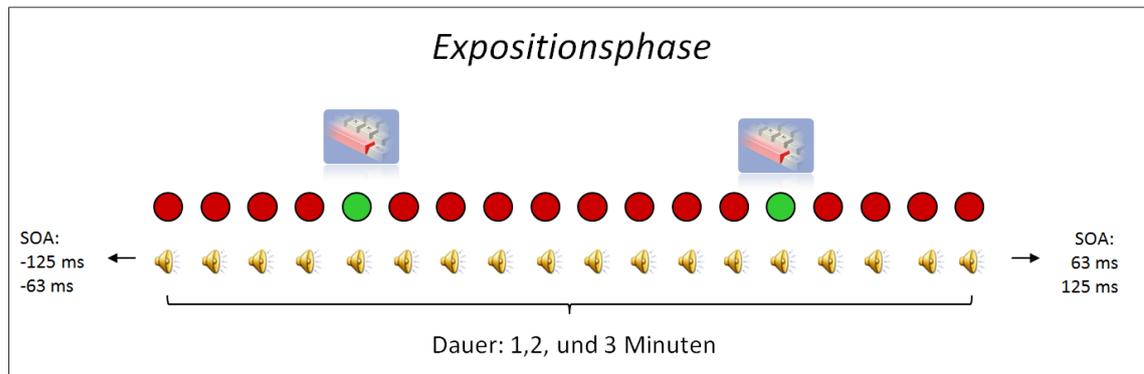
Um die oben aufgeführte Fragestellung experimentell zu untersuchen, wurden drei Untersuchungen durchgeführt. In jeder wurden Probanden instruiert, phasisch verschobene, audiovisuelle einfache Stimuli hinsichtlich ihrer Gleichzeitigkeit zu bewerten (s. Experiment 1, Kapitel 3.1.3.1). Dieses Vorexperiment manipulierte, wie lange in Folge Probanden in der Expositionsphase denselben einfachen, zeitlich verzögerten audiovisuellen Stimuli ausgesetzt wurden, die sie in der Testphase bewerten sollten. Ein Telexperiment wurde mit einer Expositionsphase von einer Minute Länge durchgeführt, eins mit einer Länge von zwei Minuten und ein drittes mit einer Länge von drei Minuten. Wie in den vorangegangenen Experimenten konnten Ton und Licht der Expositionsphase 180° und 90° jeweils in Richtung des Licht- oder Tonsignals oder 0° verschoben sein. Dies entspricht einer Verzögerung von 125 ms, 63 ms (jeweils Ton zuerst und Licht zuerst) und 0 ms (synchron). Probanden wurden dementsprechend fünf verschiedenen zeitlichen

Verzögerungen zwischen Licht und Ton ausgesetzt. Die Expositionsphase bestand aus einer Aneinanderreihung audiovisueller Ereignisse, die mit einer Frequenz von 4 Hz präsentiert wurden (s. Abb. 20). Die visuellen Stimuli bestanden aus roten, grünen und weißen LEDs. Ein weißer Fixationslichtpunkt wurde auf Augenhöhe angebracht und das rote Ziellicht 5° Sehwinkel (= 4,7 cm) darüber (60 cd/m<sup>2</sup>). Jedes auditorische Ereignis war ein deutlich hörbarer Ton, der durch den Lautsprecher (Durchmesser 2,5 cm) direkt unter dem „Ziellicht“ bei 500 Hz und 65 dB (SPL, vom Ohr der Probanden aus gemessen) präsentiert wurde (s. Abb. 8).

Licht- und Tonsignal wurden jeweils mit einer Dauer von 30 ms dargeboten. Während der Expositionsphasen wurden zwei Mal pro Minute grüne statt rote Lichtreize dargeboten, die die Probanden per Knopfdruck entdecken sollten. Diese sogenannten „catch-trials“ wurden verwendet, um die Aufmerksamkeit der Probanden während der Expositionsphase auf die relevanten Stimuli zu lenken. Nach der Expositionsphase erlosch das Fixationslicht für 1000 ms, um Probanden einen Aufgabenwechsel zu signalisieren, und es wurden anschließend 75 Testsequenzen, die aus vier aneinandergereihten audiovisuellen Ereignissen bestanden präsentiert und die die Probanden per Knopfdruck als synchron oder asynchron beurteilen sollten (Synchronizitätsurteil).

Die Testsequenzen konnten, genau wie die Expositionsstimuli, um 180°, 90° oder 0° phasisch gegeneinander in beide Richtungen verschoben sein, was einer zeitlichen Verzögerung zwischen Licht- und Tonreiz von +/- 125 ms, +/- 63 ms oder 0 ms entspricht. Bei dem hier vorliegenden Experiment handelt es sich folglich um ein 5 x 5 – Design mit fünf in ihrer zeitlichen Verzögerung verschiedenen Expositionsphasen und fünf unterschiedlich zeitlich verzögerten Teststimuli. Die Zeit zwischen zwei Stimulussequenzen (ISI) betrug durchschnittlich 1000 ms. Stimuli konnten zufällig in einem Bereich von 500 ms bis 1500 ms nach dem vorhergehenden Stimulus dargeboten werden (Jittering). Dieser Zeitraum konnte von den Probanden für das Bewerten einer Testsequenz genutzt werden. Insgesamt wurden 15 Durchgänge (und Expositionsphasen) à 75 Testsequenzen dargeboten, so dass jede der fünf Verzögerungsbedingungen während der Expositionsphasen drei Mal und jede Verzögerungsbedingung der Testsequenzen 225 Mal präsentiert werden konnte. Pro Verzögerungsbedingung der Exposition konnte je eine Verzögerungsbedingung der Testsequenzen 45 Mal gezeigt

werden (s. Tab. 1). Verzögerungsbedingungen der Testsequenzen und Expositionsphasen wurden randomisiert dargeboten.



**Abbildung 20.** Stylisierte Darstellung der Expositionsphase von Vorexperiment 2.1. Audiovisuelle Sequenzen, bestehend aus aneinandergereihten roten Lichtblitzen (rote Punkte) und Tonsignalen (Lautsprecher; Licht- und Tonsignale jeweils 30 ms lang), wurden mit einer Frequenz von 4 Hz phasengleich (synchron) oder phasenverschoben (asynchron) in fünf Experimentalbedingungen [SOA: -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran), 125 ms (Licht voran)] dargeboten. Pro Minute wurden anstatt roter zwei grüne Lichtsignale präsentiert, die per Tastendruck der Leertaste detektiert werden sollten (s. Text zur Erläuterung). Die Expositionsphase dauerte in Vorexperiment 2.1.1 eine Minute, in Vorexperiment 2.1.2 zwei Minuten und in Vorexperiment 2.1.3 drei Minuten an. Für eine Darstellung der Testsequenzen s. Abb. 13.

Probanden wurden instruiert, eine Taste auf einer Computertastatur zu drücken, wenn ein Stimulus ihnen den subjektiven Eindruck zeitlicher Synchronizität vermittelte und eine zweite Taste, um einen Reiz als asynchron zu bewerten. Sie wurden gebeten, einen Stimulus nach seiner Darbietung zu bewerten, nicht aber während einer Sequenz und nicht, nachdem die nächste Sequenz bereits präsentiert würde. Vor jedem Experiment wurde ein Übungsdurchlauf durchgeführt, während dem alle vorkommenden Stimuluskombinationen und eine verkürzte Variante der Expositionsphase dargeboten wurde, die bereits beurteilt werden sollten. Wie auch in Experiment 1 gab der Versuchsleiter während des Übungsdurchgangs Rückmeldung über korrekte bzw. nicht korrekte Urteile. Nachdem die Probanden synchrone und asynchrone Reize jeweils mindestens zweimal korrekt detektiert hatten, wurde mit dem Hauptexperiment begonnen, andernfalls wurde der Übungsdurchgang wiederholt.

**Tabelle 1.** Tabellarische Anordnung der Durchgänge und Experimentalbedingungen. Es wurden 15 Stimuli pro Durchgang (Spalten) und SOA der Testsequenzen (Zeilen) dargeboten. Insgesamt wurde diese Anordnung 15 Mal wiederholt, wobei sich jedes SOA der Expositionsbedingung drei Mal in randomisierter Reihenfolge wiederholte. Durchgänge sind in der Tabelle aus Gründen der Übersicht bestimmten SOAs der Expositionsbedingung zugeordnet. Die Abfolge verschiedener Expositionsbedingungen pro Durchgang war zufällig und unabhängig von der Reihenfolge der Durchgänge.

		Exposition															
Durchgang		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
SOA (ms)		-125	-125	-125	-63	-63	-63	0	0	0	63	63	63	125	125	125	
Testsequenzen	-125	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	225
	-63	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	225
	0	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	225
	63	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	225
	125	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	225
		75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	

Am Vorexperiment 2.1 mit einer Expositionsdauer von einer Minute nahmen 28 rechtshändige (Selbstbericht) Probanden (Durchschnittsalter 23,9; Spanne 20-30; 18 weiblich) teil, am Vorexperiment 2.1 mit zwei Minuten Exposition 32 (Durchschnittsalter 22,3; Spanne 19-27; 20 weiblich) und am Vorexperiment 2.1 mit drei Minuten Expositionsdauer 39 (Durchschnittsalter 22,6; Spanne 21-30; 20 weiblich). Alle Teilnehmer berichteten, normales Hörvermögen und normales oder korrigiertes Sehvermögen und keine psychiatrischen oder neurologischen Störungen zu haben. Sie variierten hinsichtlich ihrer Vorerfahrung mit der Teilnahme an psychophysikalischen Studien und wurden über deren Zweck aufgeklärt. Vorexperiment 2.1 bei einer Minute Exposition dauerte 75, bei zweiminütiger Exposition 90, und bei dreiminütiger 105 Minuten und wurde in einem vollständig dunklen und schallisolierten Raum durchgeführt.

### 3.2.3.2 Ergebnisse des Vorexperiments 2.1

#### 3.2.3.2.1 Synchronizitätsurteile

Mit den Synchronizitätsurteilen aller drei Vorexperimente (99 Probanden) wurde eine 5 x 5-faktorielle Varianzanalyse mit Messwiederholung für die Faktoren *Zeitverzögerung der Exposition* und *Zeitverzögerung der Testsequenzen* durchgeführt (SPSS

15.0). Die fünf Faktorstufen stellen jeweils die fünf Zeitverzögerungen dar, um die Lichtgegen Tonreize verschoben wurden. Zusätzlich wurde der Faktor *Expositionsdauer* als Zwischen-Subjekt-Faktor analysiert. Probanden, die während der Expositionsphase weniger als 80% der „catch-trials“ entdeckt hatten und mehr als 20% falsch positive Antworten gaben, wurden aus der Analyse entfernt. Denn bei diesen Probanden ist davon auszugehen, dass sie die Stimulation der Expositionsphase nicht zulänglich beachteten. Somit war es unklar, ob die gewünschte Adaptation auf bestimmte audiovisuelle Verzögerungen auftreten würde. Antworten, die während der Darbietung bis 200 ms nach Ende einer Testsequenz gegeben wurde, flossen nicht in die Analyse ein.

Probanden gaben durchschnittlich über alle drei Varianten des Vorexperiments 2.1 hinweg in 53% (2,5%), 74% (1,8%), 83% (1,5%), 84% (1,5%) und 73% (2,1%) aller Fälle ein Synchronizitätsurteil ab, wenn Licht und Ton der Teststimuli um -125 ms (Ton zuerst), -63 ms (Ton zuerst), 0 ms (synchron), 63 ms (Licht zuerst) und 125 ms (Licht zuerst) verzögert waren (s. Tab. 2 und Anhang, Abb. 43 -45). Nach Exposition bei um -125 ms (Ton zuerst) verzögerten audiovisuellen Stimuli wurden 73 % (1,4%), nach -63 ms (Ton zuerst) 73% (1,4%), nach 0 ms (synchron) 74 % (1,5%), nach 63 ms (Licht zuerst) 72% (1,5%) und nach 125 ms (Licht zuerst) 74% (1,4%) synchrone Beurteilungen abgegeben. Nach einminütiger Exposition wurden 74 % (2,7%), nach zweiminütiger 73% (2,3%) und nach dreiminütiger ebenfalls 74 % (2,5%) Synchron-Urteile abgegeben (Standardfehler in Klammern, s. Tab. 2 für alle ermittelten Werte und Anhang, Abbildungen 24 -26). Es wurden Haupteffekte der Faktoren *Zeitverzögerung der Exposition* [ $F(4;360) = 2,4$ ;  $p < 0,05$ ;  $\eta^2 = 0,03$ ] und *Zeitverzögerung der Testsequenzen* [ $F(1,8;160,3) = 93,65$ ;  $p < 0,001$ ;  $\eta^2 = 0,51$ ] ermittelt und ein Interaktionseffekt zwischen *Zeitverzögerung der Exposition* und *Expositionsdauer* [ $F(8;360) = 2,2$ ;  $p < 0,05$ ;  $\eta^2 = 0,05$ ]. Die Analyse von *Expositionsdauer* ergab keinen signifikanten Effekt [ $F(2;90) = 0,14$ ;  $p = 0,99$ ;  $\eta^2 = 0$ ]. Interaktionseffekte aus allen drei Faktoren traten nicht auf [ $F(23;1039) = 0,57$ ;  $p = 0,95$ ;  $\eta^2 = 0,01$ ]. Die Greenhouse-Geisser-Korrektur wurde angewandt, falls keine Sphärizität nach Mauchly vorlag. Post-Hoc T-Tests des Faktors *Zeitverzögerung der Exposition* zeigten, dass signifikant mehr Synchronizitätsurteile nach Exposition bei 125 ms Verzögerung (Licht zuerst) gemacht wurden als nach Exposition bei 63 ms Verzögerung (Licht zuerst;  $p < 0,05$ ) und -63 ms (Ton zuerst;  $p < 0,05$ ). Post-Hoc-Analysen des Faktors *Zeitverzögerung der Testsequenzen* ergaben, dass signifikant mehr Synchronizitätsurteile gemacht wurden für Stimuli, die

um 0 und 63 ms (Licht zuerst) gegeneinander verschoben waren ( $p < 0,05$ ) als für solche, die mit größerer Verzögerung präsentiert wurden. Für Stimuli, die um 63 ms (Licht zuerst) und 0 ms verzögert waren, wurden gleich viele ‚Synchron‘-urteile abgegeben. Darüber hinaus unterschieden sich die Urteile nicht zwischen Stimuli, die um 125 ms (Licht zuerst) und -63 ms (Ton zuerst) verzögert waren ( $p = 0,31$ ). Da die meisten Synchronizitätsurteile nach Exposition bei um 63 ms (Licht zuerst, 84%) und 0 ms (83%) verzögerten Stimuli gemacht wurden und sie nach Exposition mit Stimuli größerer Verzögerungen geringer waren, weist dies darauf hin, dass alle Verteilungen synchroner Antworten, wie bei Vorexperiment 1 (vgl. 3.1.2.3), leicht asymmetrisch linksschief verteilt sind.

**Tabelle 2.** Anteile synchroner Antworten je SOA der Expositionsbedingung, Expositionsdauer und SOA der Testsequenzen. Standardfehler sind in Klammern angegeben. Die äußere Spalte stellt die gemittelten Werte über alle SOAs der Testsequenzen dar, die unterste Zeile die gemittelten Werte kollabiert über SOA der Expositionsbedingung und Expositionsdauer. S. auch Anhang, Abbildungen 43-45.

		<i>Testsequenz</i>						<i>Mittel</i>
		<i>SOA</i>	<b>-125 ms</b>	<b>-63 ms</b>	<b>0 ms</b>	<b>63 ms</b>	<b>125 ms</b>	
<i>Exposition</i>	<i>1 Minute</i>	<b>-125 ms</b>	0,49 (0,04)	0,75 (0,03)	0,85 (0,03)	0,85 (0,03)	0,73 (0,04)	0,73
		<b>-63 ms</b>	0,49 (0,05)	0,75 (0,03)	0,87 (0,03)	0,87 (0,03)	0,72 (0,04)	0,74
		<b>0 ms</b>	0,50 (0,05)	0,74 (0,04)	0,85 (0,03)	0,85 (0,03)	0,72 (0,04)	0,73
		<b>63 ms</b>	0,48 (0,05)	0,75 (0,04)	0,85 (0,03)	0,85 (0,03)	0,71 (0,04)	0,73
		<b>125 ms</b>	0,49 (0,05)	0,76 (0,03)	0,86 (0,03)	0,88 (0,03)	0,73 (0,04)	0,74
	<i>2 Minuten</i>	<b>-125 ms</b>	0,56 (0,04)	0,74 (0,03)	0,82 (0,03)	0,81 (0,03)	0,71 (0,03)	0,73
		<b>-63 ms</b>	0,55 (0,04)	0,75 (0,03)	0,81 (0,03)	0,81 (0,03)	0,70 (0,04)	0,72
		<b>0 ms</b>	0,55 (0,05)	0,76 (0,03)	0,82 (0,03)	0,82 (0,03)	0,72 (0,04)	0,73
		<b>63 ms</b>	0,55 (0,05)	0,74 (0,03)	0,82 (0,03)	0,81 (0,03)	0,70 (0,04)	0,72
		<b>125 ms</b>	0,57 (0,04)	0,77 (0,03)	0,82 (0,03)	0,83 (0,03)	0,70 (0,04)	0,74
	<i>3 Minuten</i>	<b>-125 ms</b>	0,53 (0,04)	0,73 (0,03)	0,84 (0,02)	0,85 (0,02)	0,75 (0,03)	0,74
		<b>-63 ms</b>	0,52 (0,04)	0,72 (0,03)	0,81 (0,02)	0,85 (0,02)	0,72 (0,03)	0,72
		<b>0 ms</b>	0,57 (0,04)	0,74 (0,03)	0,83 (0,03)	0,87 (0,02)	0,78 (0,04)	0,76
		<b>63 ms</b>	0,52 (0,04)	0,72 (0,03)	0,81 (0,02)	0,84 (0,03)	0,73 (0,03)	0,72
		<b>125 ms</b>	0,54 (0,04)	0,75 (0,03)	0,84 (0,02)	0,85 (0,02)	0,76 (0,04)	0,75
<i>Mittel</i>		0,53	0,74	0,83	0,84	0,72		

Um zu klären, auf welche Expositionsdauer der oben aufgeführte Interaktionseffekt zurückgeht, wurden 5 x 5-stufige Varianzanalysen einzeln für jedes Experiment bzw. jede Expositionsdauer mit den Faktoren *Zeitverzögerung der Exposition* und *Zeitverzögerung der Testsequenzen* durchgeführt. Die Analysen ergaben keine Effekte

des Faktors *Zeitverzögerung der Exposition* für Vorexperimente 2.1.1. (Expositionsdauer: eine Minute;  $F(4;92) = 1,2$ ,  $p = 0,32$ ;  $\eta^2 = 0,05$ ) und 2.1.2. (Expositionsdauer: zwei Minuten;  $F(4;120) = 0,40$ ;  $p = 0,81$ ;  $\eta^2 = 0,01$ ), jedoch für Experiment 2.1.3. (Expositionsdauer: drei Minuten;  $F(4;148) = 5,72$ ;  $p < 0,001$ ;  $\eta^2 = 0,13$ ). Post-Hoc-Analysen des Faktors *Zeitverzögerung der Exposition* nach dreiminütiger Exposition ergaben, dass keine Unterschiede zwischen Urteilen nach Exposition bei -63 ms (Ton zuerst) und 63 ms (Licht zuerst;  $p = 0,98$ ), nach -125 ms (Ton zuerst) und 125 ms (Licht zuerst;  $p = 0,49$ ) und nach 0 ms und 125 ms (Licht zuerst,  $p = 0,23$ ) Verzögerung vorlagen. Die Anzahl der Synchronizitätsurteile nach Exposition bei allen asynchronen Verzögerungen war signifikant geringer als nach synchroner Exposition ( $p < 0,05$ ). Für den Faktor *Zeitverzögerung der Testsequenzen* wurden für jede Expositionsdauer signifikante Effekte ermittelt (eine Minute:  $F(1,6;36,5) = 28$ ;  $p < 0,001$ ;  $\eta^2 = 0,55$ ; zwei Minuten:  $F(1,6;49,3) = 26,9$ ;  $p < 0,001$ ;  $\eta^2 = 0,47$ ; drei Minuten:  $F(1,9;69,9) = 40,2$ ;  $p < 0,001$ ;  $\eta^2 = 0,52$ ).

Post-Hoc-Tests zeigten, dass sich Synchronizitätsurteile für Testsequenzen bei -63 ms (Ton zuerst) Verzögerung zwischen Licht und Ton und bei 125 ms (Licht zuerst) für keine der drei Expositionsängen voneinander unterschieden ( $p > 0,10$ ). Außerdem ergaben sich nach einer und zwei Minuten Expositionsdauer keine Unterschiede zwischen den Beurteilungen der Testsequenzen, die 0 ms und 63 ms (Ton zuerst) gegeneinander verschoben waren ( $p > 0,65$ ). Nach drei Minuten Expositionslänge waren die Synchronizitätsurteile für um 63 ms (Licht zuerst) verschobene Teststimuli jedoch häufiger (85%) als für synchrone Stimuli ( $p < 0,05$ , 82%). Synchronizitätsurteile für alle anderen Stimulusarten [-125 ms (Ton zuerst), -63 ms (Ton zuerst), 125 ms (Licht zuerst)] wurden nach allen drei Expositionsängen signifikant weniger gemacht ( $p < 0,01$ ) als für um 63 ms (Licht zuerst) und um 0 ms (synchron) verschobene Stimuli.

### 3.2.3.2.2 JNDs und PsSS

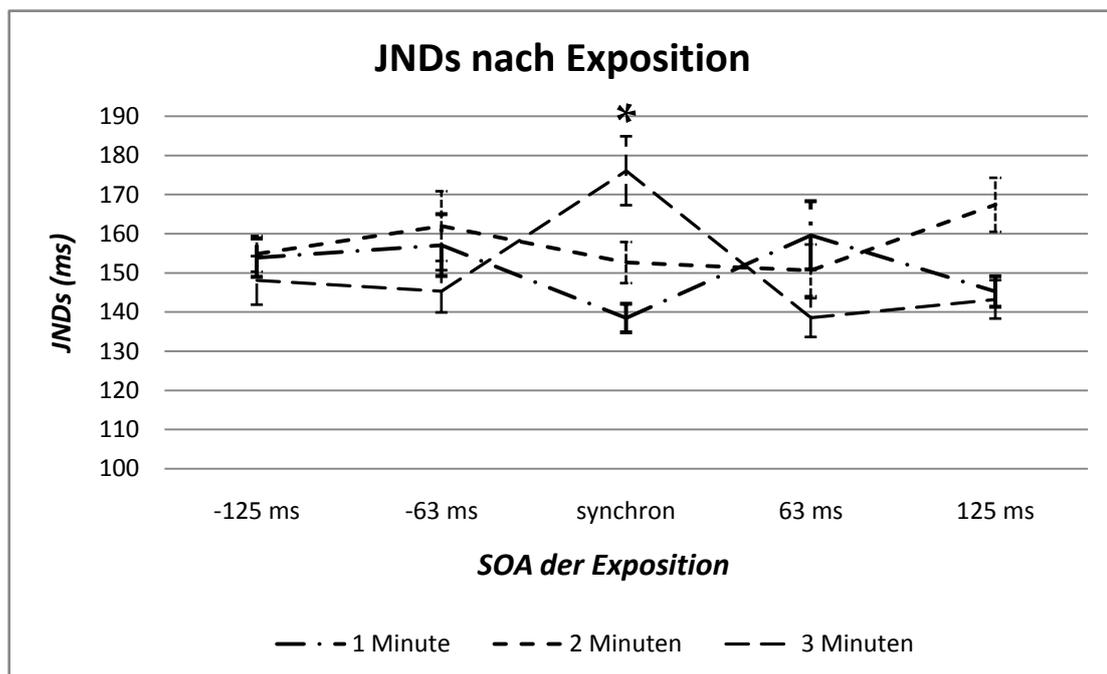
Aus den Synchronizitätsurteilen wurden PsSS und JNDs für die fünf verschiedenen Zeitverzögerungen geschätzt, denen Probanden eine, zwei oder drei Minuten ausgesetzt wurden (s. Kapitel 2.3.1). Mit den Daten aller drei Vorexperimente wurden getrennt für PsSS und JNDs zwei 5-faktorielle Varianzanalysen mit Messwiederholung für den Faktor *Zeitverzögerung der Exposition* durchgeführt. Zusätzlich wurde auch hier der Zwischen-Subjekt-Faktor *Expositionsdauer* eingefügt. Aus den

geschätzten Werten für PsSS, JNDs und Kurvenmaxima wurden standardisierte Z-Werte berechnet und alle Probanden, deren Datenpunkte zwei Standardabweichungen über- oder unterschritten, eliminiert. Dies verhindert, dass unzuverlässige Schätzungen aufgrund willkürlichen Antwortverhaltens in die Datenanalyse einfließen (s. Spence et al., 2001; Vatakis et al., 2007). Die oben besprochenen Daten der Anteile synchroner Urteile bieten diesbezüglich kein Kriterium, so dass unzulässige Fälle erst im Rahmen der Transformation in JNDs bzw. PsSS geschehen kann.

Die Analyse des Faktors *Zeitverzögerung der Exposition* der JNDs ergab keinen signifikanten Haupteffekt [ $F(2,7;125,2) = 0,3$ ;  $p = 0,9$ ;  $\eta^2 = 0,006$ ], jedoch einen Interaktionseffekt mit dem Faktor *Expositionsdauer* [ $F(5,4;125,2) = 2,82$ ;  $p < 0,05$ ;  $\eta^2 = 0,11$ ]. Die Analyse des Faktors *Expositionsdauer* ergab keinen signifikanten Effekt [ $F(2;46) = 0,11$ ;  $p = 0,9$ ;  $\eta^2 = 0,005$ ]. Durchschnittlich betragen die JNDs nach Exposition mit um -125 ms, -63 ms, 0 ms, 63 ms und 125 ms gegeneinander verschobenen Stimuli 153 ms (7,3 ms), 155 ms (8,3 ms), 155 ms (8,3 ms), 150 ms (8,4 ms) und 152 ms (6 ms, Standardfehler in Klammern; s. Abb. 21 und Tab. 3). Bei einminütiger Exposition betragen die JNDs durchschnittlich 151 ms (10,3 ms), bei zweiminütiger 158 ms (13,2 ms) und bei dreiminütiger 150 ms (9,8 ms, s. Tabelle 3 und Abb. 21). Um zu überprüfen, welche Expositionsdauer den Interaktionseffekt mit den JND-Daten erklären kann, wurden fünfstufige Varianzanalysen mit Messwiederholung getrennt nach Experimenten bzw. Expositionslängen durchgeführt. Diese ergaben keine Effekte des Faktors *Zeitverzögerung der Exposition* für die Experimente eins [Expositionslänge: eine Minute;  $F(2,1;35,5) = 1,2$ ;  $p = 0,32$ ,  $\eta^2 = 0,07$ ] und zwei [Expositionslänge: zwei Minuten;  $F(2,2;22,3) = 0,7$ ;  $p = 0,59$ ;  $\eta^2 = 0,07$ ], jedoch für Experiment drei [Expositionslänge: drei Minuten;  $F(1,8;34,2) = 4,17$ ;  $p < 0,05$ ;  $\eta^2 = 0,18$ ]. Post-Hoc-Analysen dieses Faktors bei dreiminütiger Exposition zeigten, dass JNDs nach Exposition mit audiovisuell synchronen Stimuli größer waren als JNDs nach Exposition mit allen verwendeten asynchronen Verzögerungsstufen mit Ausnahme der JNDs nach Exposition bei -125 ms (Ton zuerst). JNDs nach dieser Expositionsbedingung unterschieden sich nicht signifikant von JNDs anderer Expositionsbedingungen.

Eine Varianzanalyse mit den PSS-Daten aller drei Teilexperimente ergab keine signifikanten Effekte [Faktor *Zeitverzögerung der Exposition*:  $F(4;184) = 1,05$ ,  $p = 0,38$ ;  $\eta^2 = 0,02$ ; Interaktion mit *Expositionsdauer*:  $F(8;184) = 0,44$ ;  $p = 0,9$ ;  $\eta^2 = 0,02$ ; *Expositionsdauer*:

$F(2;46) = 0,07$ ;  $p = 0,93$ ;  $\eta^2 = 0,003$ ; s. auch Abb. 30]. Punkte subjektiver Gleichzeitigkeit (PsSS) betragen über alle drei Vorexperimente hinweg durchschnittlich 36 ms (4,2 ms) bei -125 ms, 39 ms (4,5 ms) bei -63 ms, 41 ms (5 ms) bei 0 ms, 39 ms (4,5 ms) bei 63 ms und 43 ms (5,3 ms) bei 125 ms audiovisueller Verschiebung während der Exposition (s. Tab. 4). Nach einer Minute Exposition befanden sich die PsSS durchschnittlich bei 41 ms (7 ms), nach zwei Minuten ebenfalls bei 41 ms (9 ms) und nach drei Minuten bei 38 ms (7 ms, Standardfehler in Klammern; s. Tab. 4). Ergebnisse der Varianzanalysen wurden korrigiert, falls keine Sphärizität vorlag.



**Abbildung 21.** Die Grafik zeigt JNDs des Vorexperiments 2.1 in Abhängigkeit des SOAs der Expositionsphase und der Dauer der Expositionsphase. JNDs nach Exposition bei verschiedenen zeitlichen Verzögerungen zwischen Licht und Ton veränderten sich in Abhängigkeit der Expositionsdauer. Signifikante Unterschiede der JNDs traten nach 3 Minuten Exposition mit Stimuli auf, die phasengleich dargeboten wurden. Fehlerbalken stellen Standardfehler dar. \* stellen signifikante Unterschiede der Daten zwischen den SOAs der Expositionsphase dar.

**Tabelle 3.** JNDs nach Exposition unterschiedlicher Dauer bei verschiedenen zeitlichen Verzögerungen zwischen Licht und Ton. Standardfehler in Klammern. S. auch Kapitel 3.2.7.3.1.2, Abb.29.

		<i>SOA der Expositionsphase</i>				
		<i>-125 ms</i>	<i>-63 ms</i>	<i>0 ms</i>	<i>63 ms</i>	<i>125 ms</i>
<i>Dauer</i>	<i>1 Minute</i>	154 ms (12)	157 ms (15)	138 ms (7)	160 ms (18)	145 ms (8)
	<i>2 Minuten</i>	155 ms (11)	162 ms (18)	153 ms (11)	151 ms (13)	167 ms (14)
	<i>3 Minuten</i>	148 ms (12)	145 ms (11)	176 ms (18)	139 ms (10)	143 ms (10)

**Tabelle 4.** PsSS nach Exposition unterschiedlicher Dauer bei verschiedenen zeitlichen Verzögerungen zwischen Licht und Ton. Standardfehler in Klammern. S. auch Kapitel 3.2.7.3.1.2, Abb.30.

		<i>SOA der Expositionsphase</i>				
		<i>-125 ms</i>	<i>-63 ms</i>	<i>0 ms</i>	<i>63 ms</i>	<i>125 ms</i>
<i>Dauer</i>	<i>1 Minute</i>	39 ms (7)	38 ms (15)	40 ms (8)	40 ms (8)	46 ms (9)
	<i>2 Minuten</i>	35 ms (10)	42 ms (9)	41 ms (11)	41 ms (11)	45 ms (14)
	<i>3 Minuten</i>	35 ms (5)	38 ms (6)	42 ms (7)	36 ms (5)	37 ms (6)

### 3.2.3.3 Bewertung der Ergebnisse von Vorexperiment 2.1

#### 3.2.3.3.1 Synchronizitätsurteile

Die Analyse der Synchronizitätsurteile ergab, wie schon in Vorexperiment 1, dass Probanden Asynchronizität leichter entdeckten, wenn Tonreize dem Licht vorangingen als umgekehrt (s. linksschiefe Verteilung der Urteile). Diese Asymmetrie repliziert Befunde aus anderen Studien (z. B. Dixon & Spitz, 1981; Recanzone, 2003) und reflektiert den Tatbestand, dass die menschliche Wahrnehmung natürlicherweise in Richtung visueller Ereignisse verzerrt ist (z. B. Fujisaki, et al. 2004; Grant et al. 2003; Spence & Squire, 2003). Da sich Licht schneller durch den Raum bewegt als Schall (s. 1.2), sind wir gewohnt, visuelle Ereignisse (kurz) vor auditorischen wahrzunehmen.

#### 3.2.3.3.2 JNDs und PsSS

Die weitere Analyse der Synchronizitätsurteile ließ erkennen, dass nach dreimütiger Exposition ein anderes Antwortmuster vorlag als nach einer bzw. zwei Minuten Stimulusexposition. Die Analyse der ermittelten Schätzwerte für PsSS und JNDs ergab, dass die Dauer, mit der Probanden vorab Stimuli präsentiert wurden, deren Sensitivität für Gleichzeitigkeit einfacher Stimuli beeinflusste (JNDs), nicht aber deren subjektives Gleichzeitigkeitsempfinden (PsSS). Probanden wurden nach drei Minuten Stimulusexposition mit synchronen audiovisuellen Stimuli unempfindlicher gegenüber Asynchronizität, d.h. dass es ihnen schwerer fiel, Ungleichzeitigkeit bei dargebotenen Testsequenzen zu entdecken bzw. ihr audiovisuelles Zeitfenster, um zwei Ereignisse unterschiedlicher Modalität einander zuzuordnen, vergrößerte sich. Diese Auswirkung zeigte sich nicht nach einer oder zwei Minuten Stimulusexposition. Die JNDs mit Werten von 138 ms bis 176 ms und PsSS mit Werten von 35 ms bis 46 ms waren höher als die von

Vorexperiment 1 (JND: 123 ms und PSS: 26 ms bei Ereignislänge vier). Dies kann einerseits ein Effekt der Stimulusexposition sein, die während Vorexperiment 1 nicht stattfand oder andererseits Auswirkungen der verwendeten SOAs: im vorherigen Vorexperiment wurden zusätzlich zwei asynchrone SOAs ( $\pm 360^\circ$ ) präsentiert. Dies wiederum kann sich auf den Schätzalgorithmus auswirken, der aus den Synchronizitätsurteilen JNDs und PsSS ermittelt, da in der vorliegenden Untersuchung zwei Datenpunkte weniger vorlagen und die Schätzung daher gegebenenfalls ungenauer wurde. Darüber hinaus haben die beiden zusätzlichen Bedingungen, die deutliche Verzögerungen zwischen Licht- und Tonsignal aufwiesen, eventuell als klare asynchrone Vergleichsbedingung wirken können, die dazu führten, dass Probanden eindeutiger urteilen konnten und dadurch mehr asynchrone Urteile für asynchrone Stimuli und mehr synchrone für synchrone Stimuli abgaben, was wiederum zu höherer Sensitivität gegenüber Ungleichzeitigkeit und niedrigeren JNDs führte. Jedoch konnten Zampini et al. (2005) zeigen, dass die Spanne der präsentierten SOAs in ihrer Studie keinen Einfluss auf JNDs und PsSS hat. Dies impliziert, dass Probanden keine Antwortstrategien verfolgen, um beispielsweise Antworten gleichmäßig jeder Experimentalbedingung zuzuordnen. Die Autoren schließen, dass die auch hier angewandte Methode, aus Synchronizitätsurteilen Maße für JND und PSS zu gewinnen geeignet ist, um perzeptuelle Effekte, und nicht Antworttendenzen zu erfassen (s. auch 2.1.1).

Ein weiterer Grund für die hohen Werte des vorliegenden Vorexperiments könnte darin bestehen, dass ein innerer Abgleich mit den Stimuli der Expositionsphase geschieht, wenn Probanden die Testsequenzen bewerten sollen und daher das perzeptuelle System mehr beansprucht ist. Untersuchungen, die die perzeptuelle Beanspruchung des audiovisuellen Zeitsystems systematisch beeinflussten, ergaben, dass JNDs und PsSS bei hoher Beanspruchung größer werden (z. B. Alsius, 2005; Vatakis et al., 2007). Die hier ermittelten JNDs sind allerdings in Einklang mit anderen Befunden, die denselben Schätzalgorithmus verwendeten, um PsSS und JNDs zu ermitteln (Navarra et al., 2005; Fujisaki et al., 2004; Spence et al., 2001; Stone et al., 2001; Vatakis, 2008; Zampini et al., 2005; JNDs von 114 ms – 192 ms). Es konnte im vorliegenden Vorexperiment kein Einfluss der Expositionsdauer auf die subjektiv empfundene Gleichzeitigkeit gefunden werden (PSS). Der vorliegende Befund, dass visuelle Stimuli vor auditorischen präsentiert werden mussten, damit Sequenzen als synchron empfunden werden, deckt sich jedoch mit

Ergebnissen anderer Studien (z. B. Jaskowski et al., 1990; Slutsky & Recanzone, 2001; Van de Par et al., 1999; Zampini et al., 2005). Stimulusexposition mit audiovisuellen synchronen Reizen beeinflusst demzufolge unter der vorliegenden Stimuluskonfiguration die Sensitivität gegenüber (A)synchronizität, bewirkt jedoch keine Rekalibrierung des bimodalen Zeitempfindens. Diese Ergebnislage lässt sich gegebenenfalls auf die Aufgabenstellung zurückführen, die von den Probanden verlangte, zwei Ereignisse zu verbinden, oder zu trennen (Synchronizitätsurteil). Dies wiederum könnte die Bereitschaft der Probanden erhöht haben, zwei zeitlich verzögerte Ereignisse einander zuzuordnen (Einheitlichkeitsannahme: Vatakis & Spence 2006; Welch & Warren 1980) und sie daher dazu veranlasst haben, ihr Antwortverhalten dementsprechend anzupassen. Dieser Umstand mag dazu geführt haben, dass eine Änderung der JNDs ermittelt wurde, nicht jedoch eine tatsächliche Veränderung des subjektiven Zeitlichkeitsempfindens (s. auch Vatakis et al., 2008). Ein vergleichbares Ergebnis trat in zwei Studien auf, die Probanden ebenfalls (a)synchronen audiovisuellen Stimuli aussetzten und deren Synchronizitätsempfindlichkeit (JND) und Gleichzeitigkeitsempfinden (PSS) ermittelten (Navarra et al., 2005; Vatakis et al., 2008): die JNDs veränderten sich von 120 auf 145 ms bzw. von 140 auf 157 ms, wenn Probanden asynchronen Reizen im Vergleich zu synchronen ausgesetzt wurden, nicht jedoch die PsSS. Andere Studien, die audiovisuelle Adaptation auf Zeitwahrnehmung untersuchten, fanden Effekte für PsSS, aber nicht für JNDs (Vatakis et al., 2007; Vroomen, 2004), jedoch auch Veränderungen beider Parameter (Fujisaki et al., 2004; Vatakis, 2008). In der vorliegenden Voruntersuchung liegt in bezug auf die Veränderung der JNDs, verglichen mit den Studien von Navarra et al. (2005) und Vatakis et al. (2008), ein gegensätzlicher Befund vor: nach Exposition mit synchronen Stimuli vergrößerte sich das audiovisuelle Zeitfenster der Probanden und nicht, wie in vergangenen Studien, nach Exposition mit asynchronen Stimuli. Das Ausmaß der Veränderung ist allerdings mit diesen Studien vergleichbar (hier: ca. 32 ms). In diesen Studien wurde argumentiert, dass Adaptation auf asynchrone Stimuli die Empfindlichkeit gegenüber asynchronen Reizen vermindert (also zu größeren JNDs führt, s. z. B. Navarra et al., 2005). Im Gegensatz dazu schienen Probanden in der vorliegenden Studie eine höhere Bereitschaft zu haben, Stimuli als synchron zu bewerten, nachdem sie synchronen Stimuli ausgesetzt wurden bzw. bekräftigte dies ihre Einheitlichkeitsannahme in bezug auf zeitlich verzögerte Ereignisse. D.h. die Bereitschaft, zwei zeitlich auseinander

liegende Ereignisse zu einem multisensorischen zusammenzufassen, stieg durch die Exposition an. Folglich verminderte sich die Sensitivität gegenüber Asynchronizität nach Adaptation auf synchrone Stimuli (höhere JNDs) und es trat eine Art Generalisierung der Expositionsbedingung auf die Folgestimuli auf.

Prinzipiell wird vermutet, dass multimodale Adaptation in zwei Schritten verläuft:

1. die Exposition mit asynchronen Stimuli führt initial zu einer Vergrößerung des multimodalen Zeitfensters, das sich nach Adaptation anschließend
2. um einen neuen Punkt subjektiver Gleichzeitigkeit wieder verkleinert (Navarra et al., 2005; Vatakis et al., 2008).

Dieser Theorie zufolge wird ein neuer PSS gesucht, der zuvor eine Ausdehnung des zeitlichen Rahmens verlangt, in dem sich das perzeptuelle System bewegt. Es finde eine ‚Suche‘ nach neuen zeitlichen Mustern statt.

Die allgemein höheren Werte für JNDs und PsSS in diesem Vorexperiment könnten, angelehnt an diese Theorie, dadurch erklärt werden, dass es über die Länge der Expositionsphase bis zu Anfang der Abfragephase zwar zu einem Anstieg der Zeitfenster und PsSS gekommen sein kann, die Werte im Verlauf der Abfrage aber wieder zum Ursprungsniveau zurückkehrten. Im Mittel würden sich die Veränderungen in diesem Fall nur durch insgesamt erhöhte Werte im Vergleich zu nicht durchgeführter Exposition ausdrücken, wie es den hier vorliegenden Daten entspricht. Die Adaptationseffekte würden zu schnell wieder abgeklungen sein, um durch das aktuelle Paradigma erfassbar zu sein. Dauer und Verfall erfolgter zeitlicher multimodaler Adaptation wurden bisher nicht systematisch untersucht. Die gefundenen Effekte in den JNDs nach Exposition mit synchronen Stimuli könnten daher rühren, dass synchrone Stimuli schon während der Exposition eindeutig einem synchronen audiovisuellen Perzept zuzuordnen waren und es somit zu keiner dauerhaften Veränderung des Zeitwahrnehmungssystems kam (keine beobachtete Änderung der PsSS). Während der Abfragephase könnte sich das perzeptuelle System jedoch auf der ständigen ‚Suche‘ nach neuen, ‚nicht eindeutigen‘ (asynchronen) Reizen befunden haben, um Hinweise für eine zeitliche Rekalibrierung zu finden. Somit vergrößerte es dadurch das Integrationsfenster (hohe beobachtete JNDs), fand den Bedarf für eine zeitliche Anpassung jedoch nicht vor. Dieser ‚Suchmodus‘

könnte für die nachgewiesene Erhöhung der JND nach synchroner Exposition verantwortlich gewesen sein. Eine bzw. zwei Minuten Exposition hätten dieser Vermutung zufolge dem perzeptuellen System nicht genügt, um Eindeutigkeit bezüglich der Synchronizität der Stimuli gesichert herzustellen, da JNDs nach dieser Expositionsdauer nicht signifikant verändert waren. Diese Eindeutigkeit wäre demzufolge auch nicht bei asynchroner Exposition aufgetreten, da Effekte nur für Exposition mit synchronen Stimuli beobachtet wurden. Dass synchrone Stimuli während der Exposition zeitlich eindeutig zugeordnet werden konnten, könnte den Tatbestand erklären, dass die PsSS in allen drei Experimenten nach synchroner Exposition nahezu identisch waren, nicht aber nach asynchroner Exposition (s. Tab. 4). Des Weiteren war anhand der Daten sichtbar, dass die interindividuelle Variabilität der PsSS nach dreiminütiger Exposition etwas geringer war als nach ein- bzw. zweiminütiger (s. Standardfehler in Tab. 2). Dies könnte ebenfalls ein Hinweis darauf sein, dass sich das perzeptuelle System über alle Probanden hinweg nach dreiminütiger Exposition eindeutig auf die Stimuli einstellte und es daher weniger Variabilität der PsSS zwischen den Probanden gab.

Insgesamt ist die hohe interindividuelle Variabilität der Daten ein markantes Merkmal, das in vielen Studien, die multimodale PsSS und JNDs untersuchen, auftritt (z. B. Stone et al., 2001; Vatakis et al., 2007). Bisherige Studien zeigen, dass sich PsSS und JNDs sowohl zwischen Personen (z. B. Arnold et al., 2005; Stone et al., 2001), aber auch von Studie zu Studie sehr unterscheiden (Spence et al., 2001 für eine Zusammenfassung). Diese Unterschiedlichkeit der Gleichzeitigkeitswahrnehmung bleibt bisher ungeklärt (Spence & Squire, 2003) führt jedoch dazu, dass Veränderungen in den Daten statistisch schwer ermittelt werden können.

Zusammengefasst zeigen die Daten von Vorexperiment 2.1, dass sich das perzeptuelle System erst nach dreiminütiger Exposition hinsichtlich der Zeitwahrnehmung eindeutig beeinflussen lässt. Daher sollen in folgenden Experimenten Probanden drei Minuten lang zeitlich unterschiedlich verzögerten audiovisuellen Stimuli ausgesetzt werden. Allerdings wird vermutet, dass die Adaptationseffekte nicht lange genug aufrechterhalten werden konnten, um anhand des aktuellen Paradigmas hinreichend erfasst zu werden. Daher soll anhand des folgenden Vorexperiments überprüft werden, ob die Adaptationseffekte durch ‚Re-Adaptationsdurchgänge‘

während der Abfragephase länger aufrechterhalten werden können. Die eingangs formulierte Fragestellung lässt sich anhand des aktuell durchgeführten Experiments folgendermaßen beantworten:

1. Die Dauer, mit der Probanden zeitlich verzögerten audiovisuellen Stimuli ausgesetzt werden, wirkt sich auf deren Gleichzeitigkeitsurteile gegenüber anschließend präsentierten Teststimuli aus. Während das audiovisuelle Wahrnehmungssystem nach einer bzw. zwei Minuten Expositionsdauer keine Veränderungen in bezug auf Zeitfenster (JNDs) oder subjektive Gleichzeitigkeit (PsSS) aufwies, wurden JNDs nach dreiminütiger Exposition beeinflusst.

### 3.2.4 Vorexperiment 2.2

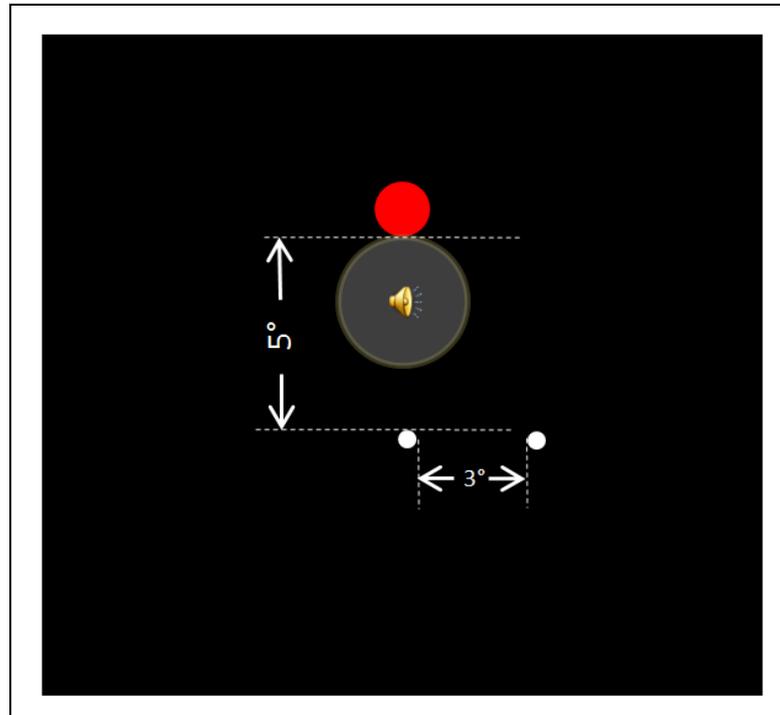
Anhand dieses Experiments soll geklärt werden, ob sich das perzeptuelle System während der Abfragephase re-adaptieren lässt und dadurch deutliche Rekalibrierungseffekte des audiovisuellen Zeitsystems nachweisbar werden. Falls dies zutrifft, soll diese experimentelle Modifikation in die folgende fMRT-Studie implementiert werden.

#### 3.2.4.1 Methodik

An diesem Experiment nahmen 25 rechtshändige (Selbstbericht) Probanden (Durchschnittsalter 24,3; Spanne 21-33; 13 weiblich) teil. Alle Teilnehmer berichteten, normales Hörvermögen und normales oder korrigiertes Sehvermögen und keine psychiatrischen oder neurologischen Störungen zu haben. Sie variierten hinsichtlich ihrer Vorerfahrung mit der Teilnahme an psychophysikalischen Studien und wurden über deren Zweck aufgeklärt. Das Experiment dauerte ca. 100 Minuten und wurde in einem vollständig dunklen und schallisolierten Raum durchgeführt.

Alle experimentellen Parameter, Design und Aufbau waren mit einer Ausnahme identisch zu denen von Vorexperiment 2.1 mit dreiminütiger Expositionslänge: während der Abfragephase wurden nun audiovisuelle Sequenzen dargeboten, die dieselbe zeitliche Verzögerung aufwiesen wie die Stimuli der jeweils vorhergehenden Expositionsphase (s. unten). Da anhand der Datenlage aus Vorexperiment 2.1 vermutet wird, dass durch

dessen Experimentalaufbau Veränderungen im Beurteilungsverhalten durch vorherige Stimulusexposition nicht lange genug persistierten, um statistisch erfassbar zu sein, soll durch das aktuelle Design versucht werden, die auftretenden Adaptationseffekte länger aufrecht zu erhalten.



**Abbildung 22.** Experimentalaufbau für Vorexperiment 2.2 und 2.3. Eine LED (Zielreiz) wurde  $5^\circ$  Sehwinkel über einem weißen Fixationspunkt angebracht. Sie konnte rote und grüne Lichtsignale senden. Töne wurden über einen Piezolausprecher direkt unter der LED präsentiert. Siehe Text für nähere Erläuterungen.

Die Re-Adaptationsdurchgänge bestanden, wie die Testdurchgänge, aus Sequenzen mit vier auditorischen und visuellen Ereignissen und wurden abwechselnd mit den Testsequenzen präsentiert. Damit Probanden zwischen Re-Adaptations- und Testsequenzen zuverlässig unterscheiden konnten, leuchtete 500 ms vor bis 200 ms nach einer Testsequenz eine weiße LED rechts unterhalb ( $5^\circ$  vertikal,  $3^\circ$  horizontal) der Zielreiz-LEDs auf und signalisierte den Probanden, dass sie die dargebotene Sequenz bewerten sollen (s. Abb. 22). Leuchtete die weiße LED nicht auf, handelte es sich um eine Re-Adaptationssequenz, die nur betrachtet, und nicht bewertet werden sollte. Die gesamte experimentelle Prozedur wurde vor Messbeginn in einem verkürzten Übungsdurchgang solange trainiert, bis die Probanden die Teststimuli korrekt

bewerteten. Da sich die Länge eines Durchgangs durch das Einfügen der Re-Adaptationssequenzen im Vergleich zu Vorexperiment eins verdoppelte, wurde die Anzahl der dargebotenen Testsequenzen auf 35 pro Durchgang beschränkt, um die Gesamtdauer des Experiments von 100 Minuten nicht zu überschreiten. Jede Verzögerungsbedingung konnte dadurch sieben Mal pro Durchgang und insgesamt 105 Mal über 15 Durchgänge gezeigt werden. So wurde jede Verzögerungsbedingung der Testsequenzen 21 Mal pro Verzögerungsbedingung der Expositionsphase dargeboten. Bedingungen der Expositionsphase und der Testsequenzen wurden in randomisierter Reihenfolge gezeigt. Der zeitliche Abstand zwischen zwei Sequenzen (ISI), unabhängig von Test- oder Re-Adaptationssequenz, betrug 2000 ms +/- 500 ms (Jittering).

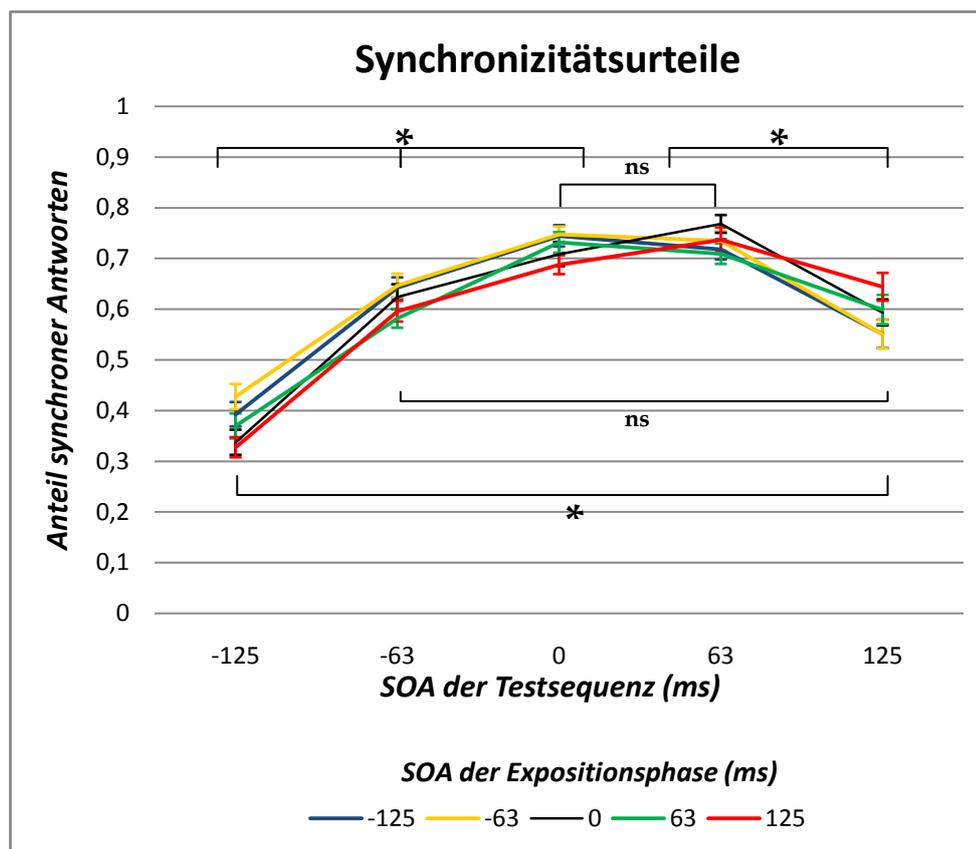
### 3.2.4.2 Ergebnisse des Vorexperiments 2.2

#### 3.2.4.2.1 Synchronizitätsurteile

Analog zu Vorexperiment 2.1 wurde eine 5 x 5-faktorielle Varianzanalyse mit den Faktoren *Zeitverzögerung der Exposition* und *Zeitverzögerung der Testsequenzen* durchgeführt. Fälle, die während der Expositionsphase weniger als 80% der „catch-trials“ entdeckt hatten und mehr als 20% falsch positive Antworten gaben, wurden aus der Analyse entfernt. Antworten, die während der Darbietung bis 200 ms nach Ende einer Testsequenz gegeben wurden, flossen nicht in die Analyse ein.

Durchschnittlich wurde in 37 % (4,2%), 62% (3,7), 72% (3,7), 73% (3,6%) und 59 % (5,1%) der Darbietungen bei Stimuli, die um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) bzw. 125 ms (Licht voran) verzögert waren, ein Synchronizitätsurteil abgegeben (s. Abb. 23). Nach Exposition mit um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) und 125 ms (Licht voran) audiovisuell gegeneinander verschobenen Sequenzen, wurden durchschnittlich 61% (3,6%), 62% (3,4%), 61% (3,9%), 60% (3,5%) und ebenfalls in 60% (3,5%) synchrone Urteile gemacht. Die Analyse ergab einen Haupteffekt des Faktors *Zeitverzögerung der Testsequenzen* [ $F(4;84) = 35,15$ ;  $p < 0,001$ ;  $\eta^2 = 0,63$ ] und einen Interaktionseffekt zwischen *Zeitverzögerung der Testsequenzen* und *Zeitverzögerung der Exposition* [ $F(16;336) = 2,4$ ;  $p < 0,01$ ;  $\eta^2 = 0,10$ ]. Die Greenhouse-Geisser-Korrektur wurde angewandt, falls keine Sphärizität nach Mauchly vorlag. Post-Hoc durchgeführte T-Tests (nach Bonferroni korrigiert) zeigten, dass mehr synchrone Antworten nach Exposition mit um 0 ms und 63

ms (Licht voran) verzögerten Stimuli gegeben wurden als nach Exposition mit größeren audiovisuellen Verzögerungen (alle  $p < 0,001$ ). Nach Exposition mit 0 ms und 63 ms verzögerten Stimuli wurden Synchronizitätsurteile vergleichbar häufig gemacht, ebenso wie nach Exposition mit um -63 ms (Ton voran) und 125 ms (Licht voran) verzögerten Stimuli ( $p = 1$ ). Wie in den vorgehenden Experimenten wurde folglich Asynchronizität besser entdeckt, wenn der Ton- dem Lichtreiz voranging (linksschiefe Verteilung der Antworten, vgl. 3.1.2.3.1, 3.1.3.2.1 und 3.2.3.2.1).

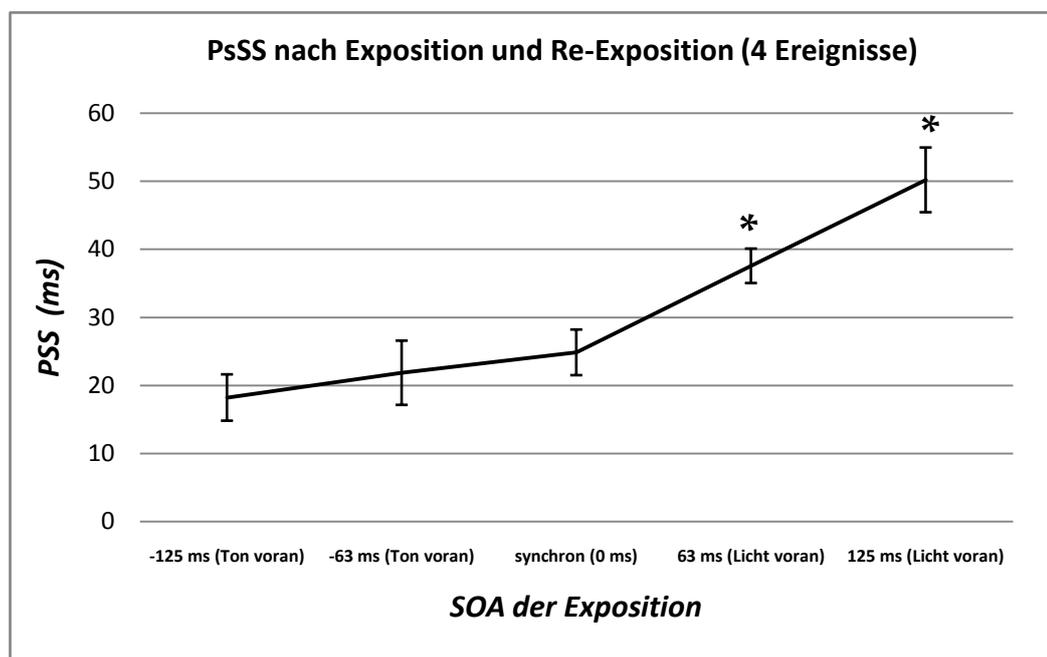


**Abbildung 23.** Anteil synchroner Antworten in Abhängigkeit des präsentierten SOAs der Testsequenz und des SOAs der Expositionsphase in Vorexperiment 2.2. Signifikante Unterschiede der Urteile, die mittels einer  $5 \times 5$ -stufigen Varianzanalyse ermittelt wurden, sind durch \* gekennzeichnet ( $p < 0,05$ ). Dargestellte Unterschiede beziehen sich auf Datenpunkte, die über alle SOAs der Expositionsphase kollabiert wurden (s. Text). Die Fehlerbalken stellen Standardfehler dar. Antworten unterschieden sich auf dem Faktor *SOA der Testsequenz* und hinsichtlich der Interaktion aus *SOA der Testsequenz* und *SOA der Expositionsphase* (s. Text). Sequenzen, die um 0 ms und 63 ms (Licht voran) verschoben waren, wurden signifikant am häufigsten als synchron beurteilt. ns = nicht signifikant.

### 3.2.4.2.2 JNDs und PsSS

Aus den Synchronizitätsurteilen wurden PsSS und JNDs geschätzt und jeweils 5-faktorielle Varianzanalysen mit Messwiederholung mit dem Faktor *Zeitverzögerung der Exposition* durchgeführt, um zu klären, ob sich diese Werte in Abhängigkeit der Verzögerungsbedingung der Expositionsphase signifikant verändern. Aus den geschätzten Werten für PsSS und JNDs wurden standardisierte Z-Werte berechnet und alle Probanden, deren Datenpunkte zwei Standardabweichungen über- oder unterschritten, eliminiert. Zusätzlich wurden Fälle entfernt, die größere PsSS als das maximale SOA (125 ms) oder größere JNDs als das Doppelte des maximalen SOAs (250 ms) aufwiesen. Dies verhindert, dass unzuverlässige Schätzungen aufgrund willkürlichen Antwortverhaltens in die Datenanalyse einfließen (s. Spence et al., 2001, Vatakis et al., 2007). Die JNDs betragen 132 ms (11,3 ms), 130 ms (10,8 ms), 125 ms (7,1 ms), 127 ms (8,7 ms) und 133 ms (5 ms) nach Exposition mit um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) und 125 ms (Licht voran) verzögerten audiovisuellen Stimuli. Sie unterschieden sich nicht signifikant voneinander [ $F(4;40) = 0,25$ ;  $p = 0,91$ ;  $\eta^2 = 0,02$ ]. PsSS stiegen von 18 ms (6,8 ms) über 22 ms (9,4 ms), 25 ms (6,7 ms) und 38 ms (5 ms) auf 50 ms (9,5 ms) nach Exposition mit um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) und 125 ms (Licht voran) verzögerten audiovisuellen Stimuli an (s. Abb. 24). Dies entspricht einer durchschnittlichen Verschiebung des PSS um ca. 8 ms pro Expositionsbedingung. *Zeitverzögerung der Exposition* ergab einen signifikanten Effekt für diese PSS-Werte [ $F(4;40) = 5,05$ ;  $p < 0,01$ ;  $\eta^2 = 0,34$ ]. Post-Hoc-Analysen zeigten, dass der mittlere PSS nach Exposition mit um -125 ms (Ton voran) verzögerten Stimuli signifikant niedriger war als nach den Expositionsbedingungen 63 ms und 125 ms (Licht voran,  $p < 0,01$ ). Er unterschied sich jedoch nicht von den mittleren PsSS der Expositionsbedingungen -63 ms (Ton voran) und 0 ms (synchron). Der mittlere PSS nach Exposition bei um -63 ms (Ton voran) verzögerten Stimuli war signifikant niedriger als nach Expositionsbedingung 125 ms (Licht voran,  $p < 0,05$ ), aber statistisch nicht unterschiedlich von den mittleren PsSS aller anderen Expositionsbedingungen. Der durchschnittliche PSS nach Exposition mit physikalisch synchronen Stimuli (0 ms Verzögerung) war signifikant kleiner als die PsSS nach Exposition mit Stimuli, bei denen Licht dem Ton voranging (63 ms:  $p < 0,05$ ; 125 ms:  $p < 0,001$ ), aber statistisch gleich mit mittleren PsSS nach Exposition mit Stimuli, bei denen der Ton- dem Lichtreiz voranging.

Der durchschnittliche PSS nach Expositionsbedingung 63 ms (Licht voran) war signifikant höher als mittlere PsSS nach Expositionsbedingung -125 ms (Ton voran) und 0 ms (synchron, alle  $p < 0,05$ ), aber statistisch gleich mit den PsSS der restlichen Expositionsbedingungen (-63 ms und 125 ms). Der höchste mittlere PSS nach Exposition mit um 125 ms (Licht voran) verzögerten Stimuli unterschied sich signifikant von den mittleren PsSS aller anderen Expositionsbedingungen. Diese Datenlage zeigt, dass sich die mittleren PsSS mit der Veränderung der Expositionsbedingung von einer Verschiebungsrichtung (Ton voran) zur anderen (Licht voran) stetig erhöhen.



**Abbildung 24.** Die Grafik zeigt PsSS des Vorexperiments 2.2 in Abhängigkeit des SOAs der Expositionsphase. PsSS nach Exposition bei verschiedenen zeitlichen Verzögerungen zwischen Licht und Ton veränderten sich. Signifikante Unterschiede der PsSS traten nach Exposition mit Stimuli auf, die um 63 ms und 125 ms (Licht voran) phasenverschoben dargeboten wurden. Fehlerbalken stellen Standardfehler dar. \* stellen signifikante Unterschiede der Daten zwischen den SOAs der Expositionsphase dar.

### 3.2.4.3 Bewertung der Ergebnisse von Vorexperiment 2.2

Anhand des vorliegenden Re-Adaptationsparadigmas konnte nachgewiesen werden, dass Stimulusexposition mit audiovisuell verzögerten Ereignissen nachhaltig das anschließende Antwortverhalten beeinflusst, wenn Probanden die Gleichzeitigkeit von dargebotenen Testreizen beurteilten. Probanden adaptierten demzufolge auf bestimmte zeitliche Intervalle zwischen den zwei dargebotenen Modalitäten. Es konnte gezeigt werden, dass sich das subjektive Gleichzeitigkeitsempfinden mit zunehmender zeitlicher

Verschiebung der auditorischen Ereignisse relativ zu den visuellen Ereignissen von der einen Asynchronizitätsrichtung (Ton voran) zur anderen (Licht voran) stetig in dieselbe Richtung verschiebt. D.h. dass Probanden ein immer größeres Intervall benötigten, um einen audiovisuellen Reiz als synchron zu empfinden, je mehr der Licht- dem Tonreiz voranging. Im Mittel verschob sich das subjektive Gleichzeitigkeitsempfinden (PSS) um 8 ms (von 18 ms bis 50 ms in vier Schritten, s. 3.2.4.2.2) bei einer durchschnittlichen Verschiebung der audiovisuellen Ereignisse während der Exposition (SOAs) gegeneinander um 62,5 ms (von -125 ms bis 125 ms in 4 Schritten). Dies macht 12,8% der während der Exposition präsentierten SOAs aus und lässt auf eine Rekalibrierung des audiovisuellen Zeitsystems schließen, die anhand einfacher Stimuli auch in anderen Untersuchungen nachgewiesen werden konnte (Fujisaki et al., 2004; Vroomen et al., 2004). Der PSS wurde während des vorliegenden Experiments von einer maximalen Verzögerungsbedingung der Exposition (-125 ms, Ton voran) zur anderen (125 ms, Licht voran) um 32 ms verändert. Dies entspricht einer Spanne, die auch Vroomen berichtete (35 ms). Fujisaki et al. (2004) konnten den PSS um 59 ms verändern, verwendeten zwar ebenfalls einfache Teststimuluspaarungen, jedoch größere audiovisuelle Verzögerungen während der Exposition. Verglichen mit den SOAs, die während der Exposition präsentierten wurden, betrug die Verschiebung der PsSS 10%. Dieses Ausmaß entspricht ungefähr dem Befund der vorliegenden Untersuchung. Allerdings sind in diesen beiden Studien (Fujisaki et al., 2004; Vroomen et al., 2004) die ermittelten PsSS geringer als in der vorliegenden. Sie bewegen sich in einem Bereich von ca. -30 ms (Ton voran) bis 30 ms (Licht voran), während hier Werte von 18 ms bis 50 ms (Licht zuerst) ermittelt wurden. Die Abweichungen in den Werten lassen sich eventuell auf die Unterschiede in den Paradigma zurückführen: Fujisaki et al. (2004) verwendeten größere und hellere Stimuli (Ringe) und präsentierten, wie auch Vroomen et al. (2004), sehr kurze Teststimuli, die lediglich aus je einem auditorischen bzw. visuellen Ereignis bestanden. Da bei kurzen audiovisuellen Ereignissen Gleichzeitigkeitssurteile erschwert werden (vgl. Vorexperiment 1), könnte dies den Unterschied mit den vorliegenden Ergebnissen erklären. Auch die größeren Verzögerungen zwischen Licht und Ton, die in den beiden Untersuchungen von Vroomen et al. (2004) und Fujisaki et al. (2004) präsentiert wurden, könnten die abweichenden Werte erklären. Zwar haben Zampini et al. (2005), ohne ein Expositionsparadigma zu verwenden, keine Veränderung der JNDs und PsSS feststellen

können, wenn größere SOAs dargeboten wurden, die PsSS eines Vergleichsexperiments bei Fujisaki et al. (2004) waren jedoch deutlich niedriger, als sie die Spanne der präsentierten SOAs vergrößerten. Es fällt auf, dass sich die subjektiven Gleichzeitigkeitspunkte (PsSS) in den Studien von Vroomen et al. (2004) und Fujisaki et al. (2004) nach Adaptation auf synchrone audiovisuelle Ereignisse um den physikalischen Nullpunkt bewegten, während sich in der vorliegenden Studie nach Exposition mit synchronen Reizen der PSS bei 25 ms (Licht vor Ton) befand und damit in einem Bereich, der unter anderem von King (1985) und Stone et al. (2001) als ‚natürlicher‘ subjektiver Gleichzeitigkeitspunkt betrachtet wird, da die menschliche Gleichzeitigkeitswahrnehmung wahrscheinlich in Richtung des Lichtereignisses verzerrt ist (vgl. 1.3.2.4, s. auch Dixon & Spitz, 1981; Spence & Squire, 2003). Werte in diesem Bereich wurden auch in Studien, die keine Adaptation verwendeten als subjektive Gleichzeitigkeitspunkte ermittelt (z. B. Jaskowski et al., 1990; Stone et al., 2001; Vatakis et al., 2007; Zampini et al., 2005; vgl. auch 3.1.2.3.1, 3.1.3.2.1 und 3.2.3.2.1). Gegebenenfalls konnten Probanden in den Studien von Vroomen und Fujisaki besser zwischen synchronen und asynchronen Teststimuli differenzieren, da die Schritte von einer Verzögerungsbedingung zur anderen dort größer waren (ca. 100 ms) als in der vorliegenden (kleinster Schritt: 63 ms), was zu weniger Synchronizitätsurteilen der Bedingungen geführt haben könnte, die dem physikalischen Nullpunkt am nächsten lagen und dadurch zu niedrigeren Schätzungen der PsSS führte.

Vatakis et al. (2008) konnten eine Veränderung des PSS um 15 ms für die Beurteilung kurzer audiovisueller Sprachstimuli nach Adaptation auf asynchrone Reize beobachten, aber auch eine Vergrößerung der Zeitfenster (JNDs). Navarra et al. (2005) fanden eine Veränderung der JNDs bei der Beurteilung einfacher audiovisueller Stimuli nach asynchroner Adaptation, jedoch keine Veränderung der PsSS. Ein wichtiger Unterschied in den Studien von Navarra et al. (2005) und Vatakis et al. (2008) im Vergleich zur vorliegenden Untersuchung und denen von Fujisaki et al. (2004) und Vroomen et al. (2004) ist, dass erstere Teststimuli *während* der Expositionsphase zeigten (‚online-Adaptation‘), während letztere diese *nach* der Exposition bewerten ließen. Da Navarra et al. (2005) und Vatakis et al. (2008) v.a. Veränderungen der JNDs feststellten und Vroomen et al. (2004), Fujisaki et al. (2008; wie auch die vorliegende Untersuchung) eine Änderung der PsSS beobachteten, entstand die Hypothese, dass das audiovisuelle

Zeitsystem, um sich zu rekalisieren, zuerst das Zeitfenster vergrößert (wie in den Studien zur ‚online-Adaptation‘ beobachtet), um sich wenig später um einen neuen, veränderten Gleichzeitigkeitsschwerpunkt wieder ‚zusammenzuziehen‘ (Navarra et al., 2005; Vatakis et al., 2008).

Veränderungen der PsSS nach Exposition mit Stimuli, bei denen der Ton- dem Lichtreiz vorangeht, sind weniger ausgeprägt als nach Exposition mit Stimuli, bei denen das Licht- dem Tonsignal vorangeht. Wie bereits erwähnt, ist dies ein Befund, wie er auch bei Vroomen et al. (2004) vorliegt und könnte damit zusammenhängen, dass Menschen sensibler gegenüber Stimuli sind, bei denen der Ton dem Licht vorangeht und die audiovisuelle Zeitwahrnehmung sich in dieser Richtung weniger beeinflussen lässt. Die Erfahrungen des menschlichen Wahrnehmungssystems mit audiovisuellen Eindrücken, die aufgrund eines vorangehenden Tonereignisses korrigiert werden müssen, sind wahrscheinlich geringer als der umgekehrte Fall.

JNDs und PsSS des vorliegenden Vorexperiments sind mit den Werten des Vorexperiments 1 bei Ereignislänge vier vergleichbar. In jenem und dem vorliegenden Vorexperiment 2.2 waren alle ermittelten Werte niedriger als diejenigen, die in Vorexperiment 2.1 ermittelt wurden. Zusätzlich fällt auf, dass die interindividuelle Varianz in bezug auf JNDs und PsSS in Vorexperiment 2.1 größer war als in den beiden anderen Vorexperimenten 1 und 2.2. Die Variabilitätsunterschiede lassen sich wahrscheinlich darauf zurückführen, dass die Effekte der Exposition in Vorexperiment 2.1 von Proband zu Proband unterschiedlich lang andauerten, was zu heterogenen Daten führte, während es im Vorexperiment 1 keine Exposition gab und im vorliegenden Vorexperiment 2.2 Probanden durch die Re-Adaptation vermutlich ‚gleichmäßig‘ adaptiert wurden. Dadurch entstand in diesen beiden Untersuchungen vermutlich eine homogenere Datenlage. Anhand der höheren Werte in Vorexperiment 2.1 lässt sich schließen, dass jenes Paradigma zu höherer Bereitschaft des audiovisuellen Zeitsystems führte, Asynchronizität zuzulassen. Vermutlich geschah dies durch fehlende, eindeutige Hinweise über das Ausmaß der erforderlichen zeitlichen Rekalibrierung wie sie wahrscheinlich durch die Re-Adaptationsdurchgänge im vorliegenden Vorexperiment 2.2 gegeben wurden.

Die vorliegenden Daten und die Befunde anderer Studien (Fujisaki et al., 2004; Navarra et al., 2005; Recanzone, 2003; Spence & Squire, 2003; Sugita & Suzuki, 2003;

Vatakis et al., 2008; Vroomen et al., 2004) legen nahe, dass ein Mechanismus existiert, der audiovisuelle Zeitverzögerungen korrigieren und die multisensorische Wahrnehmung verändern kann. Das bedeutet, dass audiovisuelle Zeitwahrnehmung erfahrungsabhängig und flexibel ist. Allerdings scheint das System nicht nur aktuelle Erfahrungen (hier vorherige Exposition) sondern auch langfristige und daher reliable und vertrauenswürdige Information einzubeziehen (alltägliche Erfahrungen). Diese Vermutung kann erklären, dass die Einschätzung der Zeitlichkeit während der Experimente nur um ca. 10-15% der präsentierten Verzögerungen korrigiert wird (s. auch Fujisaki et al., 2004). Es konnte gezeigt werden, dass sich die durch Exposition veränderte audiovisuelle Zeitwahrnehmung bestimmter audiovisueller Stimuli auch auf die zeitliche Wahrnehmung anders gearteter audiovisueller Stimuli auswirkt (Fujisaki et al., 2004; Navarra et al., 2005; Vatakis et al., 2008), was die Existenz eines generellen audiovisuellen Zeitsystems nahelegt, das entscheidend für das Verknüpfen zweier Ereignisse unterschiedlicher Modalitäten ist.

Prinzipiell konnten Adaptationsmechanismen intramodal im visuellen (Bennett & Westheimer, 1985; Stratton, 1897, zit. nach Sperry, 1950; und auditorischen System (Dix et al., 1949; Okada & Kashino, 2003) nachgewiesen werden. Im sensumotorischen Bereich konnte gezeigt werden, dass Probanden auf Verzögerungen zwischen motorischer und visueller Antwort adaptieren (Cunningham et al., 2001a; Cunningham et al., 2001b). Ebenso wurde nachgewiesen, dass audiovisuelle Adaptation in bezug auf räumliche Aspekte auftritt (Bauchrednereffekt: Canon et al., 1970; Recanzone, 1998). Zudem wurde gezeigt, dass die Verzerrung des visuellen Systems (das Tragen von Prismenbrillen) eine Veränderung der Tonlokalisation bewirkt (Knudsen & Knudsen, 1985; Zwiers et al., 2003). Diese Befunde legen zusammen mit denen des vorliegenden Experiments den Schluss nahe, dass das menschliche Wahrnehmungssystem per Adaptationsmechanismen bestrebt ist, räumlich-zeitliche Kongruenz verschiedener Sinneskanäle herzustellen, um wichtige Informationen zur Merkmalsbindung zu erhalten.

Zusammengefasst lässt sich aus dem vorliegenden Experiment schließen, dass eine nachweisbare Rekalibrierung der audiovisuellen Zeitwahrnehmung durch dreiminütige Exposition mit audiovisuellen Stimuli, die unterschiedliche Verzögerungen zwischen Licht und Ton aufwiesen und zusätzlicher intermediärer Re-Adaptation während der Abfragephase erreicht wurde. Diese Beobachtung bestätigt die oben

aufgeführte Vermutung, dass durch das Paradigma von Vorexperiment 2.1 durch fehlende Re-Exposition Adaptationseffekte nicht aufrechterhalten werden konnten. Im folgenden fMRT-Experiment sollen demzufolge Re-Adaptionsdurchgänge verwendet werden. Die eingangs formulierte Fragestellung lässt sich daher wie folgt beantworten:

2. Re-Adaptation bewirkt eine Verschiebung der PsSS in Richtung der Verzögerungsbedingung der Expositionsphase und lässt damit auf eine Rekalibrierung der audiovisuellen Zeitwahrnehmung schließen.

Aus der oben aufgeführten Datenlage ergibt sich zusätzlich die Vermutung, dass durch eine Verkürzung der Testsequenzen die auftretenden Adaptations-Nacheffekte gegebenenfalls länger aufrecht erhalten werden können, als es im vorliegenden Experiment möglich war. Dies soll anhand eines dritten Vorexperiments überprüft werden.

### 3.2.5 Vorexperiment 2.3

Dieses Vorexperiment soll Aufschluss darüber geben, ob sich das Verkürzen der Testsequenzen während der Abfragephase auf das Ausmaß der Adaptation des audiovisuellen Wahrnehmungssystems auswirkt, indem beispielsweise die Nacheffekte der Exposition länger aufrecht erhalten werden können als in den beiden vorangegangenen Vorexperimenten 2.1 und 2.2. Falls anhand dieses Paradigmas deutliche Änderungen der PsSS oder JNDs provoziert werden können, soll das vorliegende Experimentaldesign für die nachfolgende fMRT-Studie übernommen werden.

#### 3.2.5.1 Methodik

An diesem Experiment nahmen 26 rechtshändige (Selbstbericht) Probanden (Durchschnittsalter 24; Spanne 20-29; 10 weiblich) teil. Alle Teilnehmer berichteten, normales Hörvermögen und normales oder korrigiertes Sehvermögen und keine psychiatrischen oder neurologischen Störungen zu haben. Sie variierten hinsichtlich ihrer Vorerfahrung mit der Teilnahme an psychophysikalischen Studien und wurden über

deren Zweck aufgeklärt. Das Experiment dauerte ca. 100 Minuten und wurde in einem vollständig dunklen und schallisolierten Raum durchgeführt.

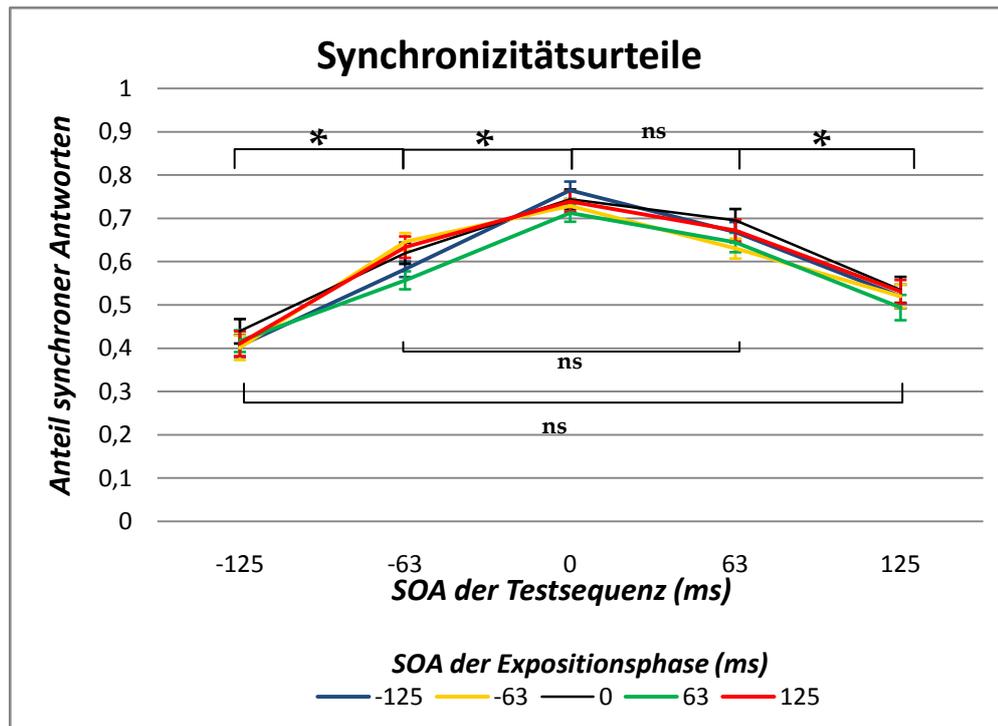
Der Aufbau dieser Untersuchung ist, mit einer Ausnahme, identisch mit Vorexperiment 2.2: Testsequenzen, die während der Abfragephase jeweils nach den Re-Adaptationsdurchgängen gezeigt werden sollen, sind hier auf eine Ereignislänge von je einem auditorischen bzw. visuellen Signal reduziert. Da sich so der zeitliche Abstand zwischen zwei Re-Adaptationsdurchgängen verkürzt, kann vermutet werden, dass die Nacheffekte des audiovisuellen Systems länger aufrecht erhalten werden als in Vorexperiment 2.1 und 2.2.

Der Abstand zwischen einem Test- und einem Re-Adaptationsdurchgang betrug 1500 ms (+/- 500 ms, Jittering). Zweihundert Millisekunden, bevor ein Testdurchgang präsentiert wurde, leuchtete eine weiße LED rechts unterhalb des Ziellichts auf (5° vertikal, 3° horizontal), um den Probanden zu signalisieren, dass es sich um einen zu bewertenden Teststimulus handelt. Diese zusätzliche weiße LED leuchtete 500 ms lang und überlappte zeitlich mit dem Teststimulus (s. Abb. 22; s 3.2.4.1). Alle anderen experimentellen Parameter sind identisch mit denen von Vorexperiment 2.2.

### **3.2.5.2 Ergebnisse des Vorexperiments 2.3**

#### **3.2.5.2.1 Synchronizitätsurteile**

Analog zu Vorexperiment eins und zwei wurde eine 5 x 5-faktorielle Varianzanalyse mit Messwiederholung für die Faktoren *Zeitverzögerung der Exposition* und *Zeitverzögerung der Testsequenzen* anhand der Synchronizitätsurteile durchgeführt. Fälle, die während der Expositionsphase weniger als 80% der „catch-trials“ entdeckt hatten und mehr als 20% falsch positive Antworten gaben, wurden aus der Analyse entfernt. Antworten, die während der Darbietung bis 200 ms nach Ende einer Testsequenz gegeben wurde, flossen nicht in die Analyse ein.



**Abbildung 25.** Anteil synchroner Antworten in Abhängigkeit des präsentierten SOAs der Testsequenz und des SOAs der Expositionsphase in Vorexperiment 2.3. Signifikante Unterschiede der Urteile, die mittels einer 5 x 5 – stufigen Varianzanalyse ermittelt wurden, sind durch \* gekennzeichnet ( $p < 0,05$ ). Dargestellte Unterschiede beziehen sich auf Datenpunkte, die über alle SOAs der Expositionsphase kollabiert wurden (s. Text). Die Fehlerbalken stellen Standardfehler dar. ns = nicht signifikant. Antworten unterschieden sich auf dem Faktor SOA der Testsequenz (s. Text). Sequenzen, die um 0 ms und 63 ms (Licht voran) verschoben waren, wurden signifikant am Häufigsten als synchron beurteilt.

Durchschnittlich wurde in 42% (5,1%), 61% (4%), 74% (3,8%), 66% (4,4%) und 52% (5,4%) der Darbietungen bei Stimuli, die um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) bzw. 125 ms (Licht voran) verzögert waren ein Synchronizitätsurteil abgegeben (s. Abb. 25). Nach Exposition mit um -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) und 125 ms (Licht voran) audiovisuell gegeneinander verschobener Sequenzen, wurden durchschnittlich 59 % (3,4%), 59% (3,7%), 61% (4,3%), 57% (3,5%) und in 60% (4%) synchrone Urteile gemacht. Die Analyse ergab einen Haupteffekt des Faktors *Zeitverzögerung der Testsequenzen* [ $F(1,6;37,7) = 17,4$ ,  $p < 0,001$ ;  $\eta^2 = 0,43$ ] und einen statistischen Trend des Faktors *Zeitverzögerung der Exposition* [ $F(2,7;61) = 2,58$ ;  $p = 0,07$ ;  $\eta^2 = 0,10$ ], jedoch keinen Interaktionseffekt beider Faktoren [ $F(8,3;190,3) = 1$ ;  $p = 0,42$ ;  $\eta^2 = 0,04$ ]. Die Greenhouse-Geisser-Korrektur wurde angewandt, falls keine Sphärizität nach Mauchly vorlag. Nach

Bonferroni korrigierte Post-Hoc-Analysen zeigten, dass Teststimuli, die um -125 ms (Ton voran) und 125 ms (Licht voran) verzögert waren, im Durchschnitt gleich häufig als synchron bewertet wurden. Stimuli, die um -63 ms (Ton voran) verzögert waren, wurden genauso häufig als synchron bewertet wie um 63 ms (Licht voran) und 125 ms (Licht voran) verzögerte Stimuli. Synchrone Testsequenzen wurden genauso oft als synchron eingestuft, wie um 63 ms (Licht voran) verzögerte Stimuli (alle  $p > 0,05$ ). Um 63 ms (Licht voran) und synchrone Testsequenzen wurden häufiger als synchron bewertet als Teststimuli, die mit größerer Verzögerung dargeboten wurden (alle  $p < 0,001$ ). Im Unterschied zu den drei vorherigen Vorexperimenten wurden, mit Ausnahme der um -125 ms (Ton voran) verzögerten Stimuli, bei allen asynchronen Reizen gleich viele Synchronizitätsurteile abgegeben. Daher weist die Verteilung der Synchronizitätsurteile dieses Vorexperiments eine weniger ausgeprägte Asymmetrie auf als in den vorgehenden Vorexperimenten.

### 3.2.5.2.2 JNDs und PsSS

Aus den Synchronizitätsurteilen wurden PsSS und JNDs geschätzt (s. 1.2.1). Aus diesen Werten wurden standardisierte Z-Werte berechnet und alle Probanden, deren Datenpunkte zwei Standardabweichungen über- oder unterschritten, eliminiert. Zusätzlich wurden Fälle entfernt, die größere PsSS als das maximale SOA (125 ms) oder größere JNDs als das Doppelte des maximalen SOAs (250 ms) aufwiesen. Dies verhindert, dass unzuverlässige Schätzungen aufgrund willkürlichen Antwortverhaltens in die Datenanalyse einfließen (s. Spence et al., 2001; Vatakis et al., 2007). Die durchschnittlichen PsSS betragen nach -125 ms (Ton voran), -63 ms (Ton voran), 0 ms (synchron), 63 ms (Licht voran) und 125 ms (Licht voran) Verzögerung während der Exposition 36 ms (9 ms), 27 ms (10 ms), 24 ms (11 ms), 25 ms (9 ms) und 27 ms (9 ms). Die mittleren JNDs betragen für diese Bedingungen 109 ms (7 ms), 106 ms (6 ms), 110 ms (7 ms), 109 ms (8 ms) und 105 ms (5 ms). Es wurde eine 5-faktorielle Varianzanalyse mit Messwiederholung für den Faktor *Zeitverzögerung der Exposition* für PSS- und JND-Daten durchgeführt, die keine Effekte für PsSS [ $F(2,6; 28,2) = 1,16; p = 0,34; \eta^2 = 0,10$ ] oder JNDs [ $F(4;44) = 0,26; p = 0,9; \eta^2 = 0,02$ ] ergab.

### 3.2.5.3 Bewertung der Ergebnisse von Vorexperiment 2.3

Anhand der Analyse der Synchronizitätsurteile wurde ein Trend dahingehend sichtbar, dass sich die Urteile um den physikalischen Nullpunkt symmetrisch verteilen. Zwar ist die Verteilung der Synchronizitätsurteile, wie in allen vorhergehenden Vorexperimenten, nach wie vor linksschief, jedoch lässt sich aus der aktuellen Antwortverteilung ableiten, dass Probanden synchrone und asynchrone Teststimuli vor allem, wenn das Licht- dem Tonsignal voranging, leichter diskriminieren konnten. Dies bedeutet, dass das audiovisuelle Wahrnehmungssystem Stimuli dieses Experiments, unabhängig davon, ob der Licht- dem Tonreiz vorausging oder vice versa, als gleichermaßen asynchron wertete. Dies steht im Gegensatz zu anderen Befunden (s. 3.1.2.3.1; 3.1.3.2.1; 3.2.3.2.1; 3.2.4.2; oder z. B. Dixon & Spitz, 1981; Stone et al., 2001; Zampini et al., 2005), bei denen Probanden Nachteile hatten, Asynchronizität zu entdecken, wenn der Licht- dem Tonreiz vorausging. Diese Asymmetrie spiegelt eine Verzerrung der audiovisuellen Wahrnehmung wider, die wahrscheinlich dadurch zustande kommt, dass bei audiovisuellen Ereignissen Lichtreize natürlicherweise die menschlichen Sinnesorgane eher erreichen als Schall (s. z. B. Spence & Squire, 2003).

Zwar verteilen sich die Synchronizitätsurteile der vorliegenden Untersuchung etwas anders als in den vorhergehenden Vorexperimenten, eine Varianzanalyse mit Messwiederholung, mittels der die Synchronizitätsurteile von Unterexperiment drei des Vorexperiments 2.1 (Expositionslänge: drei Minuten), Vorexperiment 2.2 und vom vorliegenden Vorexperiment 2.3 analysiert wurden, ergab jedoch lediglich einen statistischen Trend der Interaktion aus *Zeitverzögerung der Teststimuli* und *Experiment* [ $F(8;320) = 1,73$ ;  $p = 0,09$ ;  $\eta^2 = 0,04$ ]. Dementsprechend ist die Verteilung der synchronen Antworten, unabhängig davon, ob Abfragestimuli aus einem oder vier Ereignissen bestanden oder keine Re-Adaptation stattfand, statistisch gleich.

Da PsSS und JNDs anhand des vorliegenden Paradigmas durch Exposition mit verzögerten audiovisuellen Stimuli nicht signifikant beeinflusst wurden, ist daraus zu schlussfolgern, dass durch die Verkürzung der Abfragestimuli auftretende Nacheffekte nicht länger aufrecht erhalten werden können als durch längere Abfragesequenzen. Die PsSS-Werte dieser Untersuchung sind in ihrer Höhe mit den Befunden aus Vorexperiment 1 und 2.2 zu vergleichen. PsSS aus Vorexperiment 2.1 sind etwas höher als die vorliegenden (zur Diskussion der höheren Werte für jenes Experiment s. 3.2.3.3.2). Eine

5 x 3-faktorielle Varianzanalyse mit Messwiederholung, die PsSS aus von Vorexperiment 2.1 (Expositionslänge: drei Minuten), Vorexperiment 2.2 und aus dem aktuellen Vorexperiment 2.3 für die Faktoren *Zeitverzögerung der Exposition* und *Experiment* verglich, ermittelte einen Interaktionseffekt aus *Zeitverzögerung der Exposition* und *Experiment* [ $F(8;160) = 3,6; p < 0,001; \eta^2 = 0,15$ ]. Dies bedeutet, dass sich die PsSS, je nach dem, ob keine Re-Adaptation durchgeführt wurde oder ein oder vier Abfrageereignis(se) verwendet wurde(n), in Abhängigkeit der Expositionsbedingung unterschiedlich verschoben. Dieser Befund geht auf die Tatsache zurück, dass bei vier Abfrageereignissen eine Rekalibrierung des PSS hervorgerufen werden konnte, nicht aber durch ein Abfrageereignis oder das Weglassen der Re-Adaptation; also war eine erfolgreiche Adaptation nur durch die experimentelle Manipulation in Vorexperiment 2.2 zu erreichen. Der Faktor *Experiment* war nicht signifikant [ $F(2;40) = 0,71; p = 0,5; \eta^2 = 0,03$ ], d.h. dass PsSS, unabhängig davon, ob die Abfragestimuli aus einem oder vier Ereignis(sen) bestand(en) oder keine Re-Adaptation vorgenommen wurde, statistisch gleich waren. Die PsSS nach Exposition bei synchronen Stimuli sind zudem im Mittel nahezu identisch mit denen des Vorexperiments 1 bei einer SAPS von eins (22 ms, vgl. 3.1.2.3.2).

Interessanterweise stiegen die PsSS, ohne statistisch signifikant zu sein, im vorliegenden Vorexperiment an, wenn Probanden Stimuli ausgesetzt wurden, bei denen das Ton- dem Lichtsignal voranging, während sie in Vorexperiment 2.2 unter dieser Expositionsbedingung sanken und anstiegen, wenn Probanden Stimuli ausgesetzt wurden, bei denen der Licht- dem Tonreiz voranging. Vermutlich war der Teststimulus, der ja aus nur einem audiovisuellen Ereignis bestand zu kurz, um den Probanden ausreichend Hinweise über die Art der Asynchronizität zu liefern (Ton voran vs. Licht voran) und konnte vom Wahrnehmungssystem hinsichtlich seiner Verzögerungsrichtung nicht eingeordnet werden. Da das Wahrnehmungssystem Stimuli gegenüber sensibler ist, bei denen der Ton- dem Lichtreiz vorangeht, fand eine zeitliche Korrektur der audiovisuellen Wahrnehmung vermutlich besonders nach Exposition mit Stimuli statt, bei denen das Ton- dem Lichtsignal voranging. Zudem ist die interindividuelle Varianz in bezug auf die PsSS im vorliegenden Experiment höher als in den vorhergehenden Vorexperimenten (s. Standardfehler der PsSS). Dies könnte ebenfalls den geringen Informationsgehalt des kurzen Teststimulus als Ursache haben, der von Proband zu

Proband bzw. von Wahrnehmungssystem zu Wahrnehmungssystem wahrscheinlich sehr unterschiedlich interpretiert wurde.

JNDs sind im vorliegenden Experiment durchschnittlich kleiner als in den vorherigen Vorexperimenten. Eine 5 x 3-faktorielle Varianzanalyse mit Messwiederholung, die JNDs aus Vorexperiment 2.1 (Expositionslänge: drei Minuten), Vorexperiment 2.2 und aus dem aktuellen Vorexperiment 2.3 für die Faktoren *Zeitverzögerung der Exposition* und *Experiment* verglich ergab jedoch lediglich einen Haupteffekt für den Faktor *Experiment* [ $F(2;40) = 5,36$ ;  $p < 0,01$ ;  $\eta^2 = 0,21$ ]. Die nach Bonferroni korrigierten Post-Hoc-Analysen zeigten, dass JNDs in Vorexperiment 2.1 (Expositionsdauer: drei Minuten) größer waren als in Vorexperiment 2.2 und im aktuellen Vorexperiment 2.3 ( $p < 0,01$ ). Beide Faktoren interagieren zwar nicht signifikant miteinander, jedoch ist ein statistischer Trend hinsichtlich der Interaktion zu verzeichnen [ $F(4,7;95,5) = 1,9$ ;  $p = 0,1$ ;  $\eta^2 = 0,09$ ]. Die größeren JNDs in Vorexperiment 2.1 wurden bereits in Kapitel 3.2.3.3.2 diskutiert. Die im Trend sichtbare Interaktion zwischen JND-Verlauf und Experiment spiegelt die signifikante Änderung der JNDs in Abhängigkeit der Exposition bei Vorexperiment 2.1 wider, die im aktuellen Vorexperiment 2.3 und in Vorexperiment 2.2 nicht beobachtet wurde. Zwar unterschieden sich JNDs des aktuellen Vorexperiments 2.3 nicht signifikant von denen aus Vorexperiment 2.2 ( $p = 0,5$ ), jedoch waren sie im Mittel um 22 ms kleiner (Durchschnitt: 108 ms). Damit waren sie etwas kleiner als JNDs, die in Untersuchungen ermittelt wurden, die denselben Schätzalgorithmus für JNDs verwendeten wie das aktuelle Vorexperiment (Spence et al., 2001; Stone et al., 2001; Vatakis & Spence, 2006; Zampini et al., 2005; JNDs von 114 ms bis 192 ms). Studien, in denen nicht die Synchronizität, sondern die Reihenfolge der Asynchronizität (Ton zuerst vs. Licht zuerst; temporal order judgment: TOJ) beurteilt werden sollte, ermittelten geringere JNDs (z. B. Spence et al., 2001; Vatakis et al., 2008; Zampini et al., 2003; Werte zwischen 60 ms – 135 ms) als Studien, die, wie die vorliegende, Synchronizität beurteilen ließen. Vatakis et al. (2008) verglichen anhand einer Expositionsstudie die zwei Beurteilungsstrategien miteinander und fanden hinsichtlich daraus ermittelter JNDs und PsSS keine Unterschiede zwischen beiden Verfahren. Obwohl JNDs des vorliegenden Experiments nicht signifikant kleiner waren als die der vorhergehenden Experimente, könnten die geringeren Werte jedoch die Tatsache widerspiegeln, dass die Kürze des Teststimulus' dem Wahrnehmungssystem keine

Möglichkeit ließ, die Reihenfolge der Eingangssignale aus beiden Modalitäten zu bestimmen (Ton voran vs. Licht voran) und daher vorhergehende, ‚natürliche‘ Erfahrungen (der Licht- geht in natürlicher Umgebung oft dem Tonreiz voran; s. 1.2 oder z. B. Dixon & Spitz, 1981; Navarra et al., 2005; Spence & Squire, 2003) nicht in die Stimulusbeurteilung einfließen konnten. Alternativ könnte die kurze Spanne zwischen den Re-Adaptationsdurchgängen dazu geführt haben, dass frühere, ‚natürliche‘ Erfahrungen in bezug auf audiovisuelle Erfahrungen stärker außer Acht gelassen wurden als in vorgehenden Experimenten, in dem diese Spanne größer war. Dadurch würde die Sensibilität gegenüber Asynchronizität erhöht und führte zu kleineren JNDs. Das hieße, dass die Diskriminierbarkeit des Abfragedurchgangs oder die Spanne zwischen Re-Adaptationsdurchgängen als Moderatorvariable über die Inbezugnahme des audiovisuellen Zeitwahrnehmungssystems vorheriger Erfahrung entschiede. Die oben aufgeführte Datenlage der Synchronizitätsurteile des vorliegenden Vorexperiments 2.3, deren Verteilung, im Gegensatz zu allen anderen Vorexperimenten, eher eine symmetrische Form (im Vergleich zu einer linksschiefen) annimmt, würde diese Vermutung unterstützen, da sie zeigt, dass die Wahrnehmung der Probanden weniger zugunsten von Stimuli beeinflusst war, bei denen der Licht- dem Tonreiz vorangeht als es üblicherweise in vergleichbaren Untersuchungen gezeigt werden konnte (vgl. 3.1.2.3.1, 3.1.3.2.1 und 3.2.3.2.1, z. B. auch Dixon & Spitz, 1980). In diesem Fall würde die vorherige Stimulusexposition keine Rekalibrierung der audiovisuellen Zeitwahrnehmung erzeugen, sondern die Zeitwahrnehmung würde aufgrund der kurzen Spanne zwischen den Re-Adaptationsdurchgängen von dieser unmittelbar vorhergehenden Zeitlichkeitserfahrung dominiert. In diesem Fall müssten sich nicht nur JNDs, wie hier zu beobachten, verringern, sondern sich auch der PSS dem physikalischen Nullpunkt annähern. Letzteres ist anhand der hier aufgeführten Daten nicht zu beobachten, jedoch anhand der Daten von Fujisaki et al. (2004) und Vroomen et al. (2004), bei denen sich die PsSS nach Exposition mit synchronen Stimuli bei ca. 0 ms befanden. Das Eliminieren der länger zurückliegenden, ‚natürlichen‘ vorherigen audiovisuellen Zeitlichkeitserfahrung würde ein Zurücksetzen des Systems bewirken, auf das die verringerten JNDs im vorliegenden Experiment unter anderem hinweisen könnten. Leider wurden in den beiden oben erwähnten Studien, die audiovisuelle Nacheffekte anhand desselben Paradigmas wie das vorliegende Vorexperiment untersuchten (dreiminütige vorherige Exposition, ein

Testereignis), keine JNDs ermittelt, so dass ein Vergleich mit den vorliegenden JND-Daten nicht möglich ist (Fujisaki et al., 2004; Vroomen et al., 2004). Diese Studien konnten eine signifikante Rekalibrierung der audiovisuellen Wahrnehmung, wie sie hier durch Vorexperiment 2.2 (längere Abfragesequenzen) gezeigt wurde, demonstrieren. Wie oben bereits erwähnt, wurden in diesen Studien jedoch größere SOAs der Expositions- und Teststimuli verwendet, die ein klares Einordnen hinsichtlich der (A)Synchronizität der Abfragereize begünstigt haben könnten. Aufgrund der kleineren SOAs der vorliegenden Studie, könnte ein klares Einordnen der kurzen Teststimuli hinsichtlich ihrer Gleichzeitigkeit schwerer gewesen sein, als in den oben aufgeführten Studien, was wiederum eine nachweisbare Veränderung der PsSS verhindert haben könnte. Bei kurzen Teststimuli müssten dementsprechend die SOAs vergrößert werden, damit sie eindeutig bewertet werden können. Diese Vermutung wird auch durch Befunde von Vatakis et al. (2007) unterstützt: bei kurzen SOAs müsste der Reiz, um ausreichende Identifikation hinsichtlich seiner Gleichzeitigkeit zu ermöglichen, länger bzw. ökologisch valide sein: in der Umwelt kommen kurze Reize von 10 ms bis 30 ms Dauer vergleichsweise seltener vor, als beispielweise ‚wortlange‘ Reize von etwa einer Sekunde Dauer (s. Vorexperiment 2.2) oder länger.

Die JNDs des vorliegenden Experiments waren geringer als die des Vorexperiments 1 während ein audiovisuelles Ereignis präsentiert wurde (148 ms). Die größeren Zeitfenster in jener Studie könnten durch die dort fehlende Adaptation erklärt werden: ohne Expositionsdurchgänge würde das audiovisuelle System nicht durch neuartige Erfahrungen in bezug auf zeitliche Unterschiede zwischen Licht und Ton beeinflusst und hätte die systematische Verzerrung der Zeitwahrnehmung zugunsten von Ereignissen, bei denen der Licht- dem Tonreiz vorangeht, wahrscheinlich nicht korrigiert, sondern bisherige Erfahrungen (Licht- geht dem Tonreiz immer voraus) in Betracht gezogen. Zudem wäre das Wahrnehmungssystem wahrscheinlich durch die Präsentation der anderen, längeren Sequenzen, die während des Experiments zusätzlich präsentiert wurden und zuverlässig eingeordnet werden konnten, dahingehend verstärkt worden, diese erfahrungsbedingte Verzerrung aufrecht zu erhalten. Dies würde den Zeitrahmen vergrößern oder konstant halten, in dem zwei Ereignisse auftreten können, um immer noch als zusammengehörend wahrgenommen zu werden und damit zu größeren JNDs führen. Ein weiterer Aspekt, der zu niedrigeren JNDs im aktuellen Vorexperiment geführt

haben könnte, ist die Tatsache, dass Probanden durch die ständige, schnell aufeinanderfolgende Präsentation von einem audiovisuellen Ereignis, das ja weniger zu verarbeitende Information enthält als beispielsweise eine Sequenz, die aus vier Ereignissen besteht, rasch zu ‚Spezialisten‘ für die kurzen Teststimuli geworden sein könnten und so besser zwischen synchron und asynchron diskriminieren konnten.

Insgesamt betrachtet konnte die audiovisuelle Zeitwahrnehmung anhand des aktuellen Paradigmas, das im Vergleich zum vorhergehenden Vorexperiment nur ein Testereignis verwendete, nicht nachweisbar beeinflusst werden. Gründe dafür könnten, wie oben besprochen, ein Zurücksetzen von vorhergehenden Zeitlichkeitserfahrungen durch die verkürzten Abstände zwischen den Re-Adaptationsdurchgängen, oder die mangelnde Möglichkeit zur adäquaten Verarbeitung aufgrund der Kürze des Teststimulus sein. Die oben aufgeführte Fragestellung bezüglich des aktuellen Vorexperiments 2.3 kann daher wie folgt beantwortet werden:

3. Die Verkürzung der Teststimuli nach vorhergehender Exposition konnte keine Nacheffekte hinsichtlich der Gleichzeitigkeitsbeurteilung bewirken.

Folglich soll das Paradigma von Vorexperiment 2.2 für die nachfolgende fMRT-Studie verwendet werden.

### **3.2.6 Zusammenfassung der Vorexperimente 2.1, 2.2 und 2.3**

Anhand der drei oben beschriebenen Vorexperimente ließ sich beobachten, dass Exposition mit zeitlich verzögerten audiovisuellen Stimuli erst ab einer Dauer von drei Minuten Nacheffekte hinsichtlich der Zeitlichkeitsbeurteilung bewirkte, dass Re-Adaptationsdurchgänge, im Vergleich zur bloßen Stimulusexposition, ohne Re-Adaptation die Einheitlichkeitsannahme der Probanden verminderten, bzw. die Sensitivität gegenüber Asynchronizität erhöhten und diese mit kürzer werdenden Abständen zwischen den Re-Adaptationsdurchgängen zunahm, und eine Rekalibrierung der audiovisuellen Zeitwahrnehmung in Abhängigkeit der Zeitverzögerungen zwischen Licht und Ton, denen Probanden zuvor ausgesetzt wurden geschah, wenn das System vor jeder Beurteilung ‚re-adaptiert‘ wurde und die Re-Adaptation dieselbe Dauer bzw. Länge

hatte wie die zu bewertenden Reize. Insgesamt wurde damit gezeigt, dass audiovisuelle Wahrnehmung dynamisch ist und bei der Reizverarbeitung sowohl unmittelbare als auch mittelbare Information einbezieht. Dabei spielen die Parameter Dauer der vorherigen Reizexposition, Länge des Teststimulus und An- bzw. Abwesenheit von Re-Adaptation eine wichtige Rolle. Im Folgenden sollen neuronale Grundlagen dieser Dynamik der audiovisuellen Zeitlichkeitsverarbeitung anhand einer fMRT-Studie untersucht werden.

### 3.2.7 Hauptexperiment

#### 3.2.7.1 Einführung

Anhand dieser fMRT-Untersuchung soll ermittelt werden, welche kortikalen Gebiete aufgrund einer Stimulusexposition mit audiovisuell zeitlich verzögerten Reizen ihre neuronalen Antwortmuster über die Dauer der Exposition verändern. Da durch die oben aufgeführten psychophysischen Vorexperimente gezeigt werden konnte, dass durch Verzögerungen zwischen Licht und Ton während vorhergehender Stimulusexposition die audiovisuelle Zeitwahrnehmung rekaliert wird und die Nacheffekte der Exposition durch Re-Adaptation über die Länge der Abfragephasen aufrecht erhalten wird, ist zu vermuten, dass neuronale Modulationen dieser dynamischen Zeitkorrektur des perzeptuellen Apparates mit Hilfe eines fMRT-Paradigmas ermittelt werden können. Dementsprechend sollen neuronale Aktivierungen zu Anfang mit solchen gegen Ende der Expositionsphase verglichen werden. Die entstehenden Aktivierungskarten werden Areale lokalisieren, die in Zusammenhang mit andauernder Stimulusexposition bei verschiedenen Zeitverzögerungen zwischen Licht und Ton stehen. Vermutlich werden solche Areale durch zeitliche Adaptationsmechanismen moduliert und können demzufolge, neben wahrnehmungsabhängigen, mit stimuluspezifischen Prozessen audiovisueller Reizverarbeitung assoziiert werden. Diese Ergebnisse sollen später mit Arealen maskiert werden, die perzeptuelle Mechanismen audiovisueller Gleichzeitigkeitswahrnehmung regulieren (Ergebnisse aus Experiment 1), um Erkenntnisse über Verarbeitungshierarchien von audiovisueller Wahrnehmung (Experiment 1) und stimuluspezifischen Prozessen (hier: dynamische Regulierung von Zeitlichkeit) zu gewinnen. Dadurch, dass durch die vorliegende Studie Areale untersucht werden sollen, die nachweislich an ‚automatischen‘ Prozessen audiovisueller

Zeitverarbeitung beteiligt sind (s. Experiment 1) können innerhalb dieser Gebiete Verarbeitungsebenen identifiziert werden, die flexibel auf audiovisuelle Zeitunterschiede reagieren. Dies kann Aufschluss darüber geben, ob diese Gebiete sind, die dem STS in der Verarbeitungshierarchie folgen oder vorausgehen. Damit verbunden ist der Einblick, ab welcher Stufe audiovisueller Verarbeitung adaptive Prozesse eingreifen können. Dies gäbe präzisen Aufschluss über die Funktionsweise crossmodaler kortikaler Netzwerke.

### 3.2.7.2 Methodik

Sechszwanzig rechtshändige (Selbstbericht) Probanden (Durchschnittsalter: 25,6; Altersspanne: 21-35; 11 weiblich) ohne psychiatrische und neurologische Störungen nahmen an diesem Experiment teil. Sie wurden in einem 3 Tesla Trio-Scanner (Siemens, Erlangen) untersucht, der mit Lichtleiterkabeln, die das Signal weißer, roter und grüner LEDs in den Scanner leiteten, ausgestattet wurde. Alle Teilnehmer berichteten, normales Hörvermögen und normales oder korrigiertes Sehvermögen zu haben. Sie variierten hinsichtlich ihrer Vorerfahrung mit der Teilnahme an fMRT-Studien, gaben schriftlich ihr Einverständnis, an der Untersuchung teilzunehmen und wurden über deren Zweck aufgeklärt. Das Experiment dauerte ca. 70 Minuten plus zusätzlich der Zeit, die die Aufnahme struktureller Bilder benötigte (ca. 10 Minuten).

Ein Experimentaldurchlauf bestand aus einer dreiminütigen Expositionsphase und einer Abfragephase, in der Probanden mittels zweier MR-tauglichen Taster an der linken und rechten Hand Stimuli hinsichtlich ihrer Synchronizität bewerteten. Ton und Licht der Expositionsphase konnten um  $\pm 180^\circ$  jeweils in Richtung des Licht- oder Tonsignals oder um  $0^\circ$  verschoben sein. Dies entspricht einer Verzögerung von  $\pm 125$  ms (jeweils Ton zuerst und Licht zuerst) und 0 ms (synchron). Im Vergleich zu den Vorexperimenten wurden Bedingungen entfernt, bei denen Licht und Ton um  $90^\circ$  bzw. 63 ms gegeneinander verschoben waren, um das Gesamtexperiment zu verkürzen, damit Probanden nicht länger als 80 Minuten im Scanner lagen. Zudem soll das Hauptaugenmerk in bezug auf die Datenauswertung auf die Expositionsphase gelegt werden, da vermutet wird, dass die Rekalibrierung hauptsächlich während dieser Zeit stattfindet. Indem die Anzahl der Teststimuli verringert wurde, konnte die Anzahl der Expositionsphasen erhöht werden, was zu einer soliden Datenlage für die Auswertung der fMRT-Daten führt. Die Teststimuli sollen hauptsächlich zur Berechnung

psychophysischer Parameter herangezogen werden, aufgrund ihrer geringen Anzahl werden sie jedoch nicht separat untersucht. Probanden wurden dementsprechend drei verschiedenen zeitlichen Verzögerungen zwischen Licht und Ton ausgesetzt. Die Expositionsphase bestand aus einer Aneinanderreihung audiovisueller Ereignisse, die mit einer Frequenz von 4 Hz präsentiert wurden. Die visuellen Stimuli bestanden aus roten, grünen und weißen LEDs. Ein weißer Fixationslichtpunkt wurde auf Augenhöhe präsentiert und das rote Ziellicht  $1^\circ$  Sehwinkel darüber ( $60 \text{ cd/m}^2$ ). Jedes auditorische Ereignis war ein deutlich hörbarer Ton, der per Kopfhörer (MR-tauglich, ohne eigenes Magnetfeld) bei 500 Hz präsentiert wurde. Im Gegensatz zu Experiment 1 wurde ein Kopfhörer verwendet, um eine kontinuierliche auditorische Stimulation zu gewährleisten. Licht- und Tonsignal wurden jeweils mit einer Dauer von 30 ms dargeboten. Während der Expositionsphasen wurden zwei Mal pro Minute grüne statt rote Lichtreize dargeboten, die die Probanden per Knopfdruck des Tasters der linken Hand entdecken sollten. Diese sogenannten „Catch-trials“ wurden verwendet, um im Nachhinein zu überprüfen, ob Probanden der Expositionsphase folgten. Nach der Expositionsphase erlosch das Fixationslicht für 1000 ms, um Probanden einen Aufgabenwechsel zu signalisieren und es wurden anschließend 21 Testsequenzen, die aus vier aneinandergereihten audiovisuellen Ereignissen bestanden präsentiert und die die Probanden per Knopfdruck des Zeige- und Mittelfingers der rechten Hand als synchron oder asynchron beurteilen sollten (Synchronizitätsurteil). Nach jeder Testsequenz folgte eine Re-Adaptationsequenz, die ebenfalls aus vier audiovisuellen Ereignissen bestand und in der die Verzögerungsbedingung der jeweiligen Expositionsphase dargeboten wurde. Die Testsequenzen konnten, genau wie die Expositionsstimuli, um  $180^\circ$  oder  $0^\circ$  phasisch gegeneinander in beide Richtungen verschoben sein. Zusätzlich wurden pro Durchgang randomisiert sieben Ereignisse eingefügt, an denen kein Stimulus präsentiert wurde („blank-trials“), um eine zu starke Überlappung der BOLD-Antworten auf die Stimuli zu verhindern. Bei dem hier vorliegenden Experiment handelt es sich folglich um ein  $3 \times 3$  – Design mit drei in ihrer zeitlichen Verzögerung verschiedenen Expositionsphasen und drei unterschiedlich zeitlich verzögerten Teststimuli. Die Zeit zwischen zwei Stimulussequenzen (ISI) konnte 3000 ms bis 15000 ms betragen und durchschnittlich sieben Sekunden (eine Ereignis dauerte mit Re-Adaptation durchschnittlich vier Sekunden). Stimuli konnten zufällig in einem Bereich von 500 ms bis 1500 ms nach dem

vorhergehenden Stimulus dargeboten werden (Jittering). Dieser Zeitraum konnte von den Probanden für das Bewerten einer Testsequenz genutzt werden. Insgesamt wurden 15 Durchgänge (und Expositionsphasen) à 21 Testsequenzen dargeboten, so dass jede der drei Verzögerungsbedingungen während der Expositionsphasen fünf Mal und jede Verzögerungsbedingung der Testsequenzen 105 Mal präsentiert werden konnte. Pro Verzögerungsbedingung der Exposition konnte je eine Verzögerungsbedingung der Testsequenzen 35 Mal gezeigt werden (s. Tab. 5). Verzögerungsbedingungen der Testsequenzen und Expositionsphasen wurden randomisiert dargeboten. Probanden wurden instruiert, einen Stimulus nach seiner Darbietung zu bewerten, nicht aber während einer Sequenz und nicht, nachdem die nächste Sequenz bereits präsentiert würde. Vor jedem Experiment wurde ein Übungsdurchlauf durchgeführt, während dem alle möglichen Stimuluskombinationen und eine verkürzte Variante der Expositionsphase dargeboten wurden, während der Zielreize detektiert werden mussten. Der Versuchsleiter gab während des Übungsdurchgangs Rückmeldung über korrekte bzw. falsche Urteile. Nachdem die Probanden synchrone und asynchrone Reize jeweils mindestens zweimal korrekt detektiert hatten, wurde mit dem Hauptexperiment begonnen, andernfalls der Übungsdurchgang wiederholt.

**Tabelle 5.** Tabellarische Anordnung der Durchgänge und Experimentalbedingungen von Experiment 2. Es wurden 7 Stimuli pro Durchgang (Spalten) und SOA der Testsequenzen (Zeilen) dargeboten. Insgesamt wurde diese Anordnung 15 Mal wiederholt, wobei sich jedes SOA der Expositionsbedingung fünf Mal in randomisierter Reihenfolge wiederholte. Je SOA der Expositionsbedingung und SOA der Testsequenz wurden folglich 35 Reize (z. B. für SOA -125 ms der Testsequenz die ersten fünf Zellen der ersten Zeile). Durchgänge sind in der Tabelle aus Gründen der Übersicht bestimmten SOAs der Expositionsbedingung zugeordnet. Die Abfolge verschiedener Expositionsbedingungen war zufällig.

		Exposition															
Durchgang		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Testsequenzen	SOA (ms)	-125	-125	-125	-125	-125	0	0	0	0	0	125	125	125	125	125	
	-125	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	105
	0	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	105
	125	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	105
		21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	

Die Re-Adaptationsdurchgänge bestanden, wie die Testdurchgänge, aus Sequenzen mit vier auditorischen und visuellen Ereignissen und wurden abwechselnd mit den Testsequenzen präsentiert. Damit Probanden zwischen Re-Adaptations- und Testsequenzen zuverlässig unterscheiden konnten, leuchtete 500 ms vor bis 200 ms nach einer Testsequenz eine weiße LED rechts unterhalb ( $5^\circ$  vertikal,  $3^\circ$  horizontal) der Zielreiz-LEDs auf und signalisierte den Probanden, dass sie die dargebotene Sequenz beurteilen sollen. Leuchtete die weiße LED nicht auf, handelte es sich um eine Re-Adaptationssequenz, die nur betrachtet, und nicht bewertet werden sollte.

Die fMRT-Daten wurden mit einer 8-Kanal-Kopfspule (Bruker Bio-BioSpinMRI, Ettlingen, Deutschland) gemessen. Für die funktionellen Messungen wurde eine  $T_2^*$ -gewichtete Echoplanar-Sequenz (echo planar imaging, EPI) verwendet [räumliche Auflösung:  $3.5 \times 3.5 \times 4$  mm,  $64 \times 64 \times 32$  Schichten, TR: 2000 ms, TE: 30 ms, flip angle:  $80^\circ$ , abwechselnd (interleaved) aufgenommen]. Zusätzlich wurde ein  $T_1$ -gewichtetes inversion recovery-EPI aufgenommen, das dieselben Parameter und damit räumlichen Verzerrungen wie die funktionellen EPI-Messungen aufwies (räumliche Auflösung:  $3.5 \times 3.5 \times 4$  mm,  $64 \times 64 \times 32$  Schichten, TR: 2500 ms, TE: 17 ms, TI: 1200 ms, flip angle:  $80^\circ$ ). Dieses IR-EPI wurde zur Bestimmung der Normalisierungsparameter verwendet. Die ersten fünf Volumen wurden aus der Analyse ausgeschlossen, um eine vergleichbare Sättigung für alle Volumen zu erreichen.

Augenbewegungen wurden anhand eines MRT-kompatiblen Aufnahmesystems (Kanowski et al., 2007) und einer Software (PupilTracker, HumanScan, Erlangen) überwacht. Die Probanden trugen Ohrstöpsel; die wahrgenommene Lautstärke und Balance wurden individuell angepasst, um sicher zu stellen, dass die Sequenzen klar wahrgenommen werden konnten.

Die Bedingungen des Hauptexperiments wurden mit hämodynamischen Antwortfunktionen modelliert [gemischtes Design (Block- und ereigniskorreliertes Design), Vorverarbeitung: Bewegungskorrektur, zeitliche Korrektur der Schichtaufnahme, Normalisierung auf segmentierte weiße Substanz, Glättung:  $6 \times 6 \times 6$  mm] und zusätzlich Bewegungsparameter, Zeit- und Verteilungsableitungen geschätzt.

Für die Analyse der experimentellen Durchgänge wurden 27 Bedingungen definiert: zwei Antwortmöglichkeiten (synchron, asynchron) pro Bedingung (eine Bedingung für Ton zuerst, eine für Licht zuerst und eine für synchrone Sequenzen) für

jeweils jede der drei Abfragebedingungen als ereigniskorrelierte Bedingungen (18) und zusätzlich jede der drei Minuten für jeweils jede der drei Expositionsbedingungen als Blockbedingung (9).

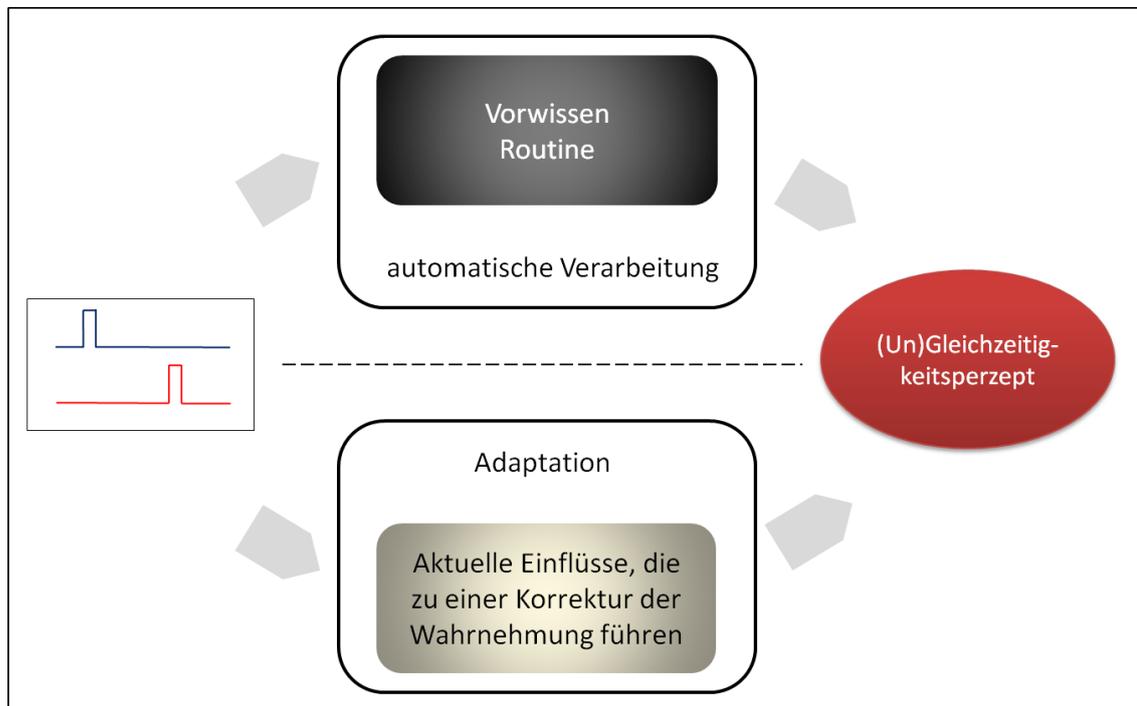
Aufgrund der geringen Anzahl von Durchgängen und dem damit verbundenen geringen Auftreten einiger Bedingungen wurden alle 15 Experimentaldurchgänge für die Datenauswertung des fMRT-Experiments zusammengeführt. Daher wurden zusätzlich 15 Regressoren modelliert, um irrelevante Varianz, die durch die getrennt durchgeführten Sitzungen entstanden sein kann, von der Varianz der Experimentalbedingungen zu trennen. Für die Analyse wurden die Daten voxelweise auf ihren globalen Mittelwert normiert, sowie hoch- und tiefpassgefiltert.

Um Effekte für die Bedingungen statistisch zu testen, wurden lineare Kontraste für jeden Probanden erstellt. Die resultierenden Kontrastbilder wurden für eine Analyse auf zweiter Ebene genutzt (Varianzanalyse, random-effects-Analyse). Die Schwellen für die parametrischen Karten der T-Statistiken für jeden Voxel wurden bei  $p < 0,05$  und  $k > 20$  Voxel gesetzt. Dabei wurden Kontraste erstellt, die den Block der dritten gegen den der ersten Minute der Expositionsphase vergleichen und umgekehrt. Dies geschah jeweils für asynchrone Stimuli [-125 ms (Ton voran) und 125 ms (Licht voran) zusammengefasst] und synchrone Stimuli während der Exposition. Da anhand der oben aufgeführten Vorexperimente 2.1 und 2.2 nachgewiesen wurde, dass sich vorherige Stimulusexposition in Abhängigkeit der Verzögerung zwischen Licht- und Tonreiz, der Probanden ausgesetzt werden, auf die darauffolgende Synchronizitätswahrnehmung auswirkt, wird angenommen dass Exposition bei asynchronen Stimuli zu einer ‚Synchronisierung‘ der Reizwahrnehmung führt. D.h. dass nach asynchroner Exposition (asynchrone) Stimuli häufiger als synchron bewertet würden als ohne Exposition. Die in den Vorexperimenten ermittelte Verschiebung der PsSS in die Richtung der Verzögerungsrichtung während der Expositionsphase, die die Bereitschaft für Gleichzeitigkeitwahrnehmung asynchroner Stimuli erhöht, unterstützt diese Vermutung. Areale, die nach drei Minuten Exposition höhere Aktivierungen aufweisen als zu Beginn der Exposition, können vermutlich subjektive Repräsentationen zeitlicher Parameter von audiovisuellen Stimuli dergestalt modulieren, dass ein synchrones Perzept eher wahrgenommen wird als ohne Exposition. Dies sollen jedoch kortikale Gebiete sein, die nach dreiminütiger Exposition bei nur asynchronen Stimuli einen Signalanstieg verzeichnen, nicht aber nach synchroner

Exposition, da nach dieser gemäß der hier aufgeführten Annahme keine ‚Synchronisierung‘ auftreten kann. Neuronale Aktivität, die einen Signalanstieg unabhängig von der Verzögerungsbedingung der Expositionsphase zeigt, könnte lediglich als Gewöhnung an die Exposition interpretiert werden. Der Kontrast, der die dritte Expositionsminute bei asynchronen Stimuli gegen die erste vergleicht, soll demzufolge mit jenem exklusive maskiert werden, der denselben Vergleich nach synchroner Exposition modelliert. Areale, die durch den umgekehrten Kontrast ermittelt werden, d.h. die während der ersten Expositionsminute bei asynchronen Stimuli eine höhere Signalstärke aufweisen als während der dritten, dieses Muster jedoch nicht nach synchroner Exposition zeigen, modulieren vermutlich das subjektive Perzept zeitlicher Komponenten von audiovisuellen Stimuli derart, dass ein asynchrones Urteil wahrscheinlich wird: Sie würden demzufolge Ereignisse ‚desynchronisieren‘. Daher soll, um diese Areale zu ermitteln, die erste Expositionsminute bei asynchronen Stimuli gegen die dritte verglichen werden und dieser Kontrast exklusiv mit demjenigen maskiert werden, der die erste gegen die dritte Minute nach synchroner Exposition vergleicht. Anschließend sollen die so ermittelten Areale, die unter anderem für stimulusabhängige Prozesse (hier: audiovisueller Zeitlichkeit) als relevant betrachtet werden können, mit Arealen maskiert werden, die durch perzeptuelle audiovisuelle (A)synchronizität moduliert werden. Diese Gebiete wurden in Experiment 1 ermittelt. Dieses Vorgehen kann Aufschluss darüber geben, ob und an welchen kortikalen Gebieten innerhalb von Arealen, die für audiovisuelle Zeitlichkeitsperzeption relevant sind, stimulusabhängige, kurzzeitig flexible Prozesse audiovisueller Zeitwahrnehmung moduliert werden. Das differenzierte Betrachten perzeptueller und stimulusabhängiger Prozesse anhand dieses fMRT-Paradigmas kann Schlüsse über Verarbeitungshierarchien bei audiovisueller Zeitwahrnehmung ermöglichen (s. Abb. 26).

Es kann vermutet werden, dass eine Adaptation auf asynchrone Stimuli, im Gegensatz zu Bedingungen, die ein Nachlassen audiovisueller Gleichzeitigkeit induzieren, unisensorische Kortices stärker moduliert, da subjektive Gleichzeitigkeit etabliert werden muss und gegebenenfalls stärkere Interaktionen auditorischer und visueller Kortices verlangt als segregierende Prozesse (s. Noesselt et al., 2007). Gemäß bisherigen Studien (Dhamala et al., 2007; Miller & D’Esposito, 2005) werden multimodale Gebiete

wie IPS oder STS wahrscheinlich unterschiedliche vermittelnde Rollen einnehmen, je nach dem, ob audiovisuelle Synchronizität hergestellt wird (Adaptation), oder nicht.



**Abbildung 26.** Schematische Darstellung des Einflusses adaptiver und automatischer Prozesse auf die zeitliche Wahrnehmung eines audiovisuellen Reizes (links im Bild). Wenn ein auditorisches und ein visuelles Ereignis beim Beobachter eintreffen, beeinflussen vermutlich Langzeiterfahrungen (schwarzer Kasten oben im Bild) und aktuelle Erfahrungen (grauer Kasten unten) die Wahrnehmung beider Reize als gleichzeitig oder asynchron. Während automatische Prozesse audiovisueller Zeitwahrnehmung in Experiment 1 untersucht wurden (oberer Pfad), sollen nun adaptive Prozesse, die kurzzeitig die Gleichzeitigkeitswahrnehmung beeinflussen, untersucht werden (unterer Pfad).

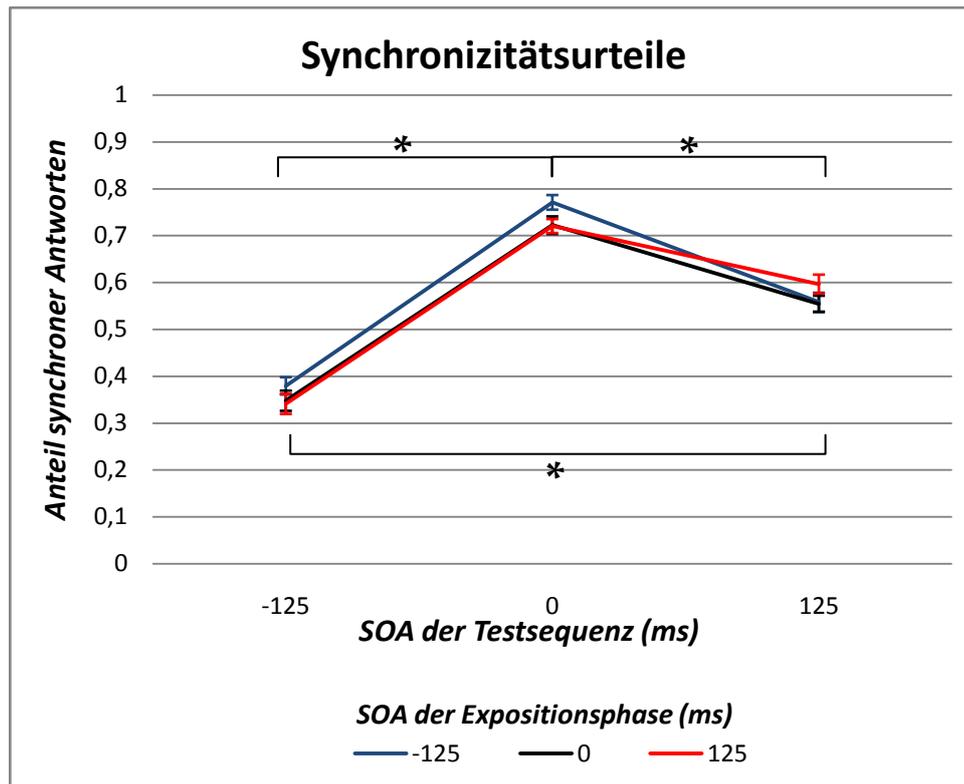
### 3.2.7.3 Ergebnisse

#### 3.2.7.3.1 Verhaltensdaten

##### 3.2.7.3.1.1 Synchronizitätsurteile

Aus den aufgezeichneten Synchronizitätsurteilen wurden diejenigen Antworten, die während der Darbietung bis 200 ms nach Ende einer Testsequenz gegeben wurden, nicht in die Analyse einbezogen. Durchschnittlich bewerteten die Probanden einen Reiz in 36% (4%), 74% (3%) bzw. 57% (4%) aller Präsentationen als synchron wenn Licht und Ton jeweils um -125 ms (Ton voran), 0 ms (synchron) und 125 ms (Licht voran) gegeneinander verschoben waren (s. Abb. 27). Nach Exposition mit um -125 ms (Ton voran) verschobenen Stimuli bewerteten die Teilnehmer die präsentierten Stimuli in 57% (2%),

nach Exposition mit Stimuli ohne Verzögerungen (synchron) in 54 % (3%) und nach Exposition mit um 125 ms (Licht voran) verzögerten Stimuli in 55% (3%) der Fälle als synchron (Standardfehler in Klammern). Mit diesen Daten wurde eine 3 x 3-faktorielle Varianzanalyse mit Messwiederholung mit den Faktoren *Zeitverzögerung der Exposition* und *Zeitverzögerung der Testsequenzen* durchgeführt. Dies soll Aufschluss darüber geben, ob sich die Synchronizitätsurteile in Abhängigkeit der präsentierten Verzögerungsbedingungen und der Verzögerungen während der Expositionsphase ändern. Ein Haupteffekt wurde für *Zeitverzögerung der Testsequenzen* gefunden [ $F(2;50) = 37$ ;  $p < 0,00$ ;  $\eta^2 = 0,60$ ]. Für *Zeitverzögerung der Exposition* wurden keine signifikanten Effekte gefunden [ $F(2;50) = 2,3$ ;  $p = 0,12$ ;  $\eta^2 = 0,08$ ] und für die Interaktion aus beiden Faktoren ein statistischer Trend [ $F(4;100) = 2,15$ ;  $p = 0,08$ ;  $\eta^2 = 0,08$ ]. Nach Bonferroni korrigierte post-hoc durchgeführte T-Tests zeigten, dass synchrone Stimuli signifikant am häufigsten als synchron beurteilt wurden ( $p < 0,001$ ) und um 125 ms (Licht voran) verzögerte Stimuli signifikant häufiger als um -125 ms (Ton voran) verzögerte Stimuli ( $p < 0,001$ ).

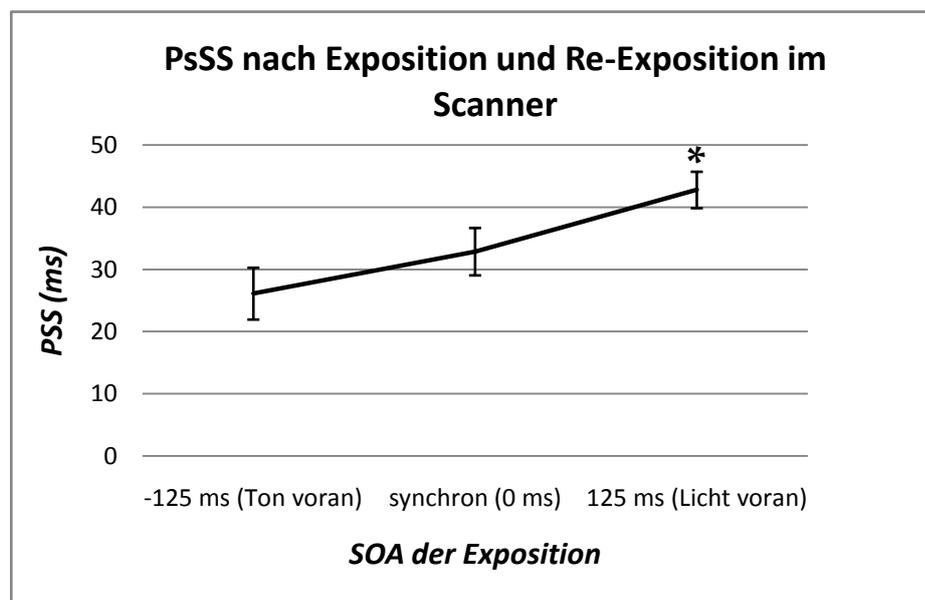


**Abbildung 27.** Anteil synchroner Antworten in Abhängigkeit des präsentierten SOAs der Testsequenz und des SOAs der Expositionsphase in Experiment 2. Signifikante Unterschiede der Urteile, die mittels einer 3 × 3-stufigen Varianzanalyse ermittelt wurden, sind durch \* gekennzeichnet ( $p < 0,05$ ). Dargestellte Unterschiede beziehen sich auf Datenpunkte, die über alle SOAs der Expositionsphase kollabiert wurden (s. Text). Die Fehlerbalken stellen Standardfehler dar. Antworten unterschieden sich auf dem Faktor *SOA der Testsequenz* (s. Text). Ein statistischer Trend für die Interaktion aus *SOA der Testsequenz* und *SOA der Expositionsphase* war erkennbar.

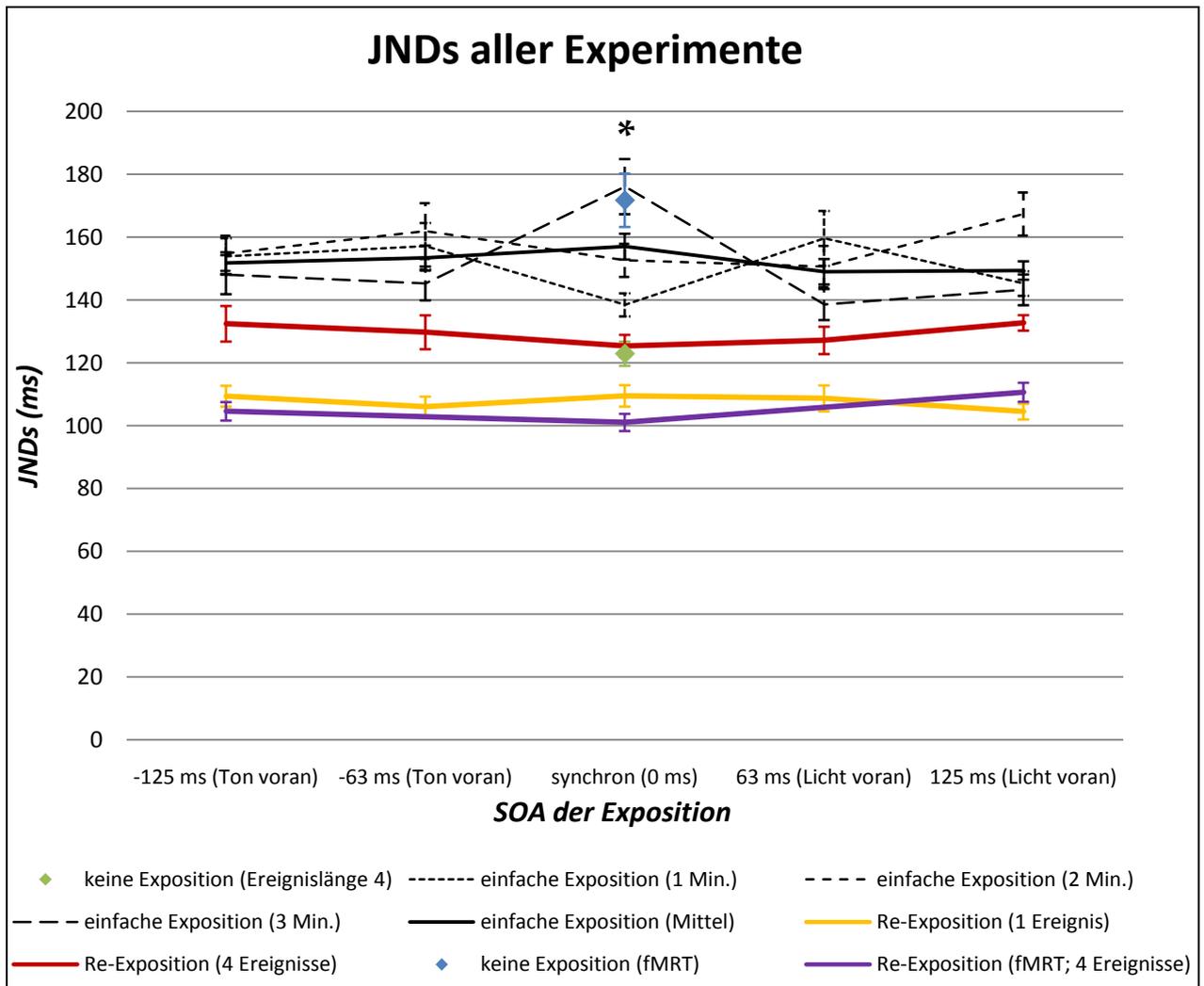
### 3.2.7.3.1.2 JNDs und PsSS

Aus den Synchronizitätsurteilen wurden JNDs und PsSS berechnet und mit diesen Daten jeweils eine 3-stufige Varianzanalyse mit Messwiederholung mit dem Faktor *Zeitverzögerung der Exposition* durchgeführt, um zu überprüfen, ob sich die Verzögerungsbedingung der Expositionsphase auf die Gleichzeitigkeitswahrnehmung der Teilnehmer signifikant auswirkt. Aus diesen Werten wurden standardisierte Z-Werte berechnet und alle Probanden, deren Datenpunkte zwei Standardabweichungen über- oder unterschritten, eliminiert. Zusätzlich wurden Fälle entfernt, die größere PsSS als das maximale SOA (125 ms) oder größere JNDs als das Doppelte des maximalen SOAs (250 ms) aufwiesen. Probanden, die während der Expositionsphase weniger als 80% der „catch-trials“ entdeckt hatten und mehr als 20% falsch positive Antworten gaben, wurden

ebenfalls aus der Analyse entfernt. Im Durchschnitt betragen die JNDs 105 ms (6 ms), 101 ms (5 ms) bzw. 111 ms (6 ms) nach Exposition mit jeweils um -125 ms (Ton voran), 0 ms (synchron) und 125 ms (Licht voran) verzögerten Stimuli. Es konnte kein signifikanter Effekt für diese Daten beobachtet werden [ $F(2;24) = 1,5$ ;  $p = 0,25$ ;  $\eta^2 = 0,11$ ]. PsSS betragen durchschnittlich 26 ms (8 ms), 33 ms (8 ms) bzw. 43 ms (6 ms) für Stimuli, die während der Expositionsphase um -125 ms (Ton voran), um 0 ms (synchron) bzw. 125 ms (Licht voran) verzögert waren (s. Abb. 28). Für diesen Anstieg ließ sich ein signifikanter Effekt nachweisen [ $F(2;24) = 5,2$ ;  $p < 0,05$ ;  $\eta^2 = 0,30$ ]. Nach Bonferroni korrigierte post-hoc durchgeführte T-Tests zeigten, dass der PSS nach Exposition bei um -125 ms (Ton voran) verzögerten Stimuli signifikant niedriger war als der nach Exposition mit Stimuli, die um 125 ms (Licht voran) verzögert waren ( $p < 0,05$ ). Ein Vergleich zwischen dem PSS nach Exposition mit synchronen Stimuli und dem nach Exposition mit Stimuli, die um 125 ms (Licht voran) verzögert waren, ergab, dass letzterer im statistischen Trend höher war ( $p = 0,08$ ).

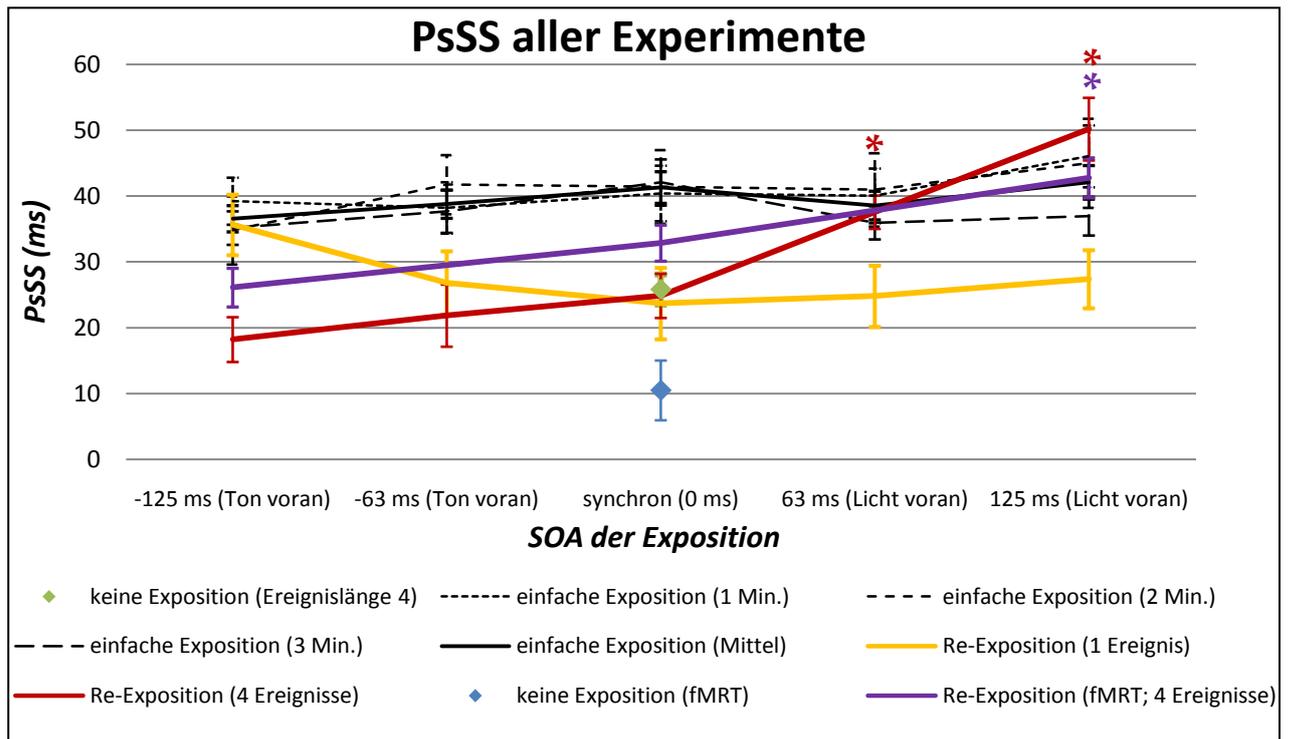


**Abbildung 28.** Die Grafik zeigt PsSS des Experiments 2 in Abhängigkeit des SOAs der Expositionsphase. PsSS nach Exposition bei verschiedenen zeitlichen Verzögerungen zwischen Licht und Ton veränderten sich. Eine 3-stufige Varianzanalyse wurde durchgeführt. Signifikante Unterschiede der PsSS traten nach Exposition mit Stimuli auf, die um 125 ms (Licht voran) phasenverschoben dargeboten wurden. Fehlerbalken stellen Standardfehler dar. \* stellen signifikante Unterschiede der Daten zwischen den SOAs der Expositionsphase dar.



**Abbildung 29.** JNDs aus Vorexperiment 1 bei Ereignislänge 4 (grün), Vorexperiment 2.1 (schwarz), Vorexperiment 2.2 (rot), Vorexperiment 2.3 (gelb), fMRT-Experiment 1 (blau), und fMRT-Experiment 2 (violett) in Abhängigkeit des SOAs der Expositionsphase (sofern durchgeführt). Die durchgezogene, schwarze Kurve stellt das arithmetische Mittel über alle Expositionszeiten (1,2, und 3 Minuten) der JNDs aus Vorexperiment 2.1 dar (für Daten nach Expositionszeit geordnet s. gestrichelte Kurven). Außer bei Vorexperiment 2.1 (drei Minuten Expositionszeit) konnten keine experimentellen Effekte ermittelt werden.

Abbildungen 24 und 25 geben eine Übersicht über alle ermittelten JNDs und PsSS der hier vorgestellten Experimente.



**Abbildung 30.** PsSS aus Vorexperiment 1 bei Ereignislänge 4 (grün), Vorexperiment 2.1 (schwarz), Vorexperiment 2.2 (rot), Vorexperiment 2.3 (gelb), fMRT-Experiment 1 (blau), und fMRT-Experiment 2 (violett) in Abhängigkeit des SOAs der Expositionsphase (sofern durchgeführt). Die durchgezogene, schwarze Kurve stellt das arithmetische Mittel über alle Expositions-längen (1,2, und 3 Minuten) der PsSS aus Vorexperiment 2.1 dar (für Daten nach Expositionsdauer geordnet s. gestrichelte Kurven). Experimentelle Effekte wurden bei Vorexperiment 2.2 (Re-Adaptation mit 4 Ereignissen) und fMRT-Experiment 2 ermittelt.

### 3.2.7.3.2 fMRT-Daten

#### 3.2.7.3.2.1 Experimentelle Kontraste

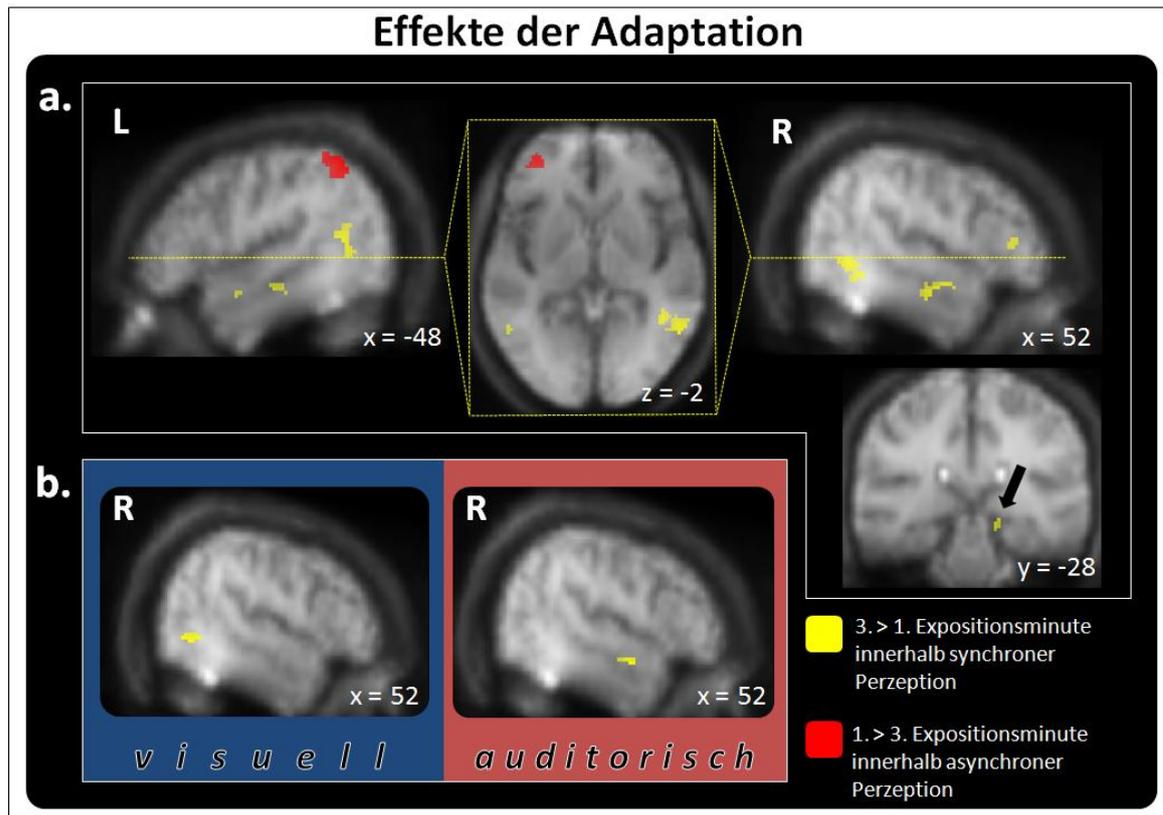
##### 3.2.7.3.2.1.1 3.>1. Minute bei asynchroner Exposition innerhalb perzeptueller Synchronizität

Wie bereits in 3.2.7.2 erwähnt, richtete sich das Hauptaugenmerk dieser Analyse auf Areale innerhalb von bimodalen Gebieten, die in Experiment 1 durch perzeptuelle Synchronizität moduliert wurden, und die nach dreiminütiger Exposition mit asynchronen Stimuli stärker aktiviert waren als während der ersten Expositionsminute, aber nicht nach Exposition mit synchronen Stimuli [s. 3.2.7.1; der entsprechende Kontrast (3. Minute > 1. Minute bei synchroner Exposition) wurde mit dem vorliegenden Vergleich exklusive maskiert]. Areale, die durch diesen Kontrast moduliert wurden, waren bilateral medial und lateral temporale (perirhinaler Kortex/Gyrus parahippocampalis und STS, STG, MTG) und frontale Regionen (ACC, IFG, MFG;  $p < 0,05$ ;  $k > 20$ ; s. Anhang, Tab. 11 c und Abb. 31 a). Diese Gebiete könnten zu einer subjektiven ‚Synchronisierung‘

wahrgenommener Stimuli beigetragen haben (s. 3.2.1). Innerhalb unimodaler Areale, die durch perzeptuelle Synchronizität in Experiment 1 moduliert wurden, waren der rechte anteriore STS (auditorisch lokalisiert) und die rechte temporo-okzipitale Verbindung (visuell lokalisiert) während der dritten Minute nach asynchroner Exposition stärker aktiviert als zu Beginn der Exposition (s. Abb. 31 b, gelb eingezeichnete Modulationen). Die Grundlage dieser Kontraste bilden die Befunde aus den Vorexperimenten 2.1 und 2.2, anhand derer demonstriert wurde, dass sich nach dreiminütiger Exposition mit phasenverschobenen audiovisuellen Sequenzen die Gleichzeitigkeitswahrnehmung im Vergleich zu phasengleichen Reizen verändert.

#### 3.2.7.3.2.1.2 1.>3. Minute bei asynchroner Exposition innerhalb perzeptueller Asynchronizität

Areale innerhalb von Gebieten, die in Experiment 1 durch perzeptuelle Asynchronizität moduliert wurden, und die während der ersten Expositionsminute höhere BOLD-Antworten aufwiesen als während der dritten und dieses Muster nur bei Exposition mit asynchronen Stimuli zeigten, nicht aber nach Exposition mit synchronen [der entsprechende Kontrast (1. Minute > 3. Minute bei synchroner Exposition) wurde mit dem vorliegenden exklusive maskiert], waren der linke Parietalkortex (IPS, SMG, Gyrus angularis) und der linke MFG ( $p < 0,05$ ;  $k > 20$ ; s. Anhang, Tab. 11 d und Abb. 31 a, rot eingezeichnete Modulationen). Die durch diesen Vergleich modulierten Gebiete könnten zu einer subjektiven ‚Desynchronisierung‘ wahrgenommener audiovisueller Reize beigetragen haben (s. 3.2.1).



**Abbildung 31.** Aktivierungskarten, die neuronale Effekte dreiminütiger Exposition mit asynchronen Stimuli abbilden. **a.** Dabei sind solche Areale innerhalb derjenigen Gebiete, die in Experiment 1 durch perzeptuelle Synchronizität moduliert wurden, gekennzeichnet, die während der dritten Expositionsminute stärker neuronal moduliert wurden, als während der ersten (gelb). Dies waren innerhalb bimodaler Gebiete von Experiment 1 der rechte perirhinale Kortex bzw. Gyrus parahippocampalis, bilateraler aSTS, pSTS, MTG und rechter PFC (s. Tab. 11 c). Darüber hinaus sind solche Areale rot dargestellt, die innerhalb von Arealen perzeptueller Asynchronizität aus Experiment 1 während der ersten Expositionsminute stärker aktiviert waren als während der dritten. Dies waren der linke Parietalkortex und der PFC (s. Tab. 11 d). **b.** Innerhalb unisensorischer (visuell: blau hinterlegt; auditorisch: rot hinterlegt) wurden rechter aSTS und ITG durch ‚3. Minute > 1. Minute‘ innerhalb perzeptueller Synchronizität aus Experiment 1 moduliert. Während die gelben Hervorhebungen Modulationen darstellen, die wahrscheinlich mit einer Verschiebung subjektiver Gleichzeitigkeit einhergehen (s. Vorexperiment 2.2), spiegeln rot markierte Modulationen wahrscheinlich Effekte subjektiver Segregation audiovisueller Informationen wider. Diese Kontraste schließen Modulationen durch synchrone Exposition aus. S. Abb. 32/33.

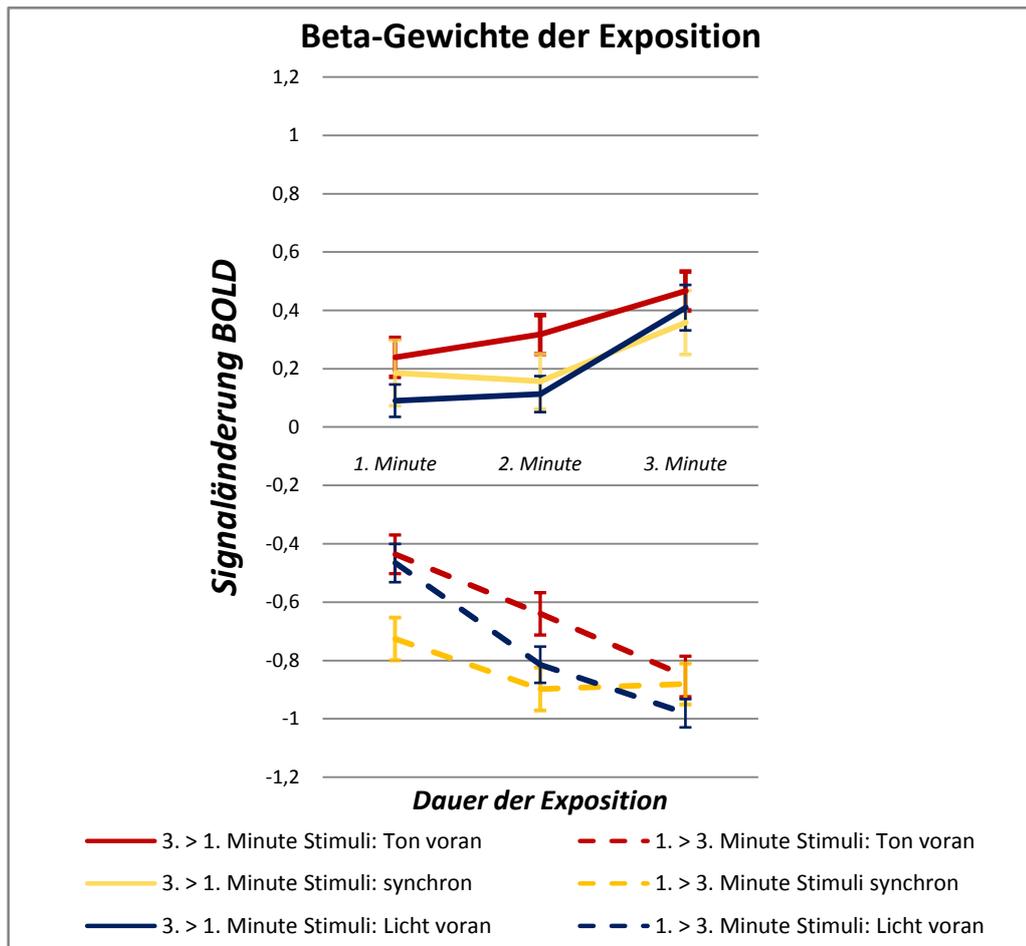
### 3.2.7.3.2.1.3 Vergleich der Expositionsphase außerhalb perzeptueller Areale

Areale, die außerhalb der perzeptuellen Masken von Experiment 1 nach dreiminütiger Exposition mit asynchronen Stimuli stärker aktiviert waren, als während der ersten Expositionsminute, aber nicht nach Exposition mit synchronen Stimuli waren lateral und medial temporale, präfrontale, parietale, okzipitale und cerebelläre Gebiete (s. Anhang, Tab. 11 a). Dieser Kontrast ergab ein breites Netz bilateraler kortikaler Aktivität. Gebiete, die außerhalb der perzeptuellen Masken von Experiment 1 während der ersten

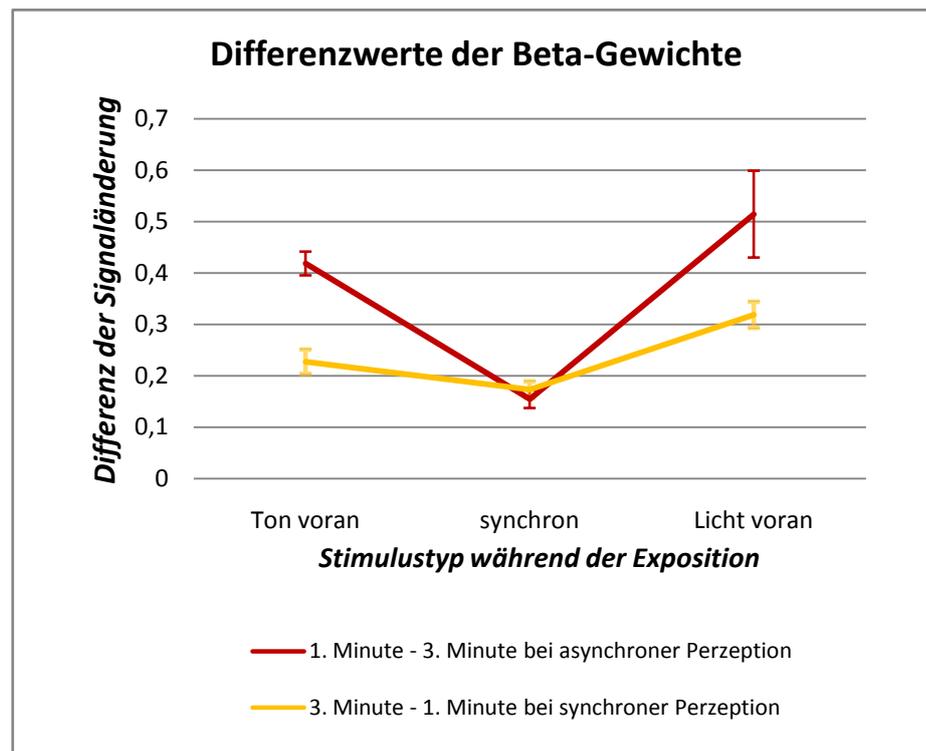
Expositionsminute höhere BOLD-Antworten aufwiesen als während der dritten und dieses Muster nur bei Exposition mit asynchronen Stimuli zeigten, nicht aber nach Exposition mit synchronen, waren linke lateral und medial temporale (Gyrus parahippocampalis, STG), präfrontale (MFG, IFG, Gyrus praecentralis), parietale (IPS, SMG) und cerebelläre Regionen (s. Anhang, Tab. 11 b). In geringerem Maß wurden rechtslaterale Gebiete moduliert, darunter der MFG, der Gyrus praecentralis und das Cerebellum. Zu Beginn der Expositionsphase mit asynchronen Sequenzen waren demzufolge vorwiegend linkstemporale Gebiete stärker aktiviert, als zum Ende der Expositionsphase.

#### 3.2.7.3.2.1.4 BOLD-Signale

Für die lokalen Maxima der oben aufgeführten Areale wurden Beta-Gewichte ermittelt und deren Mittelwerte getrennt nach Vergleich (1. > 3. Minute vs. 3. > 1. Minute) in Abbildung 26 dargestellt. Beta-Gewichte stellen die Änderung des BOLD-Signals für einzelne Voxel dar. Aus der Abbildung wird ersichtlich, dass die neuronale Antwort in den ermittelten Arealen nach Exposition mit asynchronen Stimuli signifikant stärker an- bzw. abstieg als nach synchroner Exposition. Abbildung 27 verdeutlicht dieses Muster und zeigt, dass die Unterschiede von der ersten Expositionsminute zur dritten in bezug auf die neuronale Antwort in den aufgezeigten Arealen größer waren, wenn Probanden asynchronen Stimuli ausgesetzt wurden, als im Vergleich zur Antwortänderung nach synchroner Exposition. Dieser Signalunterschied von der ersten zur dritten Expositionsminute war jedoch für den Vergleich ‚1. > 3. Minute‘ deutlicher ausgeprägt als für ‚3. > 1. Minute‘. Für beide Kontraste stieg bzw. sank die neuronale Antwort auch nach synchroner Exposition; dadurch, dass der Vergleich ‚3. > 1. Minute‘ bzw. ‚1. > 3. Minute‘ bei synchroner Exposition exklusive mit den aufgeführten Kontrasten maskiert wurde, wurde jedoch sichergestellt, dass sich die neuronale Antwort nach der dritten Expositionsminute statistisch nicht signifikant von der Antwort während der ersten Expositionsminute unterschied, wenn Probanden synchronen Stimuli ausgesetzt wurden. Signifikante Unterschiede im Antwortmuster wurden durch dieses Vorgehen nur nach asynchroner Exposition zugelassen.



**Abbildung 32.** Beta-Gewichte der in Experiment 2 verwendeten Kontraste. Durchgezogene Kurven stellen Änderungen der neuronalen Antworten für Kontraste dar, die höhere Antworten zu Ende der Expositionsphase (3. Minute) als zu Beginn (1. Minute) gegeneinander verglichen. Gestrichelte Linien repräsentieren Kontraste, die höhere Aktivität zu Beginn der Expositionsphase mit niedrigerer zum Ende der Exposition verglichen. Farben stellen durch den Vergleich ermittelte Signaländerungen in Abhängigkeit der drei Stimulustypen dar, die verwendet wurden (rot: Ton vor Licht, gelb: synchron, blau: Licht vor Ton). Die Verläufe der Signaländerungen in Abhängigkeit der Dauer der Expositionsphase verdeutlichen, dass sich die neuronale Aktivität während der Exposition mit asynchronen Sequenzen (Ton vor Licht und Licht vor Ton, rote und blaue Kurven) stärker veränderte, als nach Exposition mit synchronen, obwohl eine Signaländerung auch in dieser (synchronen) Bedingung stattfand. Dies impliziert, dass die berichteten Areale (s. Haupttext) vorwiegend durch adaptive Prozesse in bezug auf asynchrone Reize moduliert wurden und keine Habituation an die Stimulation reflektieren. Fehlerbalken stellen Standardfehler dar.



**Abbildung 33.** Die Datenpunkte stellen Differenzwerte der BOLD-Signale aus Abb. 26 in Abhängigkeit der verwendeten Stimulustypen (Licht vor Ton, synchron, Ton vor Licht) dar, wobei die Beta-Gewichte während der ersten Expositionsminute von denen der dritten Expositionsminute (gelbe Kurve) bzw. vice versa (rote Kurve) subtrahiert wurden. Der niedrigste Wert trat bei Exposition mit synchronen Reizen auf, was impliziert, dass Areale, die im Haupttext berichtet wurden vorwiegend durch Anpassungsprozesse auf asynchrone Reize moduliert wurden. Fehlerbalken repräsentieren Standardfehler.

### 3.2.7.4 Diskussion

#### 3.2.7.4.1 Verhaltensdaten

Die oben dargestellten Effekte zeigen, dass eine Rekalibrierung der subjektiven Gleichzeitigkeitswahrnehmung in Abhängigkeit der zuvor durchgeführten Exposition stattfand. Dies untermauert, dass die Stimuluskonfiguration, die anhand der Vorexperimente ermittelt worden war, adäquat war, um das audiovisuelle Zeitwahrnehmungssystem zu beeinflussen. Darüber hinaus konnte so belegt werden, dass die subjektive Zeitwahrnehmung trotz veränderter Experimentalbedingungen [im Vergleich zu den Vorexperimenten drei anstatt fünf Experimentalbedingungen (SOAs) während Expositions- und Testphase; Scannerumgebung] manipuliert werden konnte. Die Werte der prozentualen Anteile synchroner Antworten je Testsequenz und die PsSS

sind in ihrem Ausmaß in etwa mit denen des Vorexperiments 2.2 vergleichbar, obwohl PsSS nach Exposition mit Stimuli, bei denen der Ton dem Licht voranging, verglichen mit Vorexperiment 2.2, etwas höher waren. JNDs waren im fMRT-Experiment kleiner als im Vorexperiment 2.2. Diese Datenlage weist auf eine geringere Toleranz gegenüber asynchronen Stimuli hin und kann durch die veränderten Experimentalbedingungen zustande gekommen sein: den Probanden wurden im Vergleich zu Vorexperiment 2.2 zwei Expositionsbedingungen weniger präsentiert (-63 ms, Ton voran und 63 ms, Licht voran). Daher konnte das audiovisuelle Zeitsystem im Vergleich zu Vorexperiment 2.2 weniger zeitkritische Informationen sammeln, wodurch sich die subjektiven Zeitfenster audiovisueller Ereignisse gegebenenfalls nicht in dem Maß ausdehnten, wie dies in Vorexperiment 2.2 geschah. Ein weiterer Grund könnte die laute Umgebung im Scanner gewesen sein, durch die audiovisuelle Zeitwahrnehmung konservativer wurde, um Einflüsse von Störgeräuschen auszugleichen. Darüber hinaus standen dem Schätzalgorithmus, der aus den Gleichzeitigkeitsurteilen JNDs und PsSS schätzt weniger Datenpunkte zur Verfügung, was sich ebenfalls kritisch auf das Ausmaß der JNDs ausgewirkt haben könnte.

Im Durchschnitt stieg der PSS um 8,35 ms von einer Verschiebungsbedingung während der Expositionsphase zur nächsten. Dies macht 6,7% des Zeitintervalls zwischen von einer SOA zur nächsten aus (jeweils 125 ms). Damit war der Anstieg etwa halb so groß wie der in Vorexperiment 2.2. Dieser Sachverhalt könnte ebenfalls im Erfahren weniger zeitkritischer Informationen begründet sein, die durch die im Vergleich zu Vorexperiment 2.2 eliminierten SOAs dargeboten wurden: da audiovisuelle Zeitwahrnehmung, wie oben erläutert (s. 3.2.4.3) erfahrungsabhängig ist, könnte sich die Anzahl der zuvor dargebotenen SOAs auf das Ausmaß der Rekalibrierung auswirken. Darüber hinaus könnten die Scannergeräusche die subjektive Zuverlässigkeit während der Exposition erfahrener Zeitverzögerungen minimiert haben und damit den Bereich, innerhalb dessen eine Rekalibrierung stattfinden kann verkleinert haben. Analog zu Vorexperiment 2.2 konnte der PSS deutlicher beeinflusst werden, wenn der Lichtreiz während der Expositionsphase dem Tonreiz vorausging als umgekehrt. Dies könnte eine höhere Bereitschaft des audiovisuellen Wahrnehmungssystems andeuten, den PSS zu rekalibrieren, wenn Reize schwerer perzeptuell einzuordnen sind (zur Diskussion s. auch 3.2.4.3).

### 3.2.7.4.2 fMRT-Daten

Ziel der fMRT-Untersuchung war, Areale innerhalb von Gebieten, die zuvor mit audiovisueller Zeitwahrnehmung assoziiert werden konnten, zu identifizieren, die sich flexibel bzw. vorübergehend plastisch Zeitunterschieden zwischen zwei Informationsströmen anpassen. Die durch die Vorexperimente entwickelte, zugrundeliegende Annahme war, dass nach asynchroner Exposition, die mindestens drei Minuten andauert (s. Vorexperiment 2.1; Kap. 3.1.3.2.1) Stimuli zunehmend als synchron empfunden werden, d.h. dass Zeitunterschiede zwischen auditorischem und visuellem Informationsstrom ausgeglichen werden. Die psychophysischen Befunde, anhand derer nachgewiesen wurde, dass sich der PSS durch asynchrone Exposition richtungsabhängig verändert, unterstützen diese Annahme. Dieser Befund konnte auch innerhalb des Scanners repliziert werden. Areale, die diese zeitliche Adaptation modulieren, würden nach, so die Annahme, dreiminütiger Exposition höhere BOLD-Antworten aufweisen als zu Beginn der Expositionsphase. Entsprechend wurden zur Identifikation adaptionsrelevanter Areale SPM-Kontraste erstellt, die die dritte Expositionsminute mit der ersten verglichen. Für Areale, die zu einer synchronen Wahrnehmung beitragen, d.h. dass sich die beiden Informationsströme zeitlich angleichen wurde der lineare Kontrast ‚3. > 1. Expositionsminute‘ erstellt. Gleichzeitig wurde dieser Vergleich für Exposition mit synchronen Stimuli erstellt und mit dem asynchroner exklusive maskiert. Diesem Vorgehen liegt die Argumentation zugrunde, dass eine zeitliche Angleichung der Informationsströme bei Exposition mit physikalisch gleichzeitig stattfindenden auditorischen und visuellen Ereignissen nicht auftreten würde. Höhere neuronale Antworten in Arealen, die während der dritten Minute synchroner Exposition im Vergleich zur ersten auftreten würden, würden nicht auf einen Angleichungsprozess zeitrelevanter Informationen zurückzuführen sein. Die Maskierung sollte sicherstellen, dass nur Areale lokalisiert würden, die ihr Antwortmuster in Abhängigkeit anhaltender asynchroner Exposition verändern, nicht aber infolge synchroner. Diese Gebiete würden zu einer ‚Synchronisierung‘ zeitlich disparater Informationen beitragen. Die adaptiven Gebiete wurden daher innerhalb solcher Areale analysiert, die infolge der Befunde von Experiment 1 durch synchrone Perzeption moduliert wurden. Daher wurden die resultierenden Aktivierungen mit den entsprechenden Arealen aus Experiment 1

inklusive maskiert. Da die Stimuli von Experiment 1 und die der vorliegenden Untersuchung identisch waren, kann diese Maskierung durchgeführt werden.

Umgekehrt würden der Argumentation zufolge Areale, die zu Beginn asynchroner Exposition höhere Antworten aufweisen, als während der letzten (dritten) Minute (,1. > 3. Minute') anfangs asynchrone Reize verarbeiten und infolge anhaltender Exposition und damit einhergehender subjektiver Synchronizität an Einfluss auf die Zeitwahrnehmung verlieren und daher als Gebiete zu interpretieren sein, die zeitliche Unterschiede zweier Informationsströme präferentiell verarbeiten. Diese Änderung des neuronalen Antwortmusters würde nicht nach Exposition mit physikalisch synchronen Stimuli auftreten, da entsprechende Gebiete nicht Asynchronizität verarbeiten würden. Dementsprechend wurde der obige Kontrast nach asynchroner Exposition mit dem nach synchroner Exposition exklusive maskiert. Die identifizierten Areale würden Gebiete darstellen, die vorwiegend asynchrone Informationen auswerten und ihren Einfluss nach anhaltender Exposition zunehmend verlieren. Daher wurden sie innerhalb von Arealen analysiert, die zuvor (Experiment 1) durch asynchrone Perzeption moduliert wurden (inklusive Maskierung).

Areale, die gemäß den Ergebnissen möglicherweise mit einer zeitlichen Anpassung an versetzte auditorische und visuelle Informationsströme einhergehen, bilden ein Netz aus bilateral anterior und posterior gelegenen lateral temporalen Regionen (darunter der pSTS), dem rechten medialen Temporalkortex und dem DLPFC. Darüber hinaus wurden rechtsseitig unimodal visuelle und auditorische Gebiete durch diesen Kontrast moduliert. Dies bedeutet, dass bereits unimodale auditorische und visuelle Gebiete in Interaktion mit dem multimodalen STS und PFC adaptiv auf audiovisuelle Zeitinformationen reagieren. Das Netz kortikaler Modulationen, das durch subjektive zunehmende Asynchronizität moduliert wurde, besteht gemäß den Ergebnissen aus dem linken Parietal- und Präfrontalkortex.

Damit wurden distinkte kortikale Netze identifiziert, die an stimulusgetriebener Beeinflussung der Wahrnehmung subjektiver Gleichzeitigkeit bzw. Ungleichzeitigkeit beteiligt sind. Da audiovisuelle Gleichzeitigkeitswahrnehmung offenbar von lang- und kurzfristiger Erfahrung abhängt (Fujisaki et al., 2004; s. auch 3.2.4.3), stellen die identifizierten Gebiete vermutlich Regionen dar, die diese Erfahrungen (z. B.

zunehmender zeitlicher Abstand zwischen zwei Modalitäten) durch kurzzeitig plastische Prozesse in nachfolgende Wahrnehmungsmechanismen integrieren können.

Bushara et al. (2001) veränderten einfache audiovisuelle Reize hinsichtlich ihrer Gleichzeitigkeit und analysierten zugrundeliegende neuronale Korrelate. Sie identifizierten den PFC, den rechten Parietalkortex, die rechte Insula und das linke Cerebellum. Calvert et al. (2001) untersuchten zeitkritische Prozesse einfacher audiovisueller Reize und identifizierten Regionen, die auf bimodale Reize im Vergleich zu unimodalen mit veränderten BOLD-Mustern reagierten (Superadditivität und Antwortunterdrückung). Dies waren der PFC, der rechte Parietallappen, der linke STS und die Insula. Bushara et al. (2003) untersuchten anhand bewegter Balken audiovisuelle Koinzidenzwahrnehmung. Durch An- bzw. Abwesenheit eines auditorischen Stimulus wurde die Wahrnehmung hinsichtlich einer Kollision, bzw. eines aneinander Vorbeigehens beider Balken, manipuliert. Kollisionen modulierten den PFC, die Insula, den linken Parietallappen, das Kleinhirn und die SC. Gleichzeitig gingen Kollisionsperzepte im Vergleich zum Alternativperzept mit verminderter Aktivität in unisensorischen Gebieten einher.

Diese Gebiete scheinen oben aufgeführten den Studien anderer Autoren zufolge in die Auswertung physikalischer bzw. wahrnehmungsabhängiger Prozesse audiovisueller Zeitverarbeitung bei einfachen Reizen involviert zu sein. Diese Befunde decken sich größtenteils mit den Ergebnissen aus Experiment 1 (s. 3.1.3.2.2). Die vorliegende Studie erweitert diese Erkenntnisse insofern, als dass Areale identifiziert werden konnten, die neben automatischer Wahrnehmung zeitlicher Verhältnisse audiovisueller Informationen darüber hinaus durch stimulusgetriebene Mechanismen audiovisuelle Zeitwahrnehmung beeinflussen können. Dabei scheinen, in Übereinstimmung mit differentiellen neuronalen Netzen für Gleichzeitigkeits- und Ungleichzeitigkeitswahrnehmung (Dhamala et al., 2007), ‚synchronisierende‘ von ‚desynchronisierenden‘ Prozessen räumlich und funktionell voneinander differenzierbar zu sein.

Der Kontrast, der auf subjektiv zunehmende Synchronizität schließen lässt, modulierte anteriore Teile des STS und posteriore Teile des MTG und STS. Posteriore Teile des Temporallappens in der Umgebung des STS sind vermutlich an Integrationsprozessen multimodaler Informationen beteiligt (Calvert & Lewis, 2004; Hein & Knight, 2007; Macaluso et al., 2001; Meyer et al., 2007; Miller & D’Esposito, 2005;

Noesselt et al., 2007; Szycik et al., 2008a; 2008b). Seine multimodale Beteiligung wurde an integrativen Prozessen audiovisueller Sprache (Calvert et al., 1997; Meyer et al., 2007; van Atteveldt et al., 2004) und einfacher Stimuli (Beauchamp et al., 2004b; Calvert et al., 2001; Stevenson et al., 2009) festgestellt. Dabei beeinflussten räumliche (z. B. Macaluso et al., 2000), inhaltliche (z. B. Szycik et al., 2008a) und zeitliche Kongruenz (Miller & D'Esposito, 2005) diese Aktivierung. Tierstudien belegen, dass der STS über über direkte Verbindungen zu unisensorischen Kortices verfügt (Cappe & Barone, 2005; Desimone & Gross, 1979; Falchier et al., 2002; Schroeder et al., 1998; Rockland & Van Hoesen, 1994) und bildgebende Experimente am Menschen legen nahe, dass der STS über Rückkopplungsmechanismen mit unisensorischen Gebieten kommuniziert (Bushara et al., 2003; Driver & Noesselt, 2008; Giard & Peronnet, 1999; Noesselt et al., 2007; Noppeney et al., 2008; Sekiyama et al., 2003). Die hier modulierten anterioren Teile des STS sind wahrscheinlich Teil multimodal modulierter Gebiete des auditorischen Kortex. Tierstudien konnten zeigen, dass Neurone in unisensorisch auditorischen Gebieten auf multimodale (Ghazanfar et al., 2005; Kayser et al., 2007) bzw. Reize anderer Modalitäten (Brosch et al., 2005) reagierten. Weitere fMRT-Studien am Menschen zeigten, dass klassische unisensorische Regionen multisensorisch moduliert wurden (Amedi et al., 2002; Calvert et al., 1999; Macaluso et al., 2000; Martuzzi et al., 2007; Pekkola et al., 2005; Kriegstein & Giraud, 2004; Watkins et al., 2006). Dies legt nahe, dass multisensorische Inhalte im auditorischen Kortex (hier anteriorer STS) verarbeitet werden können. Des Weiteren wird vermutet, dass der anteriore STS Teil eines ventralen Pfades auditorischer Verarbeitung ist, der zur Kategorisierung akustischer Muster beiträgt (Buchsbaum & D'Esposito, 2009; Hickok & Poeppel, 2004), aber auch zur Stimmerkennung und zur Analyse wiederholter semantischer Inhalte (Belin & Zatorre, 2003).

Der oben aufgeführte Kontrast modulierte auch den rechten PFC. Aktivität im PFC ging in vergangenen Studien mit multimodaler Stimulation einher (Fuster, 1985; 1987; Sugihara et al., 2006), mit Aufmerksamkeitsprozessen in bezug auf mehrere Modalitäten (Loose et al., 2003), mit Fehlerkontrolle (Kerns et al., 2004), mit der Verarbeitung aufgabenabhängiger Ziele (Miller, 2000; Weissmann, 2004), mit dem Aufrechterhalten sensorischer Information (Baddeley, 1992; Yoo et al., 2004), aber auch mit der Verarbeitung zeitlicher Prozesse audiovisueller Sprache (Miller & D'Esposito, 2005) und einfacher Stimuli (Dhamala et al., 2007). Es wird vermutet, dass der PFC mit

unisensorischen, parietalen Arealen und dem STS interagiert (Bishop & Miller, 2008; Cusick et al., 1995), um die Relevanz unisensorischer Informationen zu modulieren.

Des Weiteren wurde der linke mediale Temporalkortex im Bereich des perirhinalen Kortex und des Gyrus parahippocampalis moduliert. Diese Gebiete scheinen in die Diskriminierung von Stimulusmerkmalen (Devlin & Price, 2007; Xu et al., 2007) und in multimodale Integrationsprozesse (Belardinelli et al., 2004; Holdstock et al., 2009; Tanabe et al., 2005) involviert zu sein und verfügen über Verbindungen zu Insula, Parietal- und Frontalkortex, zum visuellen Kortex und zum STS (z. B. Suzuki & Amaral, 1994).

Der Kontrast, der auf abnehmende subjektive Synchronizität schließen lässt, moduliert den linken PFC und Parietalkortex. Parietale Gebiete sind unter anderem in die Verarbeitung multimodaler Informationen einbezogen (Bushara et al., 2001; Dhamala et al., 2007; Macaluso et al., 2001; Miller & D'Esposito, 2005; Ojanen et al., 2005), in das Wechseln von Aufmerksamkeit zwischen den Modalitäten (Lewis et al., 2000), in Intervalldiskriminationsaufgaben (Onoe et al., 2001) und in Kategorisierungsprozesse (Corbetta et al., 2002) multimodaler Objektmerkmale (Grefkes et al., 2002).

Die hier modulierten Gebiete stellen laut bisherigen Befunden Regionen dar, die multimodale Informationsströme integrieren und außerdem miteinander kommunizieren können. Aus bestehenden Studien geht hervor, dass diese Gebiete an physikalischen und perzeptuellen Mechanismen multimodaler Integration beteiligt sind. Die vorliegende Untersuchung legt nahe, dass einige der bekannten multimodalen Regionen durch adaptive Prozesse die Wahrnehmung zeitlicher Relationen zwischen zwei Modalitäten maßgeblich beeinflussen können. So scheinen anteriorer und posteriorer STS zusammen mit dem rechten PFC und medialen Temporalkortex in Kombination mit unisensorischen auditorischen und visuellen Aktivierungen mit einer Regulierung der subjektiven Gleichzeitigkeit einherzugehen, so dass zeitliche Abweichungen zwischen Licht und Ton subjektiv ausgeglichen werden können. Dabei sind diese Areale vermutlich an einer Verstärkung synchroner Einflüsse beteiligt. Zusätzlich scheinen Gebiete in diesen Prozess involviert zu sein, die spezifisch den Einfluss asynchroner Informationen modulieren bzw. minimieren. Dies waren gemäß den Befunden der linke Parietal- und Präfrontalkortex.

Tootell et al. (1995) zeigten mittels fMRT anhand der Wasserfall-Illusion, dass Neurone in MT nach Adaptation auf eine Bewegungsrichtung höhere Antworten gaben als zu Beginn der Darbietung und erklärten den Befund mit richtungsspezifischen Zellen in MT, deren Antwortrate nach andauernder Stimulation sank und Zellen höhere Antworten zeigten, die spezifisch auf die der Stimulation entgegengesetzten Richtung reagierten und deren Einfluss zunahm. Analog dazu könnten im vorliegenden Experiment in den Gebieten, deren Antwort im Laufe der Exposition mit asynchronen Stimuli zunahm, Zellen existieren, die auf bestimmte Zeitverhältnisse spezialisiert sind. Dementsprechend würden Zellen, die asynchrone Informationen einer Verschiebungsrichtung verarbeiten (z. B. Ton vor Licht) nach Exposition mit Stimuli, bei denen der Ton dem Licht voranging, an Einfluss verlieren und solche, die die entgegengesetzte Verschiebungsrichtung verarbeiten ihre Antworten verstärken und so zu einer Verschiebung des PSS beitragen. Diesbezüglich würden Parietalkortex und linker PFC Areale darstellen, in denen Zellen, die asynchrone Zeitinformationen verarbeiten, mit andauernder asynchroner Exposition an Einfluss verlören und so gegenteilige Zeitperzepte zuließen. Im medialen Temporalkortex, STS und unisensorischen Kortices würden demzufolge Synchronizitätsperzepte verstärkt, während im Parietalkortex Asynchronizitätsperzepte inhibiert würden. Der PFC scheint beide Prozesse im Sinne einer Moderatorvariable lateralisiert zu modulieren.

Im Vergleich zu Experiment 1 (3.1.3.2.2), das rein perzeptuelle Mechanismen anhand derselben Stimuli wie die vorliegende Studie untersuchte, wurden im Bereich des PFC weniger Aktivierungscluster moduliert. Insula, Cerebellum und ACC wurden durch adaptive Prozesse nicht angesprochen. Aktivität in diesen Gebieten reflektiert womöglich entscheidungs- bzw. aufgabenabhängige Modulationen, während in den hier angesprochenen Gebieten audiovisuelle Zeitinformationen direkt, aufgabenunabhängig verarbeitet werden. Es ist bemerkenswert, dass bereits unimodale Areale adaptiv in die Repräsentation zeitlicher Relationen zweier Modalitäten eingreifen. Diese Modulationen könnten Rückkopplungsprozesse reflektieren, die gegebenenfalls durch den STS gesteuert wurden (Bushara et al., 2003; Noesselt et al., 2007; Driver & Noesselt, 2008). Dies bedeutet, dass eine Interaktion zwischen unisensorischen Kortices, multimodalem STS, linkem perirhinalen Kortex bzw. parahippocampalen Gebieten und rechtem PFC, in Abhängigkeit kurzfristiger multimodaler Zeitlichkeitserfahrungen, synchrone Perzepte

fördert, während gleichzeitig aufrechterhaltende Einflüsse asynchroner Perzeption in Parietalkortex und linken PFC minimiert werden. Diese Gebiete sind demzufolge nicht nur, wie aus bestehenden Studien bekannt, an automatischen Prozessen audiovisueller Stimulation beteiligt, sondern können regulierend in die Wahrnehmung zeitlicher Relationen beider Informationsströme eingreifen. Hervorzuheben ist dabei die Beteiligung, in der Verarbeitungshierarchie bimodalen Arealen vorangehender, unisensorischer Areale, die offenbar bereits ein Angleichen bimodal asynchroner Informationen beeinflussen können.

Anhand der vorliegenden Studie konnte gezeigt werden, dass Gebiete, die durch perzeptuelle (A)Synchronizität moduliert wurden, regulierend auf Zeitwahrnehmungsprozesse einwirken können, so dass die Vermutung naheliegend ist, dass relative Zeitinformationen zweier audiovisueller Informationsströme neuronal repräsentiert werden. BOLD-Antworten in Gebieten in der Umgebung des multimodalen pSTS, im anterioren STS, in auditorischen und visuellen Gebieten, im medialen Temporallappen und im PFC nahmen im Sinne eines Angleichens asynchroner Informationen zu und BOLD-Antworten in Parietal- und Präfrontalkortex nahmen im Sinne eines Loslösens von disparater Zeitinformation ab. Dies belegt zum ersten Mal, dass hierarchisch höhere bimodale Areale, aber auch bereits unisensorische Kortices, die automatische Prozesse audiovisueller Integration verarbeiten, auch stimulusgetriebene Prozesse, im Sinne einer flexiblen Rekalibrierung zeitlicher Relationen zweier Informationsströme, in die Verarbeitung nachfolgender perzeptueller Mechanismen einfließen lassen und analysieren. In psychophysischen Studien wurde belegt, dass audiovisuelle Informationen flexibel verarbeitet werden können (Spence & Squire, 2003; Sugihara & Suzuki, 2004; Vatakis et al., 2008), um Integrationsprozesse zu vereinfachen; im Rahmen der vorliegenden Untersuchung konnte darüber hinaus gezeigt werden, dass die Grundlage dieser Veränderungen neuronale Antwortmuster in kortikalen Gebieten sein können, die auf diese Weise plastisch auf multimodal eintreffende Informationen reagieren.

## 3.3. Experiment 3

### 3.3.1 Einführung

Bisher konnte in den Experimenten 1 und 2 gezeigt werden, welche kortikalen Areale in die Verarbeitung audiovisueller Zeitwahrnehmung involviert sind. Dies wurde anhand einfacher, kurzer, phasisch verschobener Stimuli demonstriert. Während in Experiment 1 Areale identifiziert wurden, die vorwiegend durch perzeptuelle Mechanismen moduliert werden, und in Experiment 2 solche, die darüber hinaus an stimulusgetriebenen, adaptiven Prozessen beteiligt sind, sollen nun im Rahmen von Experiment 3 Gebiete identifiziert werden, die durch Prozesse audiovisueller Zeitwahrnehmung bei sprachlichen, semantisch belegten Stimuli moduliert werden.

Die hier vorliegende Studie will in einem sogenannten „mixed-model“ perzeptuelle und entscheidungsabhängige Komponenten der Verarbeitung (a)synchroner komplexer Sprachstimuli trennen. Dabei soll hier die Verarbeitung zusammenhängender Sätze, also ökologisch valider Stimuli, und nicht einzelner Silben (z. B. Miller & D’Esposito, 2005; Mottonen et al., 2002; Ojanen et al., 2005; Tuomainen et al., 2005; van Atteveldt et al., 2004; 2007; van Wassenhove et al., 2005) oder, wie zuvor berichtet, einfacher audiovisueller Reize untersucht werden. Wesentlich dabei ist, dass in diesem Experiment Stimuli benutzt wurden, die nahe der Wahrnehmungsschwelle audiovisueller Asynchronizität lagen, so dass während der Präsentation identischer Reize unterschiedliche Wahrnehmungen auftreten und sich verändern konnten. Diese Veränderungen konnten Probanden während der Stimuluspräsentation mitteilen, so dass phasische Entscheidungen von stabilen Wahrnehmungen unterschieden werden konnten. Probandenspezifische Wahrnehmungsschwellen der Asynchronizität wurden in einem Vorexperiment individuell bestimmt, bevor die Probanden mittels fMRT gemessen wurden. Damit untersucht die vorliegende Studie zeitliche Aspekte audiovisueller Stimuli anhand komplexer sprachlicher Reize und kann aufgrund der Länge der verwendeten Stimuli (ca. 24 Sekunden, s. unten) stimulus-, wahrnehmungs- und entscheidungsabhängige Verarbeitungsmechanismen voneinander differenzieren. Besonderes Augenmerk soll dabei auf multisensorischen Gebieten (insbesondere dem STS, z. B. Calvert, 1997; Driver & Noesselt, 2008; Noesselt et al., 2008) liegen, die hinsichtlich ihrer funktionellen Neuroanatomie untersucht werden sollen. Die Länge und

Komplexität der hier verwendeten Stimuli erlaubt diese Analysen, die in den beiden oben aufgeführten Experimenten aufgrund der Kürze der dort benutzten Stimuli nicht gegeben waren. Daher kann diese Untersuchung die oben ermittelten Befunde erweitern. Der Vergleich der Ergebnisse der fMRT-Experimente 1,2 und 3 kann außerdem zeigen, wo Gemeinsamkeiten und Unterschiede der neuronalen Verarbeitung von zeitlich verzögerten einfachen und komplexen audiovisuellen Reizen liegen, bzw. ob die neuronale Basis audiovisueller Integration durch Komplexität und Länge der zu verarbeitenden Reize verändert wird.

Zusätzlich soll mittels einer Konnektivitätsanalyse ermittelt werden, mit welchen Arealen multisensorische Gebiete kommunizieren, wenn audiovisuelle Gleichzeitigkeit verarbeitet wird.

### 3.3.2 Vorexperiment 3

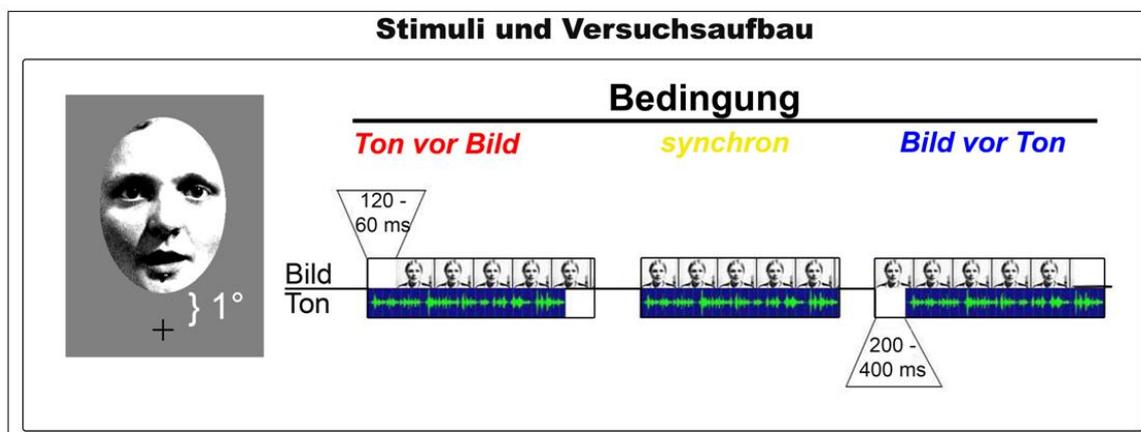
Um die Anzahl der Experimentalbedingungen für die fMRT-Studie zu reduzieren, war das Ziel dieses Vorexperiments 3 die Bestimmung der Verzögerungsschwellen für jeden fMRT-Probanden, bei denen er die gleiche Anzahl von synchronen und korrekten asynchronen Antworten gab. Nur diese Experimentalbedingungen (,Ton zuerst': gleiche Anzahl ,Ton-zuerst' und ,synchron'-Entscheidungen; ,Bild zuerst': gleiche Anzahl ,Bild zuerst' und ,synchron' -Entscheidungen) - individuell bestimmt für jeden Probanden - sowie die Synchronbedingung wurden im fMRT-Experiment verwendet, um eine größtmögliche Anzahl von Entscheidungsänderungen zu gewährleisten und so die Robustheit der Schätzung der Einzelparameter zu erhöhen. Eine identische audiovisuelle Verzögerung für alle Probanden wäre dafür nicht geeignet, da sich in einem vorangegangenen Verhaltensexperiment zeigte, dass die Probanden große Unterschiede in der Synchronizitätswahrnehmung aufwiesen.

#### 3.3.2.1 Methodik

Den Versuchspersonen wurden in zehn Durchgängen je 14 Videosequenzen in fünf verschiedenen physikalischen (A)Synchronizitätsbedingungen randomisiert präsentiert: zwei, in denen die Ton- der Bildspur unterschiedlich lang vorausging (130 ms, 60 ms), zwei, in denen die Tonspur mit unterschiedlicher Verzögerung nach der Bildspur

einsetzte (200 ms, 400 ms) und eine Synchronbedingung, in der Ton- und Bildspur gleichzeitig einsetzten (s. Abb. 34). Jede Bedingung wurde zweimal pro Sitzung präsentiert. Da Asynchronizität leichter entdeckt wird, wenn der Ton dem Bild vorausgeht als umgekehrt (Bushara et al., 2001; Dixon & Spitz, 1980, s. auch Vorexperimente 1 bis 2.2), wurden für diese Asynchronizitätsbedingungen kürzere Verzögerungen gewählt.

Die Probanden wurden mit einer Kinnstütze in 56 cm Abstand zum Bildschirm fixiert und hörten die Sprecherin über einen Kopfhörer (Panasonic, RP-HT379). Der Abstand zwischen zwei Videostimuli (ISI) betrug 10 Sekunden (s. Abb. 35).

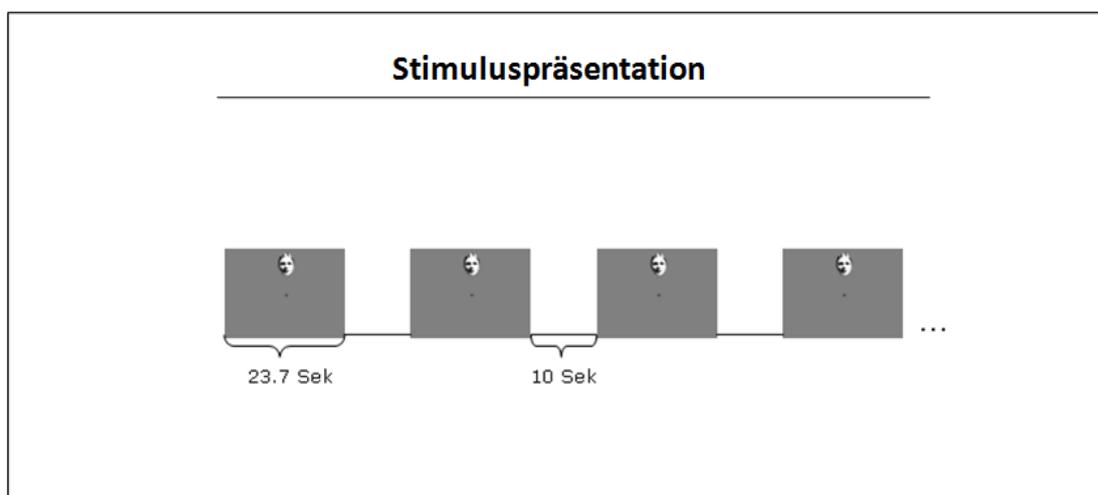


**Abbildung 34.** Videostimuli, die den Probanden präsentiert wurden. Das Umfeld der Sprecherin wurde durch eine Maske verdeckt und ihr Gesicht  $1^\circ$  -  $7^\circ$  Sehwinkel oberhalb des Fixationskreuzes präsentiert. Videosequenzen in Experiment 3 wurden in drei Bedingungen präsentiert: Ton vor Bild (120 ms oder 60 ms); Bild vor Ton (200 ms oder 400 ms) und synchron. Die Verzögerung zwischen Bild und Ton variierte infolge eines Vorexperiments, in dem individuelle Schwellen ermittelt wurden zwischen Synchronizität und Asynchronizität. Probanden gaben an, ob sie eine Videosequenz als ‚Ton voran‘, ‚Bild voran‘, oder ‚synchron‘ wahrnahmen. Sie wurden aufgefordert, Änderungen ihrer Wahrnehmung während einer Videosequenz zu berichten.

Die Probanden konnten über drei Tasten einer PC-Tastatur entscheiden, ob sie eine Stelle in der Videosequenz als synchron oder asynchron wahrnahmen. Im letzteren der beiden Fälle sollte entschieden werden, ob die Tonspur als vor oder nach der Bildspur verlaufend wahrgenommen wurde. Um Wahrnehmungswechsel bei den Probanden zu ermöglichen, wurde ihnen mitgeteilt, dass sich die Art Asynchronizität während einer Videosequenz ändern könne, und dass innerhalb einer Sequenz ein Wechsel von asynchron zu synchron oder vice versa möglich sei. Die Probanden wurden instruiert, so schnell wie möglich zu antworten und darauf hingewiesen, dass Entscheidungswechsel in derselben Videosequenz möglich seien. Dem eigentlichen Experiment gingen zwei

Videosequenzen voraus, in denen sich die Asynchronizitätsverhältnisse tatsächlich änderten (von 130 ms, die die Ton- der Bildspur bis 200 ms, die die Bild- der Tonspur vorausgehen konnte). Die Probanden sollten so für die Möglichkeit eines Bedingungswechsels innerhalb einer Videosequenz sensibilisiert werden. Jede der fünf Bedingungen wurde zweimal pro Sitzung präsentiert, die sich somit aus zehn Videosequenzen zusammensetzten.

In den Videosequenzen war eine Sprecherin zu sehen, die eine abgeänderte Form eines Märchens vorlas. Dabei war der Text so ausgewählt, dass keine /k/-, /t/- und /s/- Laute vorkamen, um zu verhindern, dass anhand dieser Konsonanten das Ausmaß der Synchronizität zu leicht entdeckt werden konnte.



**Abbildung 35.** Die Videostimuli wurden mit einer Dauer von 23,7 Sekunden präsentiert. Das Interstimulusintervall betrug 10 Sekunden. Pro Durchgang wurden 10 Videos in 5 Bedingungen (zufällige Reihenfolge) gezeigt.

Von der Sprecherin wurden 20 verschiedene Videosequenzen mit einer Sony DCR-PC3-Kamera mit 768 x 576 Pixel Auflösung bei 30 Hz im AVI-Format aufgenommen. Die Videos wurden mit Adobe Premiere 6.0 bearbeitet. Sie wurden zu ca. 24-sekündigen Sequenzen geschnitten. Zunächst wurden die Farbaufnahmen in Grauwerte umgewandelt und ihr Kontrast erhöht. Die durch die Verschiebung der Tonspur entstandenen Bildlücken wurden durch die ersten bzw. letzten Frames einer Sequenz aufgefüllt. Anschließend wurden sie anhand des MainConcept Mpeg Encoders 1.03.01.00 (MainConcept AG, Aachen, Deutschland) in MPEG-Videosequenzen mit 356 x 288 Pixel Auflösung umgewandelt. Die Tonspur wurde im MPEG2-Format gehalten.

Die Videosequenzen wurden mit Presentation 9.11 (Neurobehavioral Systems, Albany, USA) bei 800 x 600 Pixeln (60 Hz) unter Windows 98 präsentiert wobei der Hintergrund im Video mit einer Maske abgedeckt wurde, sodass nur das Gesicht der Sprecherin zu erkennen war (s. Abb. 34). Das Gesicht war ca. 1 cm oberhalb des Fixationskreuzes, das sich in der Mitte des Bildschirms befand zu sehen (1° Sehwinkel). Der Bildschirmhintergrund war grau (R: 204, B: 204, G: 204). Die Stimuli wurden mit MATLAB 6.1 (The MathWorks Inc., Natick, MA, USA) randomisiert.

Im Anschluss an das Experiment wurden für jede Verschiebungsrichtung die individuellen schwelennahen Verzögerungen aus den Verhaltensdaten bestimmt, sodass im fMRT-Experiment drei Verschiebungsbedingungen präsentiert wurden: ‚Ton zuerst‘, ‚Bild zuerst‘ und ‚gleichzeitig‘.

Eine Bedingung wurde immer dann als Schwelle identifiziert, wenn die Häufigkeiten der Antwortmöglichkeit ‚synchron‘ nahe oder gleich der Häufigkeit der Antwortmöglichkeit ‚Ton zuerst‘ bzw. ‚Bild zuerst‘ war. Für die Bedingung ‚Ton zuerst‘ konnte anhand der präsentierten Verschiebungsbedingungen für einige Probanden eine schwelennahe Tonspurverschiebung nicht eindeutig bestimmt werden, da sich die entsprechenden Antworthäufigkeiten zwischen 200 und 400 ms einander anzunähern schienen. Für diese Probanden wurde für die folgende fMRT-Studie eine neue aus den Daten approximierete Verschiebungsbedingung erstellt.

### 3.3.2.2 Ergebnisse des Vorexperiments

In Tabelle 3 sind die ermittelten Asynchronizitätsschwellen für alle Probanden dargestellt, die am nachfolgenden fMRT-Experiment teilgenommen haben. Aufgrund von Augenbewegungen sind drei Probanden aus der Datenanalyse eliminiert worden und daher nicht in Tabelle 6 aufgeführt.

**Tabelle 6.** Schwellenwerte für jeden Probanden, bei der eine Videosequenz genauso häufig als asynchron wie als synchron wahrgenommen wird. Die Tabelle bildet die Werte sowohl für Videosequenzen ab, bei denen das Bild dem Ton folgt als auch vice versa. Diese Werte wurden für das folgende fMRT-Experiment 3 verwendet.

Probanden	Verschiebung	
	<i>Ton zuerst</i>	<i>Bild zuerst</i>
1	130 ms	200 ms
2	130 ms	300 ms
3	60 ms	300 ms
4	60 ms	200 ms
5	60 ms	300 ms
6	130 ms	300 ms
7	130 ms	300 ms
8	130 ms	200 ms
9	130 ms	200 ms
10	60 ms	200 ms
11	130 ms	200 ms
<i>Mittelwert (Standardabweichung)</i>	104,6 ms (35,3)	227,3 ms (46,7)

### 3.3.2.3 Bewertung der Ergebnisse des Vorexperiments 3

Die Schwellen, um Asynchronizität zu entdecken sind für Stimuli, bei denen das Bild dem Ton vorangeht, höher als bei solchen, bei denen der Ton dem Bild vorangeht. Dies ist ein Befund, der auch schon in den vorangegangenen oben aufgeführten Experimenten mit einfachen Stimuli nachgewiesen werden konnte. Jedoch sind die Werte größer als in den vorherigen Experimenten. Bei einem durchschnittlichen PSS von ca. 32 ms und einer mittleren JND von ca. 130 ms lagen die Schwellen für audiovisuelle Sequenzen, bei denen der Ton dem Licht voranging bei ca. 33 ms und bei Stimuli, bei denen das Licht dem Ton voranging bei ca. 97 ms. Die Grenzwerte für sprachliche Stimuli sind daher ca. zwei bis drei Mal größer. Diese Befunde sind in etwa in Einklang mit denen von Dixon und Spitz (1980), die (A)synchronizitätsschwellen von 258 ms und 131 ms für semantisch belegte Stimuli bestimmten. Dagegen fanden sie ebenfalls geringere Schwellenwerte für einfache Stimuli. Diese Befunde mögen darauf zurückzuführen sein, dass multisensorische Integration von Sprache, im Vergleich zur Verarbeitung sprachlich nicht belegten Reizen, zusätzliche Prozesse involviert (z. B. Tuomainen et al., 2005).

Entsprechend den hier ermittelten Ergebnissen wurden Videosequenzen mit Verzögerungen zwischen Bild und Ton erstellt, die den individuellen Schwellen eines jeden Probanden entsprachen und während des fMRT-Experiments dargeboten wurden.

### 3.3.3 Hauptexperiment

Ziel der Studie, war es, die hirnanatomischen Korrelate der drei sensorischen Bedingungen (,Ton zuerst', ,synchron', ,Bild zuerst') in Verbindung mit den drei Wahrnehmungsmöglichkeiten der Probanden (ebenfalls ,Ton zuerst', ,synchron', ,Bild zuerst') und den dazugehörigen Entscheidungen zu identifizieren (insgesamt 21 Bedingungen). Die Analyse erfolgte innerhalb potentiell bimodaler Areale, um Erkenntnisse über die funktionelle Neuroanatomie multisensorischer Areale zu gewinnen.

Des Weiteren sollte per ROI-Analyse die funktionelle Neuroanatomie zeitlicher Verarbeitungsmuster innerhalb des STS als multisensorisches Integrationsareal untersucht werden und ausgehend von diesen Mustern interregionale Konnektivität untersucht werden.

#### 3.3.3.1 Methodik

An diesem fMRT-Experiment und dem dazugehörigen Verhaltensexperiment nahmen 14 Probanden (7 weiblich, Durchschnittsalter: 26,4 Jahre, Spanne: 19-31 Jahre) ohne psychiatrische oder neurologische Störungen teil, nachdem sie über den Zweck der Untersuchung aufgeklärt worden waren und schriftlich ihr Einverständnis zur Teilnahme gegeben hatten. Drei Probanden wurden nicht in die Datenanalyse aufgenommen, da anhand von Kameraaufzeichnungen der Augenbewegungen festgestellt werden konnte, dass die betroffenen Probanden den Instruktionen nicht folgten (Blinzeln, Augenbewegungen, Müdigkeit).

Die funktionellen und anatomischen Messungen wurden mit einem 3 Tesla-Kernspintomographen (Siemens Medical Systems, Erlangen, Deutschland) aufgezeichnet. Für die Aufnahme der anatomischen Struktur wurden ein T<sub>1</sub>-gewichtetes Volumen (räumliche Auflösung: 1 x 1 x 1 mm, 157 x 189 x 136 Schichten) mit einer Ganzkopf-Brukerspule (Bruker Bio-BioSpinMRI, Ettlingen, Deutschland) aufgenommen (TR: 2500

ms, TE: 4,77 ms, flip angle: 7°). Für die funktionellen Messungen wurde eine T<sub>2</sub>\*-gewichtete Echoplanar-Sequenz (echo planar imaging, EPI) verwendet (räumliche Auflösung: 3,5 x 3,5 x 4 mm, 64 x 64 x 32 Schichten, TR: 2000 ms, TE: 30 ms, flip angle: 80°). Zusätzlich wurde ein T<sub>1</sub>-gewichtetes inversion recovery-EPI aufgenommen, das dieselben Parameter und damit räumlichen Verzerrungen, wie die funktionellen EPI-Messungen aufwies (räumliche Auflösung: 3,5 x 3,5 x 4 mm, 64 x 64 x 32 Schichten, TR: 2500 ms, TE: 17 ms, TI: 1200 ms, flip angle: 80°). Dieses IR-EPI wurde zur Bestimmung der Normalisierungsparameter verwendet.

Die Stimuli wurden über einen Projektor (JVC, DLA-G150 CLE, 60 Hz) und eine Mattscheibe, die am Spulenrand befestigt war, auf einen Spiegel, der oberhalb der Kopfspule montiert war projiziert. Über einen elektrodynamischen Kopfhörer (Institut für Neurobiologie, Magdeburg, Deutschland), wurden die zu den jeweiligen Videosequenzen gehörenden Tonspuren präsentiert. Vor den funktionellen Messungen wurde ein Lautstärketest mit Scannermessgeräuschen absolviert, in dem die Probanden bestimmten, ob der Ton beidseitig gleichermaßen gut und laut genug wahrzunehmen war. Gegebenenfalls wurden die Lautstärkeeinstellungen korrigiert.

Vor den experimentellen Durchgängen fand ein Lokalisationsdurchgang (329 Scans) statt, in dem in sieben Blöcken ausschließlich Tonspuren der aufgenommenen Videosequenzen und in 8 Blöcken ausschließlich Bildspuren in der Länge (23,7 Sekunden, ISI: 20 Sekunden) der experimentellen Videosequenzen präsentiert wurden. Dieser Lokalisationsdurchgang diente dazu, die visuellen und auditorischen Areale der Probanden zu identifizieren, die an der Verarbeitung der unisensorischen Stimuluskomponenten beteiligt waren. Zudem wurden die Areale, die sowohl durch den visuellen als auch den auditorischen Stimulus aktiviert waren, als multimodale Areale definiert (Beauchamp, 2005; Noesselt et al., 2007; Szycik et al., 2008b). Die Probanden wurden gebeten, während des Lokalisationsdurchgangs die Bildschirmmitte zu fixieren und sonst keine Aufgabe zu bewältigen.

Für die fMRT-Messungen wurde das Stimulationsmaterial in vier Sitzungen mit je 15 Videosequenzen aufgeteilt, so dass jede Experimentalbedingung fünf Mal je Durchgang vorkommen konnte. Eine Videosequenz war 23,7 Sekunden lang. Das ISI betrug 20 Sekunden, so dass zwischen den Videos ausreichend Zeit für das Abklingen der BOLD-Antwort war.

Die Versuchspersonen erhielten dieselben Instruktionen wie im vorangegangenen Vorexperiment 3. Die Antworten wurden über einen Taster, der an ihrer rechten Hand befestigt war, mit ihren Zeige-, Mittel-, und Ringfingern aufgezeichnet.

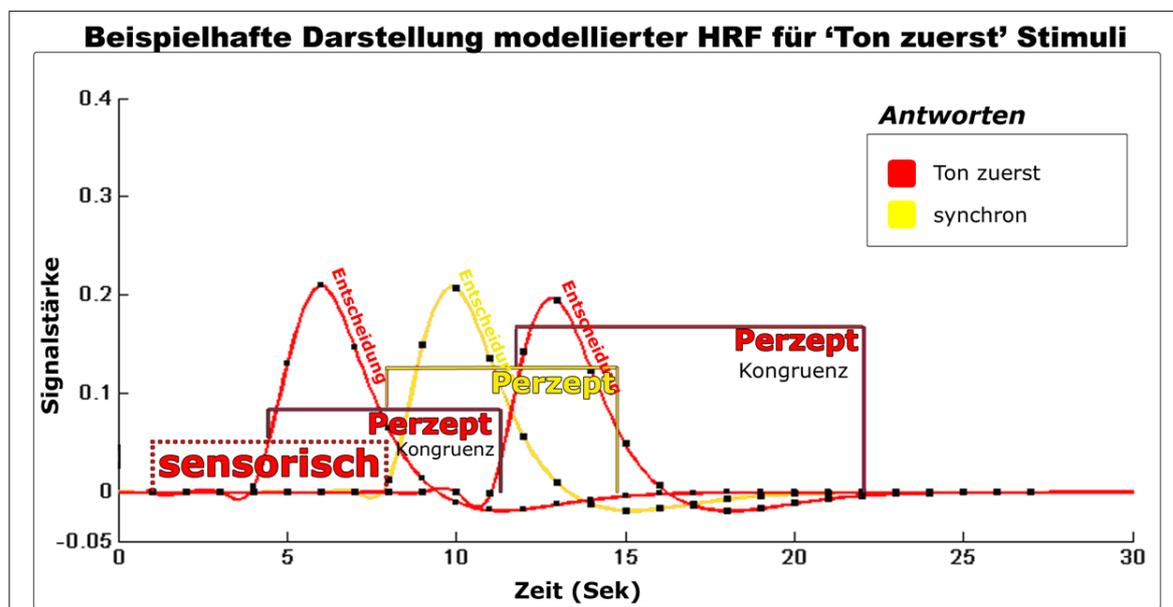
Um zu jedem Antwortzeitpunkt die visuelle Repräsentation der Videosequenzen bei den Probanden sicherzustellen, wurden während aller funktionellen Aufnahmen die Augenbewegungen des rechten Auges mit einem endoskopischen Infrarotlichtaufnahmegerät aufgezeichnet (Kanowski et al., 2007). Dieses optische Aufnahmesystem besteht aus einer Kamera, die mittels eines Endoskops ungefähr zwei Zentimeter vom Auge entfernt platziert wird, und durch ein Glasfaserkabel mit einem Aufzeichnungsgerät (DVD-Recorder, Pioneer, Pioneer Electronics GmbH, Willich, Deutschland) außerhalb des Scanners verbunden ist. Die Auflösung des Systems entspricht einer Detektierbarkeit von Augenbewegungen, die größer als  $0,5^\circ$  sind. Die Augenbewegungen wurden an einem PC mit Hilfe der PupilTracker-Software (HumanScan, Erlangen, Deutschland) verfolgt.

Für den Lokalisationsdurchgang wurde ein Modell mit zwei Bedingungen (,auditorisch' und ,visuell') erstellt (Blockdesign, Blocklänge: 23,7 Sekunden, ISI: 20 Sekunden.). Die Daten wurden voxelweise auf ihren globalen Mittelwert normiert sowie hoch- und tiefpassgefiltert. Anschließend wurden die Kontraste für beide Experimentalbedingungen relativ zu einer impliziten Baseline berechnet und diese Daten anhand einer Random-effects-Analyse auf Gruppenebene ausgewertet.

Die modellierten Bedingungen des Hauptexperiments im Rahmen des allgemeinen linearen Modells (SPM2, Friston et al., 1995) waren die drei sensorischen Bedingungen (,Ton zuerst', ,synchron', ,Bild zuerst') in Verbindung mit den drei Wahrnehmungsmöglichkeiten der Probanden (ebenfalls ,Ton zuerst', ,synchron', ,Bild zuerst') und den drei entsprechenden Entscheidungen (insgesamt 21 Bedingungen). Jede Videosequenz wurde daher in einen sensorischen Teil (vor der ersten Probandenantwort), drei perzeptuelle, sowie drei Entscheidungsteile unterteilt (s. Abb. 36). Die perzeptuellen und entscheidungsabhängigen Teile wurden nach jeder Entscheidungsänderung des Probanden, in die durch die Probandenantwort indizierte Kategorie aufgeteilt. Dabei wurde der sensorische Teil als variabler Block von Stimulusbeginn bis zu ersten Antwort definiert. Die perzeptuellen Blöcke definierten sich von einem Antwortzeitpunkt bis zum nächsten, während die Entscheidungen als ereignisabhängig zu jedem Knopfdruck

---

definiert wurden. Dieses Vorgehen ermöglicht, stabile Perzepte von entscheidungs- und stimulusabhängigen Modulationen zu trennen. Da die sensorischen Bedingungen (,Ton zuerst', ,synchron', ,Bild zuerst') feste Anfangs- und Endzeitpunkte haben, die Zeitpunkte und Dauer der Antworten bzw. der Entscheidungen jedoch nicht vorhersehbar waren, wurde ein sogenanntes ,mixed-model' (Kombination aus ereignisabhängigem – und Blockdesign) berechnet. Aufgrund dieser zeitlichen Unvorhersehbarkeit der Antworten und der daraus folgenden Abhängigkeit von Einzelbedingungen, kann Art und Verlauf der BOLD-Antworten der Einzelbedingungen jedoch nicht in allen Einzelprobanden unabhängig geschätzt werden. Daher können nur die differentiellen Kontraste zwischen Bedingungen berechnet werden.



**Abbildung 36.** Die Abbildung zeigt exemplarisch einen Versuchsdurchgang, für den hämodynamische Antwortfunktionen (HRF) für einen ,Ton voran'-Stimulus modelliert wurden. Die Kurven stellen perzeptuelle (Um-)Entscheidungen dar, die Kästchenfunktionen bilden die sensorische Verarbeitung vor der ersten Antwort und perzeptuelle Zustände ab.

Aufgrund des geringen Auftretens einiger Bedingungen wurden alle vier Sitzungen des fMRT-Experiments zusammengeführt und mit einem fixed-model analysiert. Daher wurden zusätzlich vier Regressoren modelliert, um irrelevante Varianz, die durch die vier getrennt durchgeführten Sitzungen entstanden sein kann, von der Varianz der Experimentalbedingungen zu trennen.

Für die Analyse wurden die Daten voxelweise auf ihren globalen Mittelwert normiert sowie hoch- und tiefpassgefiltert. Nach der Modellschätzung wurden die Kontraste ‚asynchron > synchron‘ und vice versa für die sensorischen (abhängig von den präsentierten Videosequenzen) und perzeptuellen (Wahrnehmung der Probanden) Bedingungen anhand einer Random-effects-Analyse berechnet. Dabei sind entscheidungsrelevante Korrelate bereits extrahiert (s. oben und Abb. 36). Diese Kontraste wurden über die Versuchsbedingungen kollabiert. Darüber hinaus wurden zusätzlich Kontraste für jeweils jede Versuchsbedingung erstellt, die Perzepte, die mit der jeweiligen Versuchsbedingung kohärent sind (‚korrekte Antwort‘), mit Perzepten verglichen, die nicht der Stimulationsbedingung entsprechen (‚falsche Antwort‘). Diese Kontraste wurden auch für eine Einzelprobanden-Analyse verwendet. Für jeden Probanden wurde für asynchrone Stimuli, bei denen der Ton dem Bild voranging ‚Ton-voran‘-Perzepte gegen synchrone Perzepte verglichen, für synchrone Stimuli ‚synchrone‘ gegen beide asynchrone Perzepte und für asynchrone Stimuli, bei denen das Bild voranging ‚Bild-voran‘-Perzepte gegen synchrone Perzepte. Die Ergebnisse wurden gemäß einer ‚Regions-of-Interest‘ (ROI)-Analyse ausgewertet, die funktionellen (innerhalb multisensorischer Masken) und anatomischen Kriterien (STS) folgte. Die Einzelanalysen bezogen sich auf die STS-Region, da dieses Gebiet als multisensorisches Integrationsareal gilt (z. B. Beauchamp, 2005; Calvert et al., 1997; Doehrmann & Naumer, 2008; Driver & Noesselt, 2008; Macaluso et al., 2004; Miller & D’Esposito, 2005; Noesselt et al., 2007; van Atteveldt et al., 2007). Innerhalb dieses Areals sollten für jeden Probanden zeitliche Muster der audiovisuellen Integration untersucht werden. Daher wurden zusätzlich Beta-Gewichte der durch die hier beschriebenen Kontraste modulierten Regionen extrahiert, um Anhaltspunkte dafür zu erhalten, wie sich das Antwortmuster der identifizierten Areale über zeitlich verschiedene audiovisuelle Stimulationen und ihrer entsprechenden Perzepte verändert.

Außerdem wurden räumliche Abstände zwischen den lokalen Maxima der oben erläuterten Kontraste perzeptueller Kohärenz innerhalb der STS-Region berechnet. Dazu wurden MNI-Koordinaten einer Dimension (x oder y oder z) verschiedener lokaler Maxima voneinander abgezogen (x-Koordinaten wurden zuvor in ihre Beträge umgewandelt). Diese Differenzen (Katheten) wurden gemäß dem Satz des Pythagoras quadriert und aufaddiert (Hypothense) und dieses Ergebnis zum Quadrat der Z-

Differenz addiert. Die Wurzel aus dieser Addition (Hypothenuse) entspricht den unten berichteten Distanzen in Millimetern.

Aufgrund der schon erwähnten teilweisen Abhängigkeit der Versuchsbedingungen untereinander, können die Parameter nicht immer unabhängig voneinander geschätzt werden (Friston et al., 1999). Daher werden die in den fixed-effects-Modellen der Einzelprobanden geschätzten differentiellen Kontraste gruppenanalytisch mittels einseitiger T-Tests analysiert.

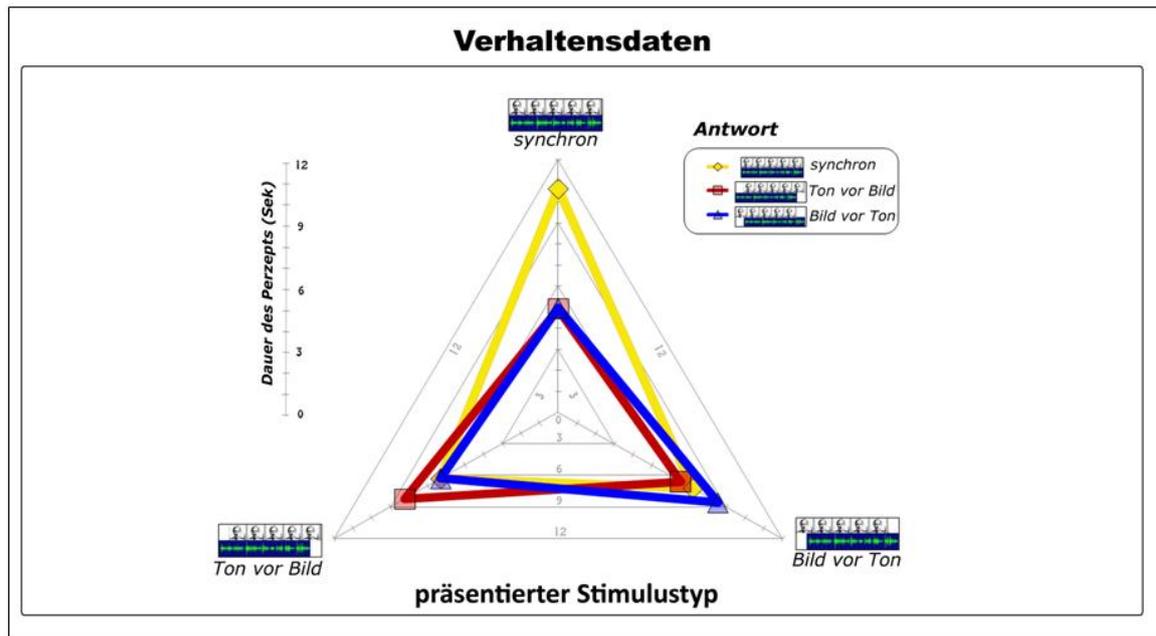
Die Kontraste wurden mit den Ergebnissen des Lokalisationsdurchgangs inklusive maskiert ( $p < 0,005$ ,  $k > 10$ ), so dass die Effekte des Hauptexperiments innerhalb bimodaler Areale analysiert werden können (z. B. Beauchamp, 2005; Beauchamp et al., 2004a; Noesselt et al., 2007; Szycik et al., 2008a; 2008b).

Die Ergebnisse der Einzelprobandenanalyse der STS-Region wurden mit ihren jeweiligen lokalen Maxima einer Konnektivitätsanalyse unterzogen [psychophysiologische Interaktion (PPI): Friston et al., 1997]. Dadurch kann ermittelt werden, welche kortikalen Gebiete mit den perzeptabhängigen Modulationen innerhalb des STS funktionell interagieren. Diese Befunde wurden durch eine Varianzanalyse auf zweiter Ebene (Gruppenebene) mit SPM 2 analysiert. Anhand psychophysiologischer Interaktionen wird versucht, die physiologische Aktivität in einem Teil des Gehirns in bezug auf eine Interaktion zwischen dem Verlauf eines sensumotorischen oder kognitiven Prozesses und der Aktivität in einem anderen Teil des Gehirns zu erklären. So erlaubt dieses Verfahren beispielsweise, Informationen über die Aktivität eines Areals, das durch eine bestimmte Aufgabe moduliert wird (beispielsweise das Beurteilen audiovisueller Stimuli) und Informationen über den Stimulus zu kombinieren und so weitere Regionen zu lokalisieren, die durch den Stimulus moduliert werden unter der Bedingung, dass die Aktivität im ersten Areal (hier: STS) hoch ist. Auf diese Weise sollen im vorliegenden Experiment Regionen lokalisiert werden, die zusammen mit dem STS durch qualitativ unterschiedliche Zeitperzepte (‘Ton vor Bild’, ‘synchron’, ‘Bild vor Ton’) komoduliert werden.

### 3.3.3.2 Ergebnisse

#### 3.3.3.2.1 Verhaltensdaten

Da die Probanden während einer Videosequenz mehrere Perzepte haben konnten bzw. mehrere Antworten (Ton zuerst/synchron/Bild zuerst) geben konnten, wurde berechnet, wie oft Probanden zwischen den Antworten wechselten und wie lang ein bestimmtes Perzept aufrechterhalten wurde. Im Durchschnitt wurde pro Videosequenz 5,72 Mal (2,34) zu einem Perzept gewechselt, das mit der Präsentationsbedingung (Ton zuerst/synchron/Bild zuerst) kohärent war (Standardabweichung in Klammern). D.h. dass in diesem Fall das subjektive (A)synchronizitätsperzept der dargebotenen Verzögerung zwischen Bild- und Tonspur entsprach. Im Vergleich dazu wurde durchschnittlich 3,97 Mal (2,0) zu einem inkohärentem Perzept gewechselt, währenddessen die subjektive Wahrnehmung nicht der präsentierten Stimulusbedingung entsprach. Darüber hinaus verblieben Probanden durchschnittlich 9,13 Sekunden (3,38) bei einem kohärenten Perzept, wogegen sie 6,04 Sekunden (2,02) ein inkohärentes aufrechterhielten (s. Abb. 37). Dies zeigt, dass Probanden öfter ein kohärentes, stabiles Perzept der dargebotenen (A)synchronizität hatten und dieses länger aufrechterhielten als sie zu einem inkohärenten wechselten bzw. dies aufrechterhielten. Anhand der stabilen Perzepte konnten im Rahmen der Auswertung der fMRT-Daten hämodynamische Antwortfunktionen mit variablen Längen in einem gemischten Modellen (mixed-model) mit Ereignissen (events; hier: Tastendruck) und variablen Blöcken (hier: Dauer der Perzepte) zuverlässig modelliert werden.



**Abbildung 37.** Die Abbildung zeigt durchschnittliche Zeitspannen (Zeit von einer Antwort zur nächsten während einer Videosequenz), die Probanden ein stabiles Perzept je experimenteller Bedingung aufrechterhielten. Perzeptuelle Zustände bestanden am längsten, wenn ein Perzept der Stimulation entsprach. Dies erlaubte die spezifische Modellierung hämodynamischer Antwortfunktionenvon (HRF) verschiedener stabiler Perzepte für die Analyse der fMRT-Daten (s. Abb. 36).

### 3.3.3.2.2 fMRT-Daten

#### 3.3.3.2.2.1 Lokalisationsdurchgang

Während dieser Sitzung wurden visuelle, auditorische und bimodale (audiovisuelle) Areale identifiziert (s. Tab. 12). Die bimodalen Areale wurden ermittelt, indem der auditorische Kontrast mit dem visuellen inklusive maskiert wurde ( $p < 0,05$ ;  $k > 100$ ). Die Kontraste des Hauptexperiments werden mit diesen drei Kontrasten des Lokalisationsdurchgangs inklusive maskiert. Dabei wurden bimodale Aktivierungen aus visuellen und auditorischen eliminiert.

Die durch den Lokalisationsdurchgang identifizierten Areale sind in Tabelle 12 a-c (s. Anhang) dargestellt. In der auditorischen Lokalisationsbedingung wurden bilateral primäre und sekundäre auditorische Areale im Bereich des superioren Temporallappens gefunden (s. Anhang, Tab. 12 a). Dazu gehörten bilateral die HG, das PT, Planum polare sowie laterale Gebiete des STG und des STS. In der visuellen Lokalisationsbedingung wurden bilaterale okzipitale Areale moduliert. Dazu zählten bilaterale Gebiete im LG, FG, MOG, IOG und der Sulcus occipitalis transversalis (s. Anhang, Tab. 12 b). Audiovisuell

bimodale Aktivierungen zeigten sich bilateral im posterioren und medialen STS und in einigen präfrontalen Gebieten (s. Anhang, Tab. 12 c).

#### 3.3.3.2.2.2 Experimentelle Kontraste

Die folgenden neuronalen Korrelate werden ausschließlich innerhalb der zuvor definierten bimodalen Maske dargestellt, da das Augenmerk dieser Untersuchung multisensorischen Gebieten gilt.

##### 3.3.3.2.2.2.1 Physikalische (A)Synchronizität

Gebiete, die während der bloßen Präsentation synchroner Videosequenzen stärker neuronal moduliert wurden als während asynchroner Stimulation, waren der rechte bimodale STS und der laterale Präfrontalkortex (s. Anhang, Tab. 13 b und Abb. 38 a). Im Vergleich zu synchroner Stimulation modulierte die Präsentation von Stimuli, bei denen der Ton dem Bild voranging, beidseits den posterioren, rechtsseitig den anterioren STS, die linke anteriore Insula, den linken PFC und den rechten prämotorischen Kortex stärker (s. Anhang, Tab. 13 a und Abb. 38 a). Während des Zeigens von Videosequenzen, bei denen das Bild dem Ton voranging waren der linke posteriore und der rechte anteriore STS, die rechte anteriore Insula und der laterale PFC beidseits stärker aktiviert als während des Zeigens synchroner Stimuli (s. Abb. 38 a). Die Präsentation asynchroner Stimuli zeigt demzufolge vergleichbare Antwortmuster beider asynchroner Bedingungen (Ton voran und Bild voran) im STS, PFC und in der anterioren Insula, während die Darbietung synchroner Stimuli den pSTS und PFC ohne insuläre Beteiligung ausschließlich rechtslateral moduliert.

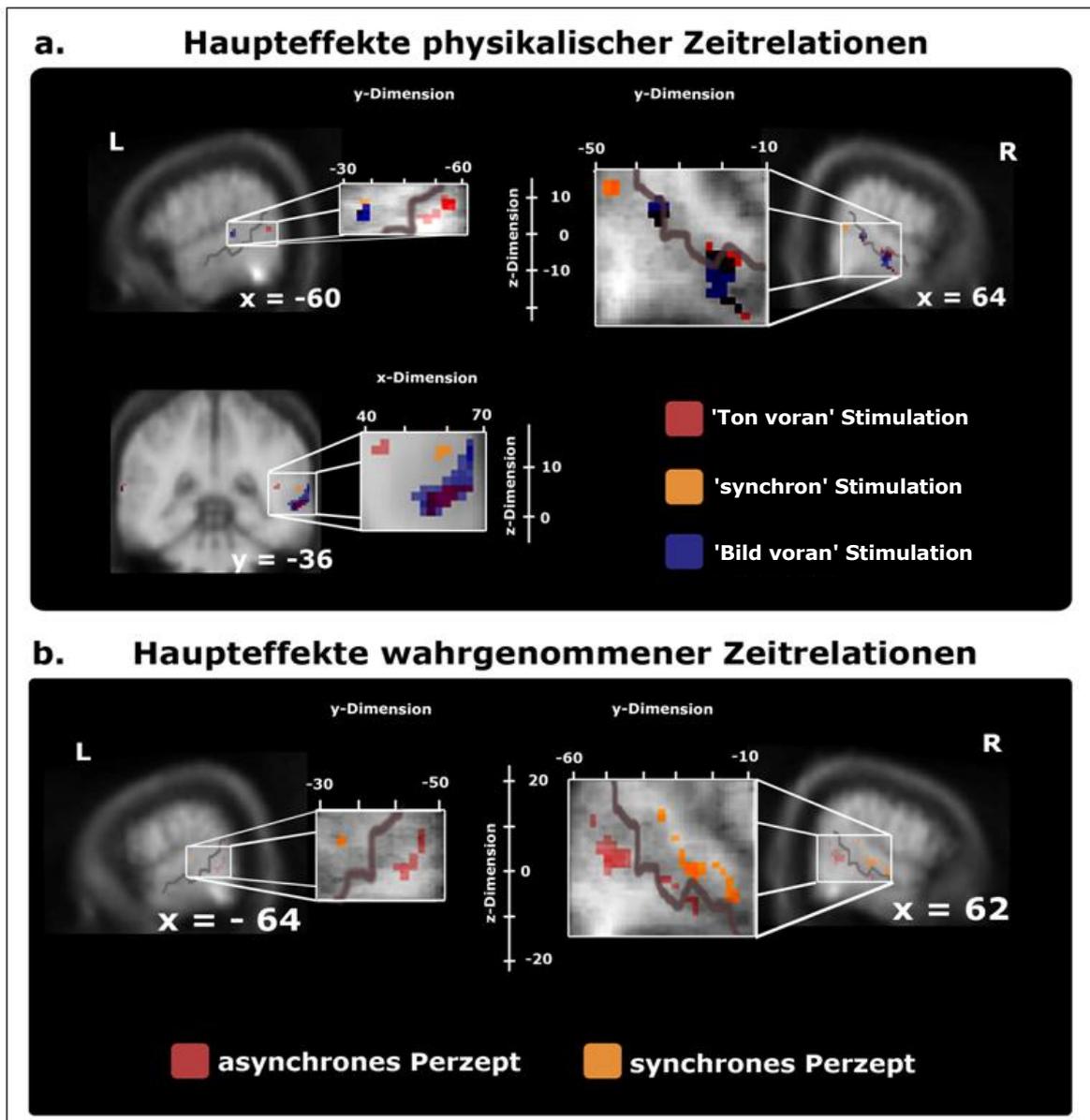
##### 3.3.3.2.2.2.2 (A)Synchrone Perzeption

Der Vergleich der Synchronizitätsperzepte unabhängig von der präsentierten Verzögerung zwischen Bild und Ton zeigte, dass der STS beidseits stärker durch synchrone Wahrnehmung angesprochen wurde als durch asynchrone (s. Anhang, Tab. 14 a und Abb. 38 b). Asynchrone Perzepte modulierten ebenfalls den STS beidseits, aber auch den bilateralen PFC und die rechte anteriore Insula stärker im Vergleich zu synchroner Wahrnehmung aller präsentierten Stimulustypen (s. Anhang, Tab. 14 b und Abb. 38 b). Die Aktivierungen synchroner und asynchroner Perzepte entlang des STS

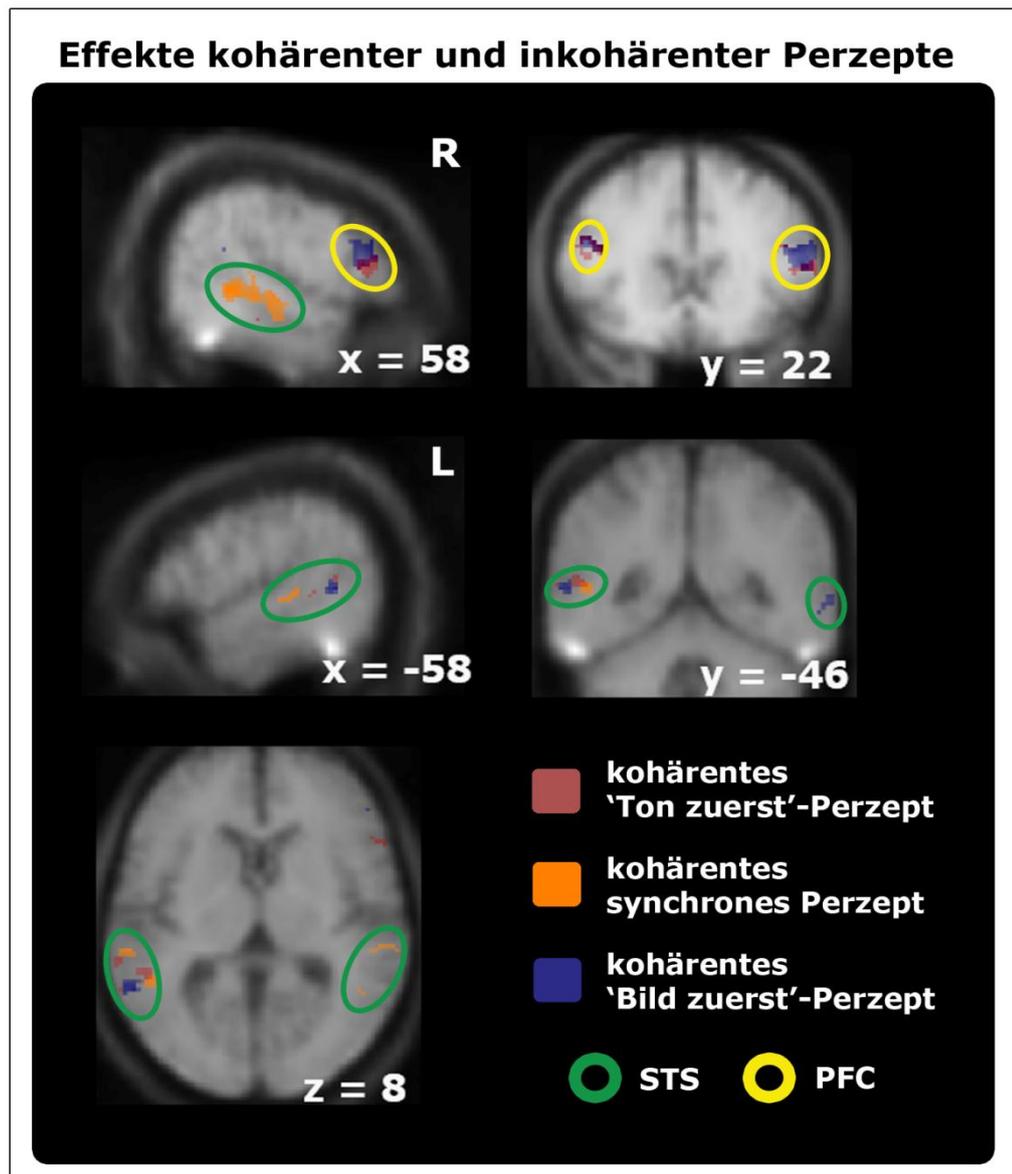
überlappten sich nicht, waren jedoch ‚benachbart‘ (s. Abb. 38 b). Sowohl asynchrone als auch synchrone Perzepte rekrutierten den Befunden zufolge bimodale Bereiche des STS, zusätzlich modulierten asynchrone Perzepte insuläre und präfrontale Regionen.

#### 3.3.3.2.2.2.3 *Kohärente Perzepte*

Des Weiteren wurden innerhalb eines Stimulustyps (Bild voran, synchron, Ton voran) stabile, anhaltende Perzepte analysiert, wenn das Perzept dem Stimulustyp entsprach (Kohärenz von Stimulustyp und Wahrnehmung). Kohärente asynchrone Perzepte, für Stimuli, bei denen der Ton dem Bild voranging, riefen bilateral höhere neuronale Antworten des STS, des Gyrus praecentralis, des PFC und der linken anterioren Insula hervor als (inkohärente) synchrone Perzepte während der Darbietung dieses Stimulustyps (Ton voran; s. Anhang, Tab. 15 a und Abb. 39). Kohärent synchrone Perzepte modulierten den STS beidseits und die linke anteriore Insula stärker als asynchrone Perzepte während der Darbietung synchroner Videosequenzen (s. Anhang, Tab. 15 b). Kohärente perzeptuelle Asynchronizität von Stimuli, bei denen das Bild dem Ton voranging wiesen bilateral höhere neuronale Aktivität in STS, Gyrus praecentralis, anteriorer Insula und PFC auf als synchrone Perzepte während der Präsentation von Stimuli, bei denen das Bild dem Ton voranging (s. Abb. 39 und Anhang, Tab. 15 c). Gemäß diesen Befunden ging kohärente Wahrnehmung synchroner Stimuli mit erhöhter neuronaler Aktivität in STS und anteriorer Insula einher, während kohärente Wahrnehmung asynchroner Stimuli zusätzlich zu STS und anteriorer Insula den PFC und prämotorische Areale modulierte. Innerhalb des linkslateralen pSTS grenzten die BOLD-Modulationen der beiden kohärenten asynchronen und synchronen Perzepte aneinander an, so dass die beiden durch kohärente asynchrone Perzepte modulierten Regionen die durch kohärente synchrone Perzepte modulierte Region umgaben (s. Abb. 39).



**Abbildung 38.** Gruppenanalyse unterschiedlicher Stimulustypen und stabiler perzeptueller Zustände (die Zeit von einer Antwort während einer Sequenz zur nächsten) innerhalb zuvor als bimodal lokalisierter Areale. **a.** Differentielle BOLD-Antworten der drei verwendeten Stimulustypen über alle Perzepte zusammengefasst. Die Analyse wies (neben prämotorischen Modulationen, nicht dargestellt, s. Tab. 14) signifikante Effekte im rechten STS (orange) für synchrone im Vergleich zu asynchronen Reizen auf, im STS beidseitig für ‚Bild voran‘-Stimuli im Vergleich zu synchronen (blau, auch PFC, nicht dargestellt) und im linken pSTS und rechten anterioren und posterioren STS für ‚Ton voran‘-Stimuli im Vergleich zu synchronen (rot, auch PFC und prämotorische Areale, nicht dargestellt). **b.** Der Vergleich synchroner gegen asynchrone Perzepte über alle Stimulustypen zusammengefasst (orange) zeigt Modulationen, die vom mSTS zum pSTS reichen. Nebeneinanderliegende Regionen innerhalb des STS wurden darüber hinaus für den Vergleich asynchroner gegen synchrone Perzepte identifiziert (rot; zusätzlich wurde der PFC durch diesen Kontrast moduliert, nicht dargestellt, s. Tab. 38)



**Abbildung 39.** Die dargestellten Kontraste vergleichen stabile Perzepte, die mit der Stimulation kohärent waren gegen inkohärente innerhalb zuvor als bimodal lokalisierter Areale. Während der STS ausschließlich durch kohärent synchrone Perzepte moduliert wurde (orange, oben links), wurden präfrontale Areale und der linke STS durch kohärente, asynchrone Perzepte [sowohl ‚Ton voran‘ (rot), als auch ‚Bild voran‘ (blau)] moduliert. Innerhalb des linken STS wiesen die Vergleiche räumlich getrennte, nebeneinanderliegende Aktivierungen auf. S. auch Tab. 15

#### 3.3.3.2.2.4 Region-of-Interest-Analyse

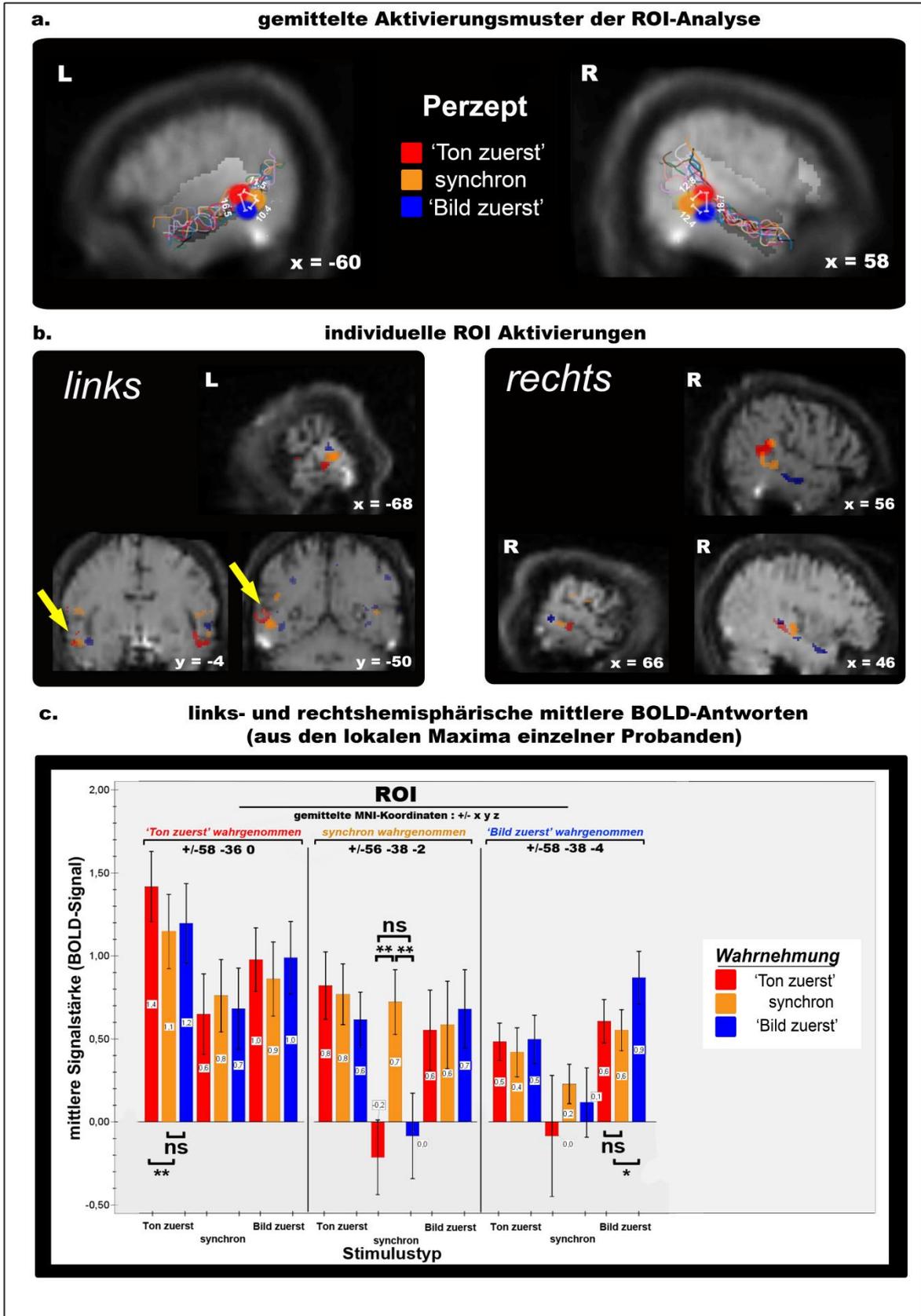
Um die funktionelle Anordnung dieses Gruppenbefundes von zeitlichkeitsbedingter Wahrnehmung innerhalb des multimodalen pSTS näher zu untersuchen, und aufgrund individueller anatomischer Unterschiede in dieser Region (Ochiai et al., 2004), wurde auf Einzelprobandenebene eine Region-of-Interest (ROI)-Analyse innerhalb des bimodalen STS durchgeführt. Dafür wurden, wie schon auf

Gruppenebene, für jeden Stimulustyp kohärente mit inkohärenten Perzepten verglichen und die Ergebnisse durch die bereits auf die vorherigen Befunde angewandte bimodale Maske gefiltert. Zusätzlich wurden für diese drei Kontraste die mittleren Beta-Gewichte ( $\beta$ , die sich proportional zur Signaländerung der BOLD-Antwort verhalten) der Perzepte aller Stimulusbedingungen aus den jeweiligen lokalen Maxima extrahiert (s. Abb. 40 c).

Mit allen extrahierten Beta-Gewichten aller untersuchten Kontraststufen wurde eine  $2 \times 3 \times 3 \times 3$ -faktorielle Varianzanalyse mit Messwiederholung für die Faktoren *Hemisphäre*, *funktionelle Region* (asynchrone Kohärenz bei Ton voran, synchrone Kohärenz und asynchrone Kohärenz bei Bild voran), *Perzept* und *Stimulustyp* durchgeführt. Da kein Effekt für *Hemisphäre* [ $F(1;10) = 0,25$ ;  $p = 0,63$ ;  $\eta^2 = 0,02$ ] gefunden wurde, wurden die Beta-Gewichte beider Hemisphären in Abbildung 40 c zusammengefasst. Haupteffekte ergaben sich für *funktionelle Region* [ $F(2;20) = 4,9$ ;  $p < 0,05$ ;  $\eta^2 = 0,33$ ] und *Perzept* [ $F(1,33;13,26) = 10,9$ ;  $p < 0,01$ ;  $\eta^2 = 0,52$ ], Interaktionseffekte für *funktionelle Region*, *Perzept* und *Stimulustyp* [ $F(8;80) = 3,1$ ;  $p < 0,01$ ;  $\eta^2 = 0,24$ ], für *Hemisphäre* und *funktionelle Region* [ $F(2;20) = 8,04$ ;  $p < 0,01$ ;  $\eta^2 = 0,45$ ], für *funktionelle Region* und *Perzept* [ $F(4;40) = 3,48$ ;  $p < 0,05$ ;  $\eta^2 = 0,26$ ], für *funktionelle Region* und *Stimulustyp* [ $F(1,73;17,3) = 9,17$ ;  $p < 0,01$ ;  $\eta^2 = 0,48$ ] und für *Perzept* und *Stimulustyp* [ $F(1,7; 17,06) = 4,7$ ;  $p < 0,05$ ;  $\eta^2 = 0,32$ ]. Obwohl nach Bonferroni korrigierte Post-Hoc-Tests keine signifikanten Unterschiede aufwiesen, sind BOLD-Antworten innerhalb der *funktionellen Region*, die durch kohärente Perzepte von Stimuli moduliert werden, bei denen das Bild dem Ton voran geht, im statistischen Trend geringer als BOLD-Modulationen in den beiden anderen funktionelle Regionen ( $p < 0,1$ ). Post-Hoc-Analysen des Faktors *Perzept* zeigten, dass BOLD-Antworten auf synchrone Stimuli signifikant niedriger waren als auf asynchrone ( $p < 0,01$ ). Weiterführende Post-hoc T-Tests ergaben, dass BOLD-Antworten auf einen Stimulustyp am höchsten waren, wenn Antworten bzw. Perzepte dem Stimulustyp (Ton voran/synchron/Bild voran) entsprachen, Perzepte also kohärent mit den präsentierten Stimulustypen waren (s. Abb. 40 c). Die mittlere BOLD-Antwort auf ein kohärentes Perzept von Stimuli, bei denen der Ton dem Bild voranging innerhalb der funktionellen Region ‚Ton voran‘ war signifikant höher als die beiden BOLD-Antworten der Alternativperzepte ( $p < 0,05$ ; rote Balken). Beta-Gewichte der inkohärenten Perzepte des Stimulustyps ‚Ton voran‘ unterschieden sich dagegen nicht signifikant voneinander. Dieses Muster trat auch für die Stimulustypen ‚synchron‘ ( $p < 0,001$ ; gelbe Balken) und ‚Bild voran‘ ( $p < 0,05$ ; blaue

Balken) auf. Die BOLD-Antworten inkohärenter Perzepte innerhalb der funktionellen Region ‚synchron‘ unterschieden sich nicht signifikant von den Signalstärken der Perzepte anderer Stimulustypen. Diese Ergebnislage unterscheidet die funktionelle Region, die durch kohärente synchrone Perzepte moduliert wird von den beiden anderen, in denen kohärente Perzepte generell die höchste Signalstärke erzeugten.

Die funktionelle Analyse einzelner Probanden bestätigte den Gruppenbefund, dass die räumliche Anordnung der drei funktionellen Regionen eine Systematik aufzuweisen schien (durch kohärent wahrgenommene Asynchronizität modulierte Regionen umgeben die durch kohärent wahrgenommene Synchronizität modulierten Gebiete). D.h. bei jedem Probanden grenzten die funktionellen Regionen kohärenter Asynchronizität an Regionen kohärenter Synchronizität, beide Regionen kohärenter Asynchronizität (für jeweils Ton voran und Bild voran) grenzten jedoch nicht aneinander (s. Abb. 40 a, b). Um diese räumliche Anordnung statistisch zu erfassen, wurden Distanzen zwischen den einzelnen funktionellen Regionen ermittelt (s. 3.3.3.1 und Abb. 40 a) und anhand einer 2 x 3-faktoriellen Varianzanalyse mit Messwiederholung für die Faktoren *Hemisphäre* und *Distanz* analysiert. Der mittlere Abstand zwischen Arealen, die durch kohärent wahrgenommene Stimuli moduliert wurden, bei denen der Ton bzw. das Bild voranging und Arealen, die durch kohärent wahrgenommene Synchronizität moduliert wurden betrug, 12,1 mm (von ‚Ton voran‘ zu ‚synchron‘) bzw. 11,1 mm (von ‚Bild voran‘ und ‚synchron‘), der Abstand zwischen beiden Regionen, die durch kohärent wahrgenommene Asynchronizität moduliert wurden dagegen 17,6 mm (von ‚Ton voran‘ zu ‚Bild voran‘; s. Abb. 40 a). Die Varianzanalyse ergab einen Haupteffekt für *Distanz* [ $F(2;20) = 10,2$ ;  $p < 0,001$ ;  $\eta^2 = 0,51$ ], aber keine Effekte für *Hemisphäre* und die Interaktion aus beiden Faktoren [*Hemisphäre*:  $F(1;10) = 0,81$ ;  $p = 0,39$ ;  $\eta^2 = 0,08$ ; *Interaktion*:  $F(2,20) = 0,03$ ;  $p = 0,97$ ;  $\eta^2 = 0,00$ ]. Post-Hoc T-Tests ergaben, dass der Abstand zwischen beiden Regionen, die durch kohärente Asynchronizität moduliert wurden, signifikant größer war als die Abstände dieser beiden Regionen zu Regionen, die durch kohärent wahrgenommene Synchronizität moduliert wurden (Ton voran und Bild voran:  $p < 0,05$ ). Die beiden Abstände von Regionen, die durch kohärent wahrgenommene Asynchronizität moduliert wurden, zu Regionen, die durch kohärent wahrgenommene Synchronizität moduliert wurden, unterschieden sich dagegen nicht signifikant voneinander ( $p = 0,55$ ).

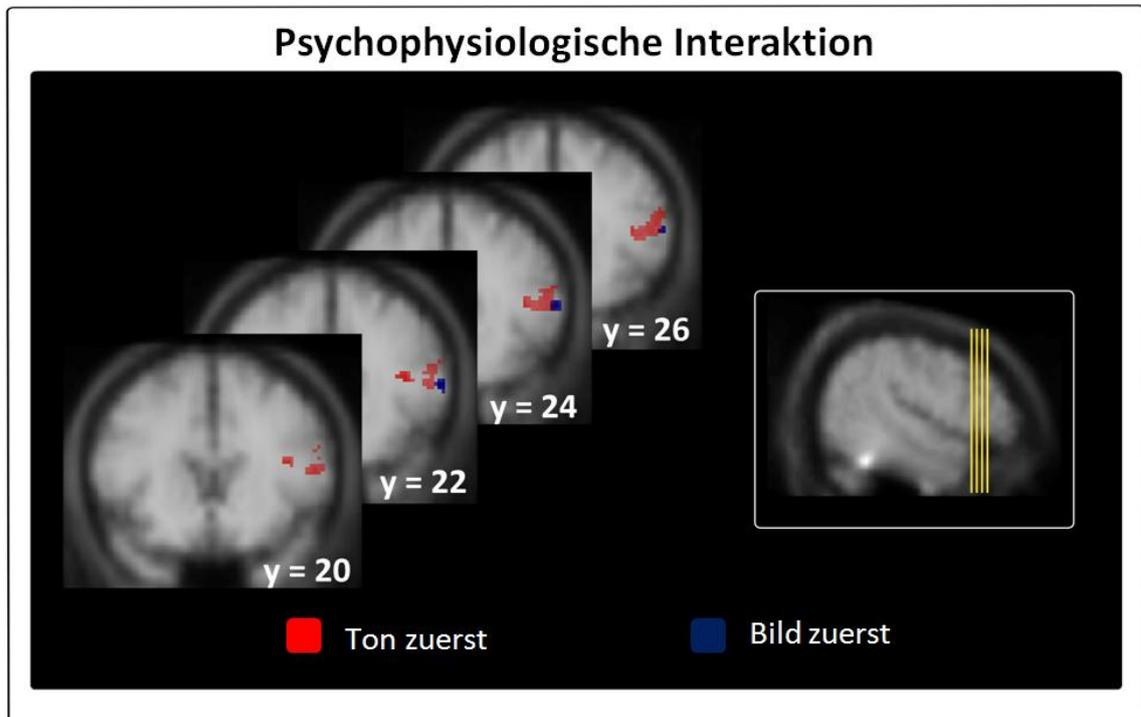


**Abbildung 40.** Die Abbildungen zeigen die ROI-Analyse auf Ebene einzelner Probanden. Die abgebildeten Kontraste stellen innerhalb bimodal lokalisierter Gebiete (grau schattierte Bereiche) subjektive Perzepte dar, die kohärent mit der jeweiligen Stimulation waren (rot: kohärentes ‚Ton voran‘-Perzept bei ‚Ton voran‘-Stimulation > restliche Perzepte bei ‚Ton voran‘-Stimulation; blau: analoger Kontrast in Bezug auf ‚Bild voran‘-Stimulation; orange: analoger Kontrast zu synchroner Stimulation). **a.** Die farblichen Bereiche repräsentieren mittlere lokale Maxima (11 Probanden) von Gebieten, die höhere Aktivität für kohärente Perzepte als für inkohärente aufwiesen. Die weißen Linien stellen, zusammen mit den dazugehörigen Zahlen, mittlere Abstände zwischen den lokalen Maxima dar. Abstände zwischen asynchronen Bereichen waren immer größer, als deren Abstände zu synchronen Bereichen. Die farbigen Linien bilden individuelle anatomische Verläufe des STS einzelner Probanden nach Normalisierung ab. **b.** Die Abbildungen zeigen Aktivierungskarten der oben aufgeführten Kontraste dreier Probanden. Synchrone Modulationen waren von asynchronen umgeben. Diese Aktivierungsmuster wurden sowohl links-, als auch rechtshemisphärisch lokalisiert. **c.** Gemittelte  $\beta$ -Gewichte (proportional zur prozentualen Signaländerung) der lokalen Maxima aus **a.** wurden über beide Hemisphären zusammengefasst. Die Balken bilden das Ausmaß der BOLD-Antwort (y-Achse) für jedes stabile Perzept (rot: ‚Ton vor Bild‘; orange: ‚synchron‘; blau: ‚Bild vor Ton‘) für die drei verwendeten Stimulustypen (‚Ton vor Bild‘; ‚synchron‘; ‚Bild vor Ton‘; x-Achse) innerhalb jedes lokalen Maximums aus **a.** ab. Die BOLD-Effekte asynchroner Perzepte innerhalb asynchroner lokaler Maxima (äußere Bereiche des Balkendiagramms) waren immer höher als die aller anderen Perzepte in Bezug auf die unterschiedlichen Stimulustypen. Innerhalb des lokalen Maximums für synchrone Perzepte waren die BOLD-Antworten auf synchrone Perzepte höher als auf asynchrone Perzepte, wenn synchrone Reize präsentiert wurden.

#### 3.3.3.2.2.5 Psychophysiologische Interaktionen

Da durch die oben aufgeführten Befunde einzelne funktionelle Regionen für kohärent wahrgenommene (A)Synchronizität innerhalb des bimodalen STS identifiziert und voneinander abgegrenzt werden konnten, wurde im Folgenden überprüft, ob diese Gebiete interregionale Verknüpfungen zu anderen multisensorischen Arealen aufweisen. Dazu wurden die lokalen Maxima der zuvor identifizierten funktionellen Regionen entlang des STS anhand einer psychophysiologischen-Interaktionsanalyse (PPI, Friston et al., 1997) untersucht und verglichen, ob sich funktionelle Verknüpfungen der einzelnen, oben identifizierten Areale, in ihrer Stärke voneinander unterscheiden. Die Analyse ergab, dass die funktionellen Regionen, die kohärente asynchrone Perzepte von Stimuli modulieren, bei denen der Ton oder das Bild vorangeht (s. Anhang, Tab. 16 a, c und Abb. 41) stärker mit rechtsseitigen präfrontalen Gebieten verknüpft waren als diejenigen, die oben durch kohärente synchrone Perception moduliert wurde (s. Anhang, Tab. 16 b). Dieses letzte Gebiet wiederum, im mSTS gelegen, zeigte stärkere psychophysiologische Interaktionen mit dem linken pSTS als durch asynchrone kohärente Perception modulierte Regionen. Durch perzeptuelle Asynchronizität modulierte funktionelle

Regionen wiesen im Temporallappen verstärkte interregionale Verknüpfungen mit rechtem aSTS und pSTS auf (s. Abb. 41).



**Abbildung 41.** Konnektivitätsanalyse stabiler, mit der Stimulation übereinstimmender, asynchroner Zeitperzepte. Stärkere Verknüpfungen temporaler Gebiete wurden ausschließlich im Kontext asynchroner, mit der Stimulation kohärenter Perzepte, im Vergleich zu kohärent synchronen Perzepten, identifiziert (links im Bild). Die Analyse identifizierte rechtslaterale präfrontale, getrennte Regionen für ‚Ton vor Bild‘-(rot), und ‚Bild-vor-Ton‘-Perzepte (blau). S. auch Tab. 16.

### 3.3.3.3 Diskussion

Die letzte aufgeführte Studie untersuchte anhand audiovisueller Sprachreize die neuronalen Grundlagen physikalischer und subjektiv perzeptueller Eigenschaften zeitlicher Verhältnisse innerhalb multisensorischer Areale des Menschen, die auch in den vorangegangenen fMRT-Studien durch einfache Stimuli moduliert wurden. Drei wesentliche kortikale multimodale Regionen wurden sowohl durch die Verarbeitung physikalischer Zeitlichkeitsunterschiede, als auch durch aufrechterhaltene, stabile Perzepte der Stimuli moduliert: der STS, der insuläre, und der Präfrontalkortex. Dabei konnte neuronale Aktivität der anterioren Insula und des PFC vorwiegend der Präsentation asynchroner Stimuli, unabhängig von der subjektiven Wahrnehmung derselben, und asynchronen Perzepten, unabhängig von physikalischen Eigenschaften der Stimuli, zugeordnet werden. Der multimodale STS wurde durch physikalische und

perzeptuelle Synchronizität und Asynchronizität moduliert, wobei aneinander grenzende funktionelle Regionen innerhalb des STS ausgemacht wurden, die in Abhängigkeit der dargebotenen Art der Verschiebung von Bild und Ton (Ton voran, synchron, Bild voran) an kohärente Perzepte (Übereinstimmung von Stimulustyp und seiner Wahrnehmung) geknüpft waren. Eine ROI-Analyse aller Probanden bestätigte dieses räumlich-funktionelle Antwortmuster und wies zusätzlich auf, dass stabile, kohärente Perzepte die neuronalen Antworten innerhalb dieser funktionellen Subregionen verstärkten. Des Weiteren ergab die psychophysiologische Analyse kortikaler Interkonnektivität, dass die funktionellen Regionen, die auf asynchrone Perception spezialisiert waren, stärker mit dem rechten lateralen PFC und Gebieten des rechten posterioren STS verknüpft sind, während funktionelle Regionen, die auf synchrone Perception spezialisiert waren, stärkere Verknüpfungen mit dem linken posterioren STS aufwiesen. Zusätzlich konnte gezeigt werden, dass über alle Probanden hinweg eine regelmäßige räumliche Anordnung der funktionellen Subregionen vorliegt, in der Regionen, die auf perzeptuelle Asynchronizität spezialisiert sind, solche umgeben, die auf perzeptuelle Synchronizität spezialisiert sind.

Bisherige Studien, die audiovisuelle Zeitverarbeitung anhand bildgebender Verfahren untersuchten, konnten zeigen, dass der STS (neben anderen Strukturen) an der Verarbeitung von Zeitlichkeitsprozessen beteiligt ist (Calvert, 2001; Dhamala et al., 2007; Driver & Noesselt, 2008; Macaluso et al., 2004; Miller & D'Esposito 2005; Noesselt et al., 2007). Jedoch wurde im Rahmen derjenigen Studien, die audiovisuelle Integration semantisch belegter multimodaler Reize untersuchten (Calvert & Campbell, 2003; Calvert et al., 2000; Macaluso et al., 2004; van Atteveldt et al., 2007) nicht zwischen aufgaben- und wahrnehmungsbezogenen Effekten differenziert, so dass die dort berichteten Ergebnisse nicht eindeutig auf stimulus-, entscheidungs-, oder wahrnehmungsbezogene Prozesse zurückgeführt werden können.

Wie bereits erwähnt, untersuchten Miller und D'Esposito (2005) das ‚Verschmelzen‘ kurzer audiovisueller Silben, das anhand fMRT untersucht wurde und berichtete Effekte für wahrgenommene Gleichzeitigkeit asynchroner Reize im linken, und für wahrgenommene Ungleichzeitigkeit synchroner Reize, im rechten STS. Unterschiede im Stimulusmaterial könnten die dazu abweichenden Befunde des vorliegenden Experiments erklären. Lateralisierte Effekte wurden im vorliegenden Experiment zwar

nicht für die primären Kontraste gefunden, nichtsdestotrotz stimmen die Befunde der Konnektivitätsanalyse mit den berichteten Befunden von Miller und D'Esposito überein.

Van Atteveldt et al. (2004; 2007) identifizierten lateral temporale Gebiete (PT, STG, STS) als wichtige Integrationsareale für die Bindung audiovisueller Graphem-Morphem Paare. Die Intensität der BOLD-Modulationen stieg innerhalb auditorischer Areale für semantisch kongruente Bedingungen und die Position der Modulationen innerhalb des STS veränderte sich mit zunehmender Verzögerung zwischen auditorischem und visuellem Reiz: asynchrone Stimuli wurden vorwiegend am Rand des Aktivierungsmusters des STS verarbeitet und geringer ausgeprägte Verzögerungen zwischen den beiden Modalitäten im Kern der Aktivierung. Dabei überlappten die Modulationen der verschiedenen Verzögerungsbedingungen jedoch. Für die Verarbeitung synchroner audiovisueller Buchstaben wurden keine Effekte im STS berichtet.

Weitere Studien untersuchten audiovisuelle Zeitverarbeitung anhand einfacher Stimuli. Calvert et al. (2001) untersuchten multimodale Interaktionsmuster, indem einfache audiovisuelle Reize entweder synchron oder zeitlich verschoben dargeboten wurden. Diese bimodalen Reize führten bei synchroner Präsentation zur Verstärkung und bei asynchroner Präsentation zur Unterdrückung der jeweiligen BOLD-Antwort in multisensorischen Gebieten, u.a. im STS, aber auch in frontalen Arealen. Noesselt et al. (2007) berichteten kontralaterale Effekte des STS, wenn synchrone, semantisch nicht belegte Stimuli lateralisiert dargeboten wurden. Sie fanden jedoch keine Effekte für asynchrone Stimuli.

Die vorliegende Studie konnte zeigen, dass wahrgenommene Asynchronizität (stabile Perzepte) mit erhöhter bilateraler neuronaler Aktivität im pSTS, der anterioren Insula und im PFC einhergeht. Dabei wurden im Gegensatz zu den berichteten, bereits durchgeführten Untersuchungen anderer Autoren zum ersten Mal lang anhaltende Stimuli verwendet, die pro Darbietung verschiedene Perzepte hervorrufen konnten und damit eine Annäherung an natürliche Situationen darstellen. Des Weiteren wurden die Ergebnisse dergestalt dargestellt, dass perzeptuelle von physikalischen und entscheidungsbezogenen Prozessen differenziert werden konnten und die berichteten Befunde eindeutig multisensorischen Arealen zugeordnet werden konnten. Die Ergebnisse decken sich mit Befunden anderer Untersuchungen, die audiovisuelle

Wahrnehmung anhand bildgebender Verfahren analysierten und assoziierte neuronale Modulationen im rechten STS (Miller & D'Esposito, 2005) und präfrontalen Gebieten (SFG, MFG, IFG) fanden (Bushara et al., 2001; Dhamala et al., 2007). Die aktuellen Ergebnisse der vorliegenden Untersuchung untermauern diese Befunde und lassen darauf schließen, dass der STS und präfrontale Areale an aufrechterhaltenen asynchronen Perzepten beteiligt sind.

Darüber hinaus deuten die hier vorgelegten Ergebnisse darauf hin, dass multimodale Regionen innerhalb des STS in einzelne Subregionen unterteilt werden können, die unterschiedliche zeitliche Muster verarbeiten. Durch Studien an Affen, die die anatomische Textur der temporal-parietal-okzipitalen Region (TPO; ein dem menschlichen STS wahrscheinlich homologes Gebiet; Beauchamp, 2005) untersuchten, konnte Belege für eine Dreiteilung, die sich von caudalen zu rostralen Teilen des STS erstreckt, gefunden werden (Cusick et al., 1995). Diese Unterteilungen unterschieden sich hinsichtlich ihrer Chemoarchitektur. Seltzer und Pandya (1991) konnten zeigen, dass die TPO-Region aus zytoarchitektonischen Unterteilungen besteht, von denen hauptsächlich der rostrale Teil über direkte Verbindungen zur Insula verfügt. Andere chemoarchitektonische Befunde legen nahe, dass die obere Wölbung der TPO-Region des Rhesus-Affen aus verschiedenen anatomischen und funktionellen Zonen besteht (Padberg et al., 2003). Dabei konnte gezeigt werden, dass der STS innerhalb dieser getrennten, neurochemischen Module in Bereiche aufgeteilt ist, die Verbindungen zu anderen kortikalen Gebieten aufweisen. Diese Bereiche innerhalb des STS könnten funktionelle Relevanz besitzen. Eine Untersuchung anhand fMRT von Beauchamp et al. (2004a) ergab, dass der STS in unisensorisch auditorische, visuelle und multisensorische Bereiche eingeteilt werden kann. Die vorliegenden Ergebnisse erweitern diese Befunde, indem sie multisensorische Unterteilungen des STS aufzeigen, die unterschiedliche audiovisuelle Muster enkodieren, wenn (A)synchronizität von audiovisueller Sprache beurteilt wird. In Anbetracht der Tatsache, dass die Unterbereiche, die durch perzeptuelle Synchronizität moduliert wurden zwischen solchen lagen, die perzeptuelle Asynchronizität (Ton voran oder Bild voran) verarbeiteten, stellen die funktionellen Regionen eine chronologische Anordnung dar (von einer Asynchronizität über Synchronizität hin zur anderen Asynchronizität) und lassen auf einen funktionellen ‚Zeitstrahl‘ entlang des STS schließen. Eine aktuelle Studie (Fairhall & Macaluso, 2009) berichtete Modulationen des mittleren

STS durch attendierte kongruente audiovisuelle Sprachstimuli, jedoch nicht im posterioren Bereich des STS und vermutet eine funktionelle Unterteilung des STS von anterior nach posterior. Die vorliegenden Ergebnisse, die spezifische Effekte im mittleren STS für unterschiedliche zeitliche Muster und deren Wahrnehmung aufweisen, und mit stärkeren Verbindungen zu posterioren Gebieten des STS einhergehen, sind in Einklang mit dieser Hypothese.

Die oben aufgeführten Befunde könnten Hinweise auf neuronale Pfade bzw. Mechanismen geben, wie multimodale Areale mit anderen kortikalen Regionen interagieren, um perzeptuelle Integration von auditorischen und visuellen Reizen zu ermöglichen. Ergebnisse aus Tierstudien zur audiovisuellen Bindung von Gesichtern und Stimmen bei Affen (Ghazanfar et al., 2008) zeigen beispielsweise, dass Neurone im STS verstärkt mit auditorischen Arealen interagieren, wenn diese Stimuli verarbeitet werden und dass der STS über Rückkopplungsmechanismen mit dem auditorischen Kortex kommuniziert, wenn audiovisuelle Reize verarbeitet werden (Schroeder et al., 2008). Zusammen mit den Befunden anderer Studien lässt die hier vorliegende Ergebnislage die Vermutung zu, dass innerhalb des STS voneinander abgegrenzte funktionelle Subregionen existieren, die darüber hinaus über Rückkopplungsmechanismen mit unisensorischen Kortices interagieren (Driver & Noesselt, 2008).

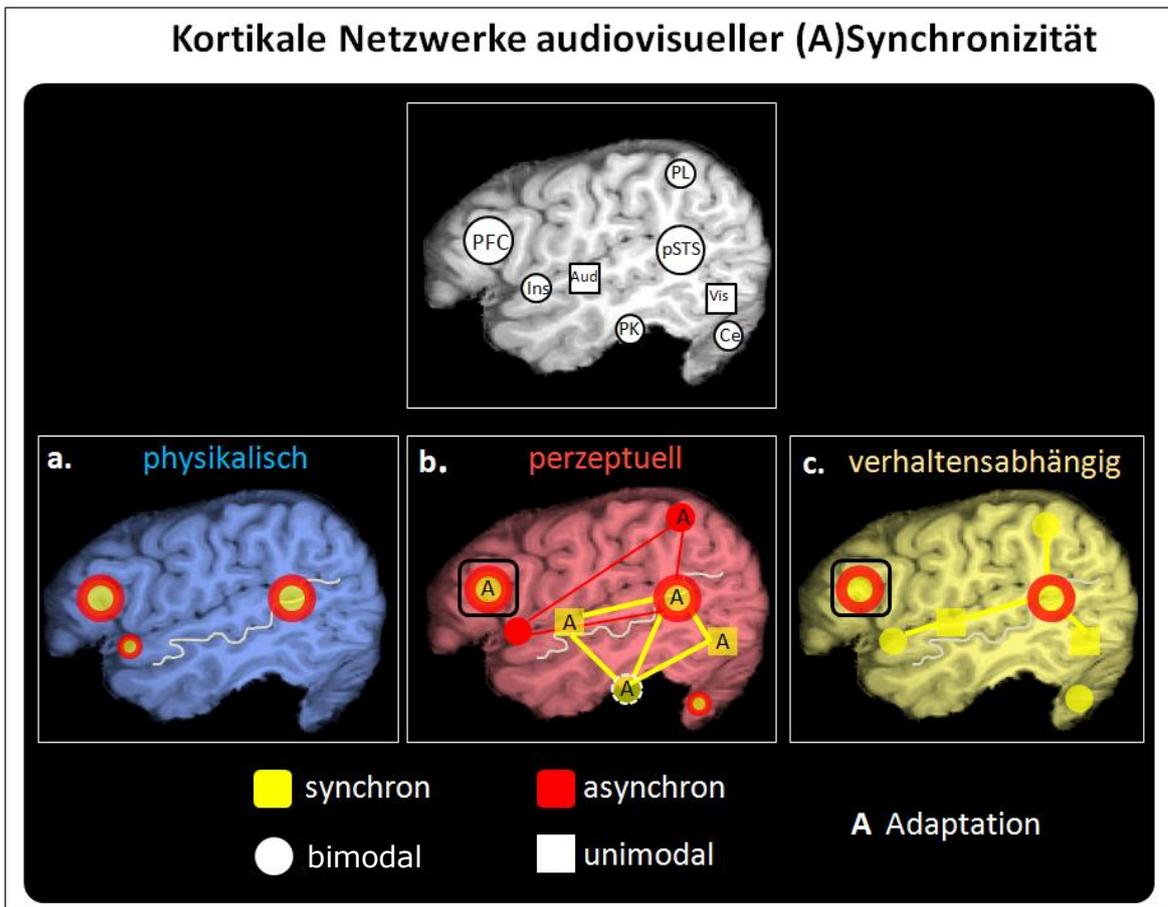
Zusammengefasst wurde in der oben aufgeführten Untersuchung ein funktionelles Muster innerhalb des STS ermittelt, das einer zeitlichen Anordnung folgt und eine Interaktion aus perzeptuellem Zustand und physikalischen Eigenschaften audiovisueller Sprachstimuli reflektiert. Diese Datenlage weist daher auf eine Repräsentation audiovisueller, zeitlicher Muster hin, die eine bestimmte räumliche Anordnung aufweist und den menschlichen multisensorischen STS darin in unterschiedliche Subregionen unterteilt, die über differentielle neuronale Verbindungen zu multisensorischen präfrontalen Regionen verfügen.

## 4 Allgemeine Diskussion

Wenn ein Lichtreiz die Retina stimuliert und ein Tonsignal gleichzeitig das Trommelfell, tritt die entsprechende Aktivierung des Gehirns 30 – 50 ms früher auf als die des auditorischen Signals (King & Palmer, 1985). Menschen scheinen neuronale und physikalische Zeitverschiebungen anhand von Rekalibrierungsmechanismen kompensieren zu können, um eine optimale Verarbeitung audiovisueller Gleichzeitigkeit zu gewährleisten. Die vorliegende Arbeit untersuchte die zugrundeliegenden neuronalen Prozesse audiovisueller Zeitverarbeitung. Dabei wurden physikalische (wahrnehmungsunabhängige, s. Abb. 42 a), perzeptuelle (stimulusunabhängige, s. Abb. 42 b) und verhaltenskorrelierte Effekte (s. Abb. 42 c) anhand einfacher und sprachlicher Reize untersucht. Darüber hinaus wurden die neuronalen Korrelate zeitlicher Kompensationsmechanismen untersucht. Die Ergebnisse zeigen ein kortikales Netzwerk audiovisueller Zeitlichkeit, das aus Präfrontal-, Temporal-, Parietalkortex, Insula, Cerebellum und unisensorischen auditorischen und visuellen Arealen besteht (s. Abb. 42). Die anteriore Insula wurde durch asynchrone Perzepte, verhaltenskorrelierte Synchronizität und durch bloße Anwesenheit synchroner und asynchroner Reize moduliert. Dieses Gebiet könnte demzufolge Abgleichungsprozesse mit subjektiv repräsentierter Gleichzeitigkeit koordinieren. Das Cerebellum wurde ausschließlich durch aufgabenabhängige Prozesse (Antwortverhalten und Perzeption) moduliert und reflektiert in diesem Zusammenhang vermutlich mit Gleichzeitigkeit verbundene Entscheidungsprozesse.

Lateraler Temporalkortex, PFC und Insula sind an der Verarbeitung sowohl synchroner, als auch asynchroner Ereignisse beteiligt. Diese Areale wurden durch alle untersuchten Effekte moduliert und repräsentieren daher vermutlich ein Basisnetzwerk audiovisueller Zeitprozesse (s. Abb. 42 b). Parietale Beteiligung variiert in Abhängigkeit asynchroner Perzeption, unisensorische und medial temporale in Abhängigkeit synchroner Perzeption. Parietale, medial temporale und unisensorische Gebiete sind außerdem in zeitliche Adaptationsprozesse im Sinne einer Kurzzeitplastizität involviert. Im Untersuchungskontext der vorliegenden Arbeit stellen diese Areale und der PFC Gebiete dar, die Änderungen zeitlicher Relationen audiovisueller Ereignisse kompensieren können. Im Zusammenhang mit ihren spezifischen Aktivierungen durch

Synchronizität (pSTS, medialer Temporalkortex, unisensorische Kortices) und Asynchronizität (Parietalkortex) stellen sie vermutlich Regionen dar, die im Sinne dynamischer Muster (s. 1.2.3) an der Erstellung stabiler synchroner und asynchroner Perzepte beteiligt sind. Da der PFC sowohl synchrone, als auch asynchrone Reize verarbeitet und an Adaptationsprozessen beider Stimuluskonstellationen beteiligt ist, kann vermutet werden, dass synchrone bzw. asynchrone Muster durch kognitive Prozesse im PFC reguliert werden. Dies könnten Mechanismen wie das Einbeziehen von Vorerfahrungen in bezug auf Gleichzeitigkeit oder Analysen hinsichtlich der Zuverlässigkeit der etablierten Muster im Sinne des Bayes'schen Ansatzes multimodaler Verarbeitung sein.



**Abbildung 42.** Vorschlag über kortikale Netzwerke bei audiovisueller Zeitwahrnehmung einfacher und sprachlicher Stimuli auf Grundlage der Ergebnisse der vorliegenden Arbeit. Obere Reihe: Legende aller beteiligten Hirnareale *Aud*: auditorischer Kortex; *Ce*: Cerebellum; *Ins*: anteriore Insula; *PFC*: Präfrontalkortex; *PL*: Parietallappen; *PK*: Perirhinaler Kortex/Gyrus parahippocampalis; *pSTS*: posteriorer Sulcus temporalis superior; *Vis*: visueller Kortex. Untere Reihe: Der anatomische Verlauf des STS wurde durch einen grauen Linienzug hervorgehoben. **a.** Die bloße Präsentation gleichzeitiger (gelbe Formen) und ungleichzeitiger (rote Formen) audiovisueller Ereignisse kann durch bimodale Aktivität in pSTS, anteriorer Insula und PFC verarbeitet werden. Dieses Netzwerk scheint Ereignisse unspezifisch von ihrer zeitlichen Relation zueinander verarbeiten zu können. **b.** Wahrnehmungen audiovisueller Gleichzeitigkeit (gelbe Formen) modulieren ein Netzwerk aus unisensorischen visuellen und auditorischen Arealen (eckige Formen) und bimodalem (runde Formen) medialen Temporalkortex (Gyrus parahippocampalis/perirhinaler Kortex) und pSTS. Asynchrone Perzepte (rote Formen) modulieren ein Netzwerk aus bimodalem Parietalkortex, anteriorer Insula und pSTS. Dabei scheint der pSTS sowohl synchrone, als auch asynchrone Perzepte zu verarbeiten, stabile Perzepte entstehen jedoch vermutlich durch medial temporale (synchrone Perzepte) und parietale (asynchrone Perzepte) Interaktionen mit dem pSTS (farbige Linien). Dabei nimmt der PFC vermutlich eine moderierende Funktion (schwarzer Rahmen) ein, die etwa Berücksichtigung von Vorwissen über audiovisuelle Zeiterfahrungen vermitteln kann. Die beteiligten Areale (Parietalkortex, pSTS, medialer Temporalkortex, unisensorische Areale und PFC) können außerdem plastisch kurzfristige Änderungen der zeitlichen Relationen der Umwelt im Sinne einer Adaption (A) berücksichtigen und gewährleisten auf diese Weise die zuverlässige, flexible Integration audiovisueller Ereignisse hinsichtlich ihrer Gleichzeitigkeit. Anteriore Insula und Cerebellum werden wahrscheinlich durch entscheidungsabhängige bzw. antizipatorische Prozesse moduliert. **c.** Effekte, die mit dem Antwortverhalten in bezug auf synchrone Wahrnehmung korrelieren, so dass Stimulation und Wahrnehmung vorwiegend übereinstimmen, wurden in unisensorischen Kortices, bimodalem pSTS, Parietalkortex, anteriorer Insula und Cerebellum identifiziert. Dabei wurden PFC und pSTS auch durch Antwortverhalten hinsichtlich (mit der Stimulation vorwiegend übereinstimmender) asynchroner Perzepte moduliert. Der subjektive Bezug auf Synchronizität scheint dabei bevorzugt verarbeitet zu werden, da er mehr neuronale Ressourcen beansprucht als asynchrone Antwortprozesse. Der PFC könnte dabei durch moderierende Einflüsse (schwarzer Rahmen) an der Weiterverarbeitung der eingetroffenen audiovisuellen Zeitinformation beteiligt sein, die ggf. für Asynchronizität und Synchronizität unterschiedlich beurteilt würde.

Semantische und semantisch nicht belegte Reize verwenden gemäß der hier vorgestellten Ergebnisse zur Etablierung zeitlicher Perzepte größtenteils überlappende Netzwerke, wobei an der Verarbeitung sprachlicher Reize neben posterioren auch anteriore Gebiete des STS beteiligt waren, die anderen Untersuchungen zufolge durch stimmenabhängige Prozesse zu erklären sind (Kriegstein & Giraud, 2004). In bezug auf physikalische Prozesse wurde sichtbar, dass synchrone Sprachreize ausschließlich rechtslaterale Modulationen hervorriefen, wohingegen lateralisierte Muster während der

bloßen Präsentation einfacher Reize nur bei Asynchronizität auftraten (linkslateral). Des Weiteren wurde die anteriore Insula durch synchrone Sprachreize moduliert, bei einfachen Reizen jedoch ausschließlich durch Asynchronizität. Falls dieses Gebiet, wie oben erwähnt, aktuelle externe Reize mit subjektiver Gleichzeitigkeit abgleicht, könnte dies ein Hinweis darauf sein, dass Synchronizität bei Sprachreizen im Sinne der Einheitlichkeitsannahme robuster etabliert wird, als bei (,unnatürlichen') einfachen Reizen. Ein Abgleich würde bei synchronen Sprachreizen aufgrund der Eindeutigkeit audiovisueller Synchronizität nicht initiiert, jedoch bei einfachen synchronen Reizen, da eine Annahme über die Zusammengehörigkeit zweier Informationsströme bei diesem Stimulustyp wahrscheinlich nicht unmittelbar offensichtlich ist.

Einen weiteren Unterschied im Aktivierungsmuster sprachbesetzter und einfacher Reize stellt die Aktivierung des medialen Temporalkortex während perzeptueller Synchronizität bei einfachen Reizen dar, die durch Sprachstimuli nicht nachgewiesen werden konnte. Darüber hinaus wurden durch perzeptuelle Synchronizität (innerhalb bimodaler Masken) unisensorische Gebiete bei einfachen Stimuli moduliert, in denen durch denselben Vergleich bei Sprachreizen keine erhöhten Antwortmuster gefunden wurden. Dies könnte ein Hinweis darauf sein, dass der mediale Temporalkortex an der Beurteilung kurzer zeitlicher Muster beteiligt ist und mit unisensorischen Kortices kommuniziert, um eine Gleichzeitigkeitsinterpretation dieser kurzen Sequenzen zu ermöglichen. Lange Sprachreize würden die zeitlichen Verarbeitungszyklen des medialen Temporalkortex überschreiten. Diese Stimulustypen werden gemäß den oben berichteten Ergebnissen vermehrt durch pSTS und PFC analysiert, und nicht im medialen Temporalkortex. Asynchrone Perzepte bei einfachen Stimuli modulierten im Gegensatz zu Sprachreizen den ACC und Parietalkortex. Dies könnte ein Hinweis darauf sein, dass diese Gebiete insbesondere in die Fehlerregulation zeitlicher Intervallperzepte bei einfachen Stimuli involviert sind, wohingegen bei Sprachreizen, wenn ein kontinuierlicher Informationsstrom vorliegt, dieser Modus des zeitlichen Abgleichs bei Asynchronizität nicht möglich ist.

Anhand semantischer Reize konnte über den Befund eines kortikalen Netzwerkes audiovisueller (Un-)Gleichzeitigkeit hinaus eine funktionelle Trennung des STS in Subregionen identifiziert werden, die qualitativ unterschiedliche Zeitperzepte enkodieren (,auditorisch vor visuell', ,synchron', und ,visuell vor auditorisch').

Die Daten der vorliegenden Arbeit legen nahe, dass eine räumliche Reorganisation neuronaler Aktivität für die erfolgreiche Verarbeitung audiovisueller Zeitinformationen bestimmend zu sein scheint und dabei chronotopischen Prinzipien gefolgt wird.

## 5 LITERATURVERZEICHNIS

- Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nat Neurosci*, 7(10), 1057-1058.
- Adams, W. J., & Mamassian, P. (2004). Bayesian combination of ambiguous shape cues. *J Vis*, 4(10), 921-929.
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol*, 14(3), 257-262.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn Sci*, 4(7), 267-278.
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Curr Biol*, 15(9), 839-843.
- Amedi, A., Jacobson, G., Hendler, T., Malach, R., & Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb Cortex*, 12(11), 1202-1212.
- Anastasio, T. J., Patton, P. E., & Belkacem-Boussaid, K. (2000). Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Comput*, 12(5), 1165-1187.
- Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annu Rev Neurosci*, 25, 189-220.
- Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Brain Res Cogn Brain Res*, 21(3), 301-308.
- Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Res*, 45(10), 1275-1284.
- Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *J Neurophysiol*, 84(1), 451-459.
- Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Res Brain Res Rev*, 22(3), 229-244.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556-559.
- Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, 41(3), 473-487.
- Bahrack, L. E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *J Exp Child Psychol*, 53(2), 180-199.
- Banich, M. T., Milham, M. P., Jacobson, B. L., Webb, A., Wszalek, T., Cohen, N. J., et al. (2001). Attentional selection and the processing of task-irrelevant information: insights from fMRI examinations of the Stroop task. *Prog Brain Res*, 134, 459-470.

- Barbas, H., Medalla, M., Alade, O., Suski, J., Zikopoulos, B., & Lera, P. (2005). Relationship of prefrontal connections to inhibitory systems in superior temporal areas in the rhesus monkey. *Cereb Cortex*, 15(9), 1356-1370.
- Barense, M. D., Bussey, T. J., Lee, A. C., Rogers, T. T., Davies, R. R., Saksida, L. M., et al. (2005). Functional specialization in the human medial temporal lobe. *J Neurosci*, 25(44), 10239-10246.
- Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. (1993). Coloured speech perception: is synaesthesia what happens when modularity breaks down? *Perception*, 22(4), 419-426.
- Barracough, N. E., Xiao, D., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J Cogn Neurosci*, 17(3), 377-391.
- Bartels, A., & Zeki, S. (2000). The architecture of the colour centre in the human visual brain: new results and a review. *Eur J Neurosci*, 12(1), 172-193.
- Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The 'when' pathway of the right parietal lobe. *Trends Cogn Sci*, 11(5), 204-210.
- Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1987). Functional subdivisions of the temporal lobe neocortex. *J Neurosci*, 7(2), 330-342.
- Beauchamp, M. S. (2005). Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics*, 3(2), 93-113.
- Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004a). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci*, 7(11), 1190-1192.
- Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004b). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, 41(5), 809-823.
- Beeli, G., Esslen, M., & Jancke, L. (2005). Synaesthesia: when coloured sounds taste sweet. *Nature*, 434(7029), 38.
- Beeli, G., Esslen, M., & Jancke, L. (2008). Time course of neural activity correlated with colored-hearing synesthesia. *Cereb Cortex*, 18(2), 379-385.
- Beer, J., Blakemore, C., Previc, F. H., & Liotti, M. (2002). Areas of the human brain activated by ambient visual motion, indicating three kinds of self-movement. *Exp Brain Res*, 143(1), 78-88.
- Belardinelli, M. O., Sestieri, C., Di Matteo, R., Delgou, F., del Gratta, C., Ferretti, A., et al. (2004). Audio-visual crossmodal interactions in environmental perception: an fMRI investigation. *Cogn Process*, 5, 167-174.
- Belin, P., McAdams, S., Thivard, L., Smith, B., Savel, S., Zilbovicius, M., et al. (2002). The neuroanatomical substrate of sound duration discrimination. *Neuropsychologia*, 40(12), 1956-1964.

- Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, *14*(16), 2105-2109.
- Benevento, L. A., Fallon, J., Davis, B. J., & Rezak, M. (1977). Auditory--visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Exp Neurol*, *57*(3), 849-872.
- Bennett, R. G., & Westheimer, G. (1985). A shift in the perceived simultaneity of adjacent visual stimuli following adaptation to stroboscopic motion along the same axis. *Vision Res*, *25*(4), 565-569.
- Bermudez-Rattoni, F., & McGaugh, J. L. (1991). Insular cortex and amygdala lesions differentially affect acquisition on inhibitory avoidance and conditioned taste aversion. *Brain Res*, *549*(1), 165-170.
- Bernstein, L. E., Auer, E. T., Jr., Moore, J. K., Ponton, C. W., Don, M., & Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *Neuroreport*, *13*(3), 311-315.
- Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept Psychophys*, *29*(6), 578-584.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex*, *10*(5), 512-528.
- Bischoff, M., Walter, B., Blecker, C. R., Morgen, K., Vaitl, D., & Sammer, G. (2007). Utilizing the ventriloquism-effect to investigate audio-visual binding. *Neuropsychologia*, *45*(3), 578-586.
- Bishop, C. W., & Miller, L. M. (2008). A Multisensory Cortical Network for Understanding Speech in Noise. *J Cogn Neurosci*.
- Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007a). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb Cortex*, *17*(9), 2172-2189.
- Bizley, J. K., Nodal, F. R., Parsons, C. H., & King, A. J. (2007b). Role of auditory cortex in sound localization in the midsagittal plane. *J Neurophysiol*, *98*(3), 1763-1774.
- Bodner, M., Kroger, J., & Fuster, J. M. (1996). Auditory memory cells in dorsolateral prefrontal cortex. *Neuroreport*, *7*(12), 1905-1908.
- Bolognini, N., Rasi, F., & Ladavas, E. (2005). Visual localization of sounds. *Neuropsychologia*, *43*(11), 1655-1661.
- Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H. J., et al. (2007). Neural basis of the ventriloquist illusion. *Curr Biol*, *17*(19), 1697-1703.
- Bonda, E., Frey, S., & Petrides, M. (1996). Evidence for a dorso-medial parietal system involved in mental transformations of the body. *J Neurophysiol*, *76*(3), 2042-2048.

- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev*, *108*(3), 624-652.
- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, *29*(1), 287-296.
- Bresciani, J. P., & Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *Neuroreport*, *18*(11), 1157-1161.
- Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci*, *25*(29), 6797-6806.
- Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118-1121.
- Brown, J. W., & Braver, T. S. (2007). Risk prediction and aversion by anterior cingulate cortex. *Cogn Affect Behav Neurosci*, *7*(4), 266-277.
- Brown, S. W. (1985). Time perception and attention: the effects of prospective versus retrospective paradigms and task demands on perceived duration. *Percept Psychophys*, *38*(2), 115-124.
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol*, *46*(2), 369-384.
- Buchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*(6690), 274-277.
- Buchsbaum, B. R., & D'Esposito, M. (2009). Repetition suppression and reactivation in auditory-verbal short-term recognition memory. *Cereb Cortex*, *19*(6), 1474-1485.
- Buckley, M. J., Booth, M. C., Rolls, E. T., & Gaffan, D. (2001). Selective perceptual impairments after perirhinal cortex ablation. *J Neurosci*, *21*(24), 9824-9836.
- Budinger, E., Heil, P., Hess, A., & Scheich, H. (2006). Multisensory processing via early cortical stages: Connections of the primary auditory cortical field with other sensory systems. *Neuroscience*, *143*(4), 1065-1083.
- Budinger, E., Laszcz, A., Lison, H., Scheich, H., & Ohl, F. W. (2008). Non-sensory cortical and subcortical connections of the primary auditory cortex in Mongolian gerbils: bottom-up and top-down processing of neuronal information via field AI. *Brain Res*, *1220*, 2-32.
- Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: mental imagery of complex sounds. *Neuroimage*, *26*(4), 1119-1127.
- Buonomano, D. V., & Karmarkar, U. R. (2002). How do we tell time? *Neuroscientist*, *8*(1), 42-51.

- Buonomano, D. V., & Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, 267(5200), 1028-1030.
- Bushara, K. O., Grafman, J., & Hallett, M. (2001). Neural correlates of auditory-visual stimulus onset asynchrony detection. *J Neurosci*, 21(1), 300-304.
- Bushara, K. O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., & Hallett, M. (2003). Neural correlates of cross-modal binding. *Nat Neurosci*, 6(2), 190-195.
- Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat Neurosci*, 2(8), 759-766.
- Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proc Natl Acad Sci U S A*, 102(51), 18751-18756.
- Bussey, T. J., Saksida, L. M., & Murray, E. A. (2002). Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *Eur J Neurosci*, 15(2), 365-374.
- Caclin, A., & Fonlupt, P. (2006). Functional and effective connectivity in an fMRI study of an auditory-related task. *Eur J Neurosci*, 23(9), 2531-2537.
- Caclin, A., Soto-Faraco, S., Kingstone, A., & Spence, C. (2002). Tactile "capture" of audition. *Percept Psychophys*, 64(4), 616-630.
- Callan, D. E., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. *Brain Res Cogn Brain Res*, 10(3), 349-353.
- Callan, D. E., Jones, J. A., Munhall, K., Kroos, C., Callan, A. M., & Vatikiotis-Bateson, E. (2004). Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. *J Cogn Neurosci*, 16(5), 805-816.
- Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb Cortex*, 11(12), 1110-1123.
- Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, 10(12), 2619-2623.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276(5312), 593-596.
- Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: the neural substrates of visible speech. *J Cogn Neurosci*, 15(1), 57-70.
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol*, 10(11), 649-657.

- Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, *14*(2), 427-438.
- Calvert, G. A., & Lewis, J. W. (2004). Hemodynamic Studies of Audiovisual Integration. In G. A. Calvert, C. Spence & B. E. Stein (Eds.), *The Handbook of Multisensory Processes* (pp. 483-502). Cambridge, Massachusetts; London: The MIT Press.
- Campbell, R., Garwood, J., Franklin, S., Howard, D., Landis, T., & Regard, M. (1990). Neuropsychological studies of auditory-visual fusion illusions. Four case studies and their implications. *Neuropsychologia*, *28*(8), 787-802.
- Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J Exp Psychol*, *84*(1), 141-147.
- Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur J Neurosci*, *22*(11), 2886-2902.
- Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol*, *363*(4), 615-641.
- Carr, C. E. (1993). Processing of temporal information in the brain. *Annu Rev Neurosci*, *16*, 223-243.
- Casini, L., & Ivry, R. B. (1999). Effects of divided attention on temporal processing in patients with lesions of the cerebellum or frontal lobe. *Neuropsychology*, *13*(1), 10-21.
- Cavada, C., & Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J Comp Neurol*, *287*(4), 422-445.
- Chavis, D. A., & Pandya, D. N. (1976). Further observations on corticofrontal connections in the rhesus monkey. *Brain Res*, *117*(3), 369-386.
- Cocosco, C. A., Zijdenbos, A. P., & Evans, A. C. (2003). A fully automatic and robust brain MRI tissue classification method. *Med Image Anal*, *7*(4), 513-527.
- Cohen, Y. E., Batista, A. P., & Andersen, R. A. (2002). Comparison of neural activity preceding reaches to auditory and visual stimuli in the parietal reach region. *Neuroreport*, *13*(6), 891-894.
- Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761-773.
- Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *J Cogn Neurosci*, *14*(3), 508-523.

- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J Neurosci*, *11*(8), 2383-2402.
- Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *J Neurosci*, *13*(3), 1202-1226.
- Coull, J., & Nobre, A. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr Opin Neurobiol*, *18*(2), 137-144.
- Coull, J. T. (2004). fMRI studies of temporal attention: allocating attention within, or towards, time. *Brain Res Cogn Brain Res*, *21*(2), 216-226.
- Cunningham, D. W., Billock, V. A., & Tsou, B. H. (2001). Sensorimotor adaptation to violations of temporal contiguity. *Psychol Sci*, *12*(6), 532-535.
- Cunningham, D. W., Chatziastros, A., von der Heyde, M., & Bulthoff, H. H. (2001). Driving in the future: temporal visuomotor adaptation and generalization. *J Vis*, *1*(2), 88-98.
- Cusick, C. G., Seltzer, B., Cola, M., & Griggs, E. (1995). Chemoarchitectonics and corticocortical terminations within the superior temporal sulcus of the rhesus monkey: evidence for subdivisions of superior temporal polysensory cortex. *J Comp Neurol*, *360*(3), 513-535.
- Dawson, K. A. (2004). Temporal organization of the brain: Neurocognitive mechanisms and clinical implications. *Brain Cogn*, *54*(1), 75-94.
- De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, *7*(10), 460-467.
- Deneve, S., Latham, P. E., & Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nat Neurosci*, *4*(8), 826-831.
- Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behav. Brain Sci.*, *15*, 183-247.
- Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Res*, *178*(2-3), 363-380.
- Devlin, J. T., & Price, C. J. (2007). Perirhinal contributions to human visual perception. *Curr Biol*, *17*(17), 1484-1488.
- Dhamala, M., Assisi, C. G., Jirsa, V. K., Steinberg, F. L., & Kelso, J. A. (2007). Multisensory integration for timing engages different brain networks. *Neuroimage*, *34*(2), 764-773.
- Dix, M. R., Hallpike, C. S., & Hood, J. D. (1949). Auditory adaptation in the human subject. *Nature*, *164*(4158), 59.
- Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*(6), 719-721.

- Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Res*, 1242, 136-150.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat Neurosci*, 3(3), 277-283.
- Dreher, J. C., & Berman, K. F. (2002). Fractionating the neural substrate of cognitive control processes. *Proc Natl Acad Sci U S A*, 99(22), 14595-14600.
- Driver, J., & Mattingley, J. B. (1998). Parietal neglect and visual awareness. *Nat Neurosci*, 1(1), 17-22.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, 57(1), 11-23.
- Driver, J., & Spence, C. (1998). Crossmodal attention. *Curr Opin Neurobiol*, 8(2), 245-253.
- Droit-Volet, S., Meck, W. H., & Penney, T. B. (2007). Sensory modality and time perception in children and adults. *Behav Processes*, 74(2), 244-250.
- Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J Neurophysiol*, 79(1), 126-136.
- Durston, S., Davidson, M. C., Thomas, K. M., Worden, M. S., Tottenham, N., Martinez, A., et al. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *Neuroimage*, 20(4), 2135-2141.
- Efron, R. (1970). The minimum duration of a perception. *Neuropsychologia*, 8(1), 57-63.
- Eimer, M. (1999). Can attention be directed to opposite locations in different modalities? An ERP study. *Clin Neurophysiol*, 110(7), 1252-1259.
- Eimer, M., & Schröger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology*, 35(3), 313-327.
- Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. In G. Knoblich, I. M. Thornton, M. Grosjean & M. Shiffrar (Eds.), *Human Body Perception From The Inside Out* (pp. 105-131). New York: NY: Oxford University Press.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429-433.
- Esterman, M., Verstynen, T., Ivry, R. B., & Robertson, L. C. (2006). Coming unbound: disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *J Cogn Neurosci*, 18(9), 1570-1576.
- Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *Eur J Neurosci*, 29(6), 1247-1257.

- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci*, 22(13), 5749-5759.
- Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Percept Psychophys*, 63(4), 719-725.
- Fischer, H., Andersson, J. L., Furmark, T., & Fredrikson, M. (1998). Brain correlates of an unexpected panic attack: a human positron emission tomographic study. *Neurosci Lett*, 251(2), 137-140.
- Fonlupt, P. (2003). Perception and judgement of physical causality involve different brain structures. *Brain Res Cogn Brain Res*, 17(2), 248-254.
- Formisano, E., Kim, D. S., Di Salle, F., van de Moortele, P. F., Ugurbil, K., & Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron*, 40(4), 859-869.
- Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res Cogn Brain Res*, 10(1-2), 77-83.
- Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16(5), 419-423.
- Frahm, J., Merboldt, K. D., & Hanicke, W. (1994). The influence of the slice-selection gradient on functional MRI of human brain activation. *J Magn Reson B*, 103(1), 91-93.
- Fraisse, P. (1984). Perception and estimation of time. *Annu Rev Psychol*, 35, 1-36.
- Frassinetti, F., Bolognini, N., & Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp Brain Res*, 147(3), 332-343.
- Friedman, D. P., Murray, E. A., O'Neill, J. B., & Mishkin, M. (1986). Cortical connections of the somatosensory fields of the lateral sulcus of macaques: evidence for a corticolimbic pathway for touch. *J Comp Neurol*, 252(3), 323-347.
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, 6(3), 218-229.
- Friston, K. J., Frith, C. D., Frackowiak, R. S., & Turner, R. (1995). Characterizing dynamic brain responses with fMRI: a multivariate approach. *Neuroimage*, 2(2), 166-172.
- Friston, K. J., Holmes, A. P., Price, C. J., Büchel, C., & Worsley, K. J. (1999). Multisubject fMRI studies and conjunction analyses. *Neuroimage*, 10(4), 385-396.
- Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Exp Brain Res*, 166(3-4), 455-464.
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci*, 7(7), 773-778.

- Fuster, J. M. (1985). The prefrontal cortex and temporal integration. In J. M. Fuster (Ed.), *The Cerebral Cortex* (Vol. 4, pp. 151-177). New York: Plenum Press.
- Fuster, J. M. (1987). Single-Unit Studies of the Prefrontal Cortex. In E. Perecman (Ed.), *The Frontal Lobes Revisited* (pp. 109-120). New York: IRBN Press.
- Fuster, J. M., Bodner, M., & Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, *405*(6784), 347-351.
- Galaburda, A. M., & Pandya, D. N. (1983). The intrinsic architectonic and connectional organization of the superior temporal region of the rhesus monkey. *J Comp Neurol*, *221*(2), 169-184.
- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11-21.
- Gebhard, J. W., & Mowbray, G. H. (1959). On discriminating the rate of visual flicker and auditory flutter. *Am J Psychol*, *72*, 521-529.
- Geiser, E., Zaehle, T., Jancke, L., & Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *J Cogn Neurosci*, *20*(3), 541-552.
- Ghazanfar, A. A., Chandrasekaran, C., & Logothetis, N. K. (2008). Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J Neurosci*, *28*(17), 4457-4469.
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci*, *25*(20), 5004-5012.
- Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn Sci*, *10*(6), 278-285.
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci*, *11*(5), 473-490.
- Gil-da-Costa, R., Braun, A., Lopes, M., Hauser, M. D., Carson, R. E., Herscovitch, P., et al. (2004). Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque. *Proc Natl Acad Sci U S A*, *101*(50), 17516-17521.
- Goerke, U., Moller, H. E., Norris, D. G., & Schwarzbauer, C. (2005). A comparison of signal instability in 2D and 3D EPI resting-state fMRI. *NMR Biomed*, *18*(8), 534-542.
- Gonzalo, D., Shallice, T., & Dolan, R. (2000). Time-dependent changes in learning audiovisual associations: a single-trial fMRI study. *Neuroimage*, *11*(3), 243-255.
- Gorno-Tempini, M. L., Price, C. J., Josephs, O., Vandenberghe, R., Cappa, S. F., Kapur, N., et al. (1998). The neural systems sustaining face and proper-name processing. *Brain*, *121* (Pt 11), 2103-2118.

- Goulet, S., & Murray, E. A. (2001). Neural substrates of crossmodal association memory in monkeys: the amygdala versus the anterior rhinal cortex. *Behav Neurosci*, *115*(2), 271-284.
- Grant, K. W., von Wassenhove, V., & Poeppel, D. (2003). Discrimination of Auditory-Visual Synchrony. *unknown*.
- Graybiel, A. M. (1973). The thalamo-cortical projection of the so-called posterior nuclear group: a study with anterograde degeneration methods in the cat. *Brain Res*, *49*(2), 229-244.
- Graziano, M. S., Andersen, R. A., & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. *J Neurosci*, *14*(1), 54-67.
- Graziano, M. S., Reiss, L. A., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, *397*(6718), 428-430.
- Grefkes, C., Weiss, P. H., Zilles, K., & Fink, G. R. (2002). Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron*, *35*(1), 173-184.
- Grèzes, J., Costes, N., & Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain*, *122* ( Pt 10), 1875-1887.
- Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. *Percept Psychophys*, *49*(3), 245-256.
- Grossberg, S., & Schmajuk, N. A. (1989). Neural Dynamics of Adaptive Timing and Temporal Discrimination During Associative Learning. *Neural Networks*, *2*, 79-102.
- Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends Cogn Sci*, *5*(1), 36-41.
- Guldin, W. O., & Markowitsch, H. J. (1984). Cortical and thalamic afferent connections of the insular and adjacent cortex of the cat. *J Comp Neurol*, *229*(3), 393-418.
- Hamilton, R. H., Shenton, J. T., & Coslett, H. B. (2006). An acquired deficit of audiovisual speech processing. *Brain Lang*, *98*(1), 66-73.
- Hanggi, J., Beeli, G., Oechslin, M. S., & Jancke, L. (2008). The multiple synaesthete E.S.: neuroanatomical basis of interval-taste and tone-colour synaesthesia. *Neuroimage*, *43*(2), 192-203.
- Harrington, D. L., Haaland, K. Y., & Hermanowicz, N. (1998a). Temporal processing in the basal ganglia. *Neuropsychology*, *12*(1), 3-12.
- Harrington, D. L., Haaland, K. Y., & Knight, R. T. (1998b). Cortical networks underlying mechanisms of time perception. *J Neurosci*, *18*(3), 1085-1095.
- Hasegawa, T., Matsuki, K., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., et al. (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Brain Res Cogn Brain Res*, *20*(3), 510-518.

- Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *J Neurosci*, 28(10), 2539-2550.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, 4(6), 223-233.
- Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci*, 9(6), 467-479.
- Heilman, K. M., & Valenstein, E. (1972). Frontal lobe neglect in man. *Neurology*, 22(6), 660-664.
- Hein, G., & Knight, R. T. (2008). Superior temporal sulcus--It's my area: or is it? *J Cogn Neurosci*, 20(12), 2125-2136.
- Helmholtz, H. v. (1896). *Handbuch der physiologischen Optik* (2 ed.). Hamburg: Voss.
- Herdener, M., Lehmann, C., Esposito, F., di Salle, F., Federspiel, A., Bach, D. R., et al. (2009). Brain responses to auditory and visual stimulus offset: shared representations of temporal edges. *Hum Brain Mapp*, 30(3), 725-733.
- Hetherington, R., Dennis, M., & Spiegler, B. (2000). Perception and estimation of time in long-term survivors of childhood posterior fossa tumors. *J Int Neuropsychol Soc*, 6(6), 682-692.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99.
- Hicks, T. P., Benedek, G., & Thurlow, G. A. (1988). Modality specificity of neuronal responses within the cat's insula. *J Neurophysiol*, 60(2), 422-437.
- Hikosaka, K., Iwai, E., Saito, H., & Tanaka, K. (1988). Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *J Neurophysiol*, 60(5), 1615-1637.
- Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science*, 298(5598), 1627-1630.
- Hinrichs, H., Scholz, M., Tempelmann, C., Woldorff, M. G., Dale, A. M., & Heinze, H. J. (2000). Deconvolution of event-related fMRI responses in fast-rate experimental designs: tracking amplitude variations. *J Cogn Neurosci*, 12 Suppl 2, 76-89.
- Holdstock, J. S., Hocking, J., Notley, P., Devlin, J. T., & Price, C. J. (2009). Integrating Visual and Tactile Information in the Perirhinal Cortex. *Cereb Cortex*.
- Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. *Neuropsychologia*, 45(14), 3340-3345.
- Holmes, N. P. (2009). The Principle of Inverse Effectiveness in Multisensory Integration: Some Statistical Considerations. *Brain Topogr*.

- Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nat Neurosci*, 3(3), 284-291.
- Howard, R. J., Brammer, M., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. *Curr Biol*, 6(8), 1015-1019.
- Hoyt, D. F., & Taylor, R. C. (1981). Gait and the energetics of locomotion in horses. *Nature*, 292, 239-240.
- Hu, D., Yan, L., Liu, Y., Zhou, Z., Friston, K. J., Tan, C., et al. (2005). Unified SPM-ICA for fMRI analysis. *Neuroimage*, 25(3), 746-755.
- Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron*, 48(3), 509-520.
- Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., et al. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *AJNR Am J Neuroradiol*, 19(4), 659-671.
- Ivry, R. (1993). Cerebellar involvement in the explicit representation of temporal information. *Ann N Y Acad Sci*, 682, 214-230.
- Ivry, R. (1997). Cerebellar timing systems. *Int Rev Neurobiol*, 41, 555-573.
- Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Curr Opin Neurobiol*, 6(6), 851-857.
- Ivry, R. B., Keele, S. W., & Diener, H. C. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. *Exp Brain Res*, 73(1), 167-180.
- Ivry, R. B., & Spencer, R. M. (2004). Evaluating the role of the cerebellum in temporal processing: beware of the null hypothesis. *Brain*, 127(Pt 8), E13; author reply E14.
- Jancke, L. (2005). *Methoden der Bildgebung in der Psychologie und den kognitiven Neurowissenschaften* (1 ed.). Stuttgart: Kohlhammer.
- Jancke, L., Beeli, G., Eulig, C., & Hanggi, J. (2009). The neuroanatomy of grapheme-color synesthesia. *Eur J Neurosci*, 29(6), 1287-1293.
- Jancke, L., & Shah, N. J. (2004). Hearing syllables by seeing visual stimuli. *Eur J Neurosci*, 19(9), 2603-2608.
- Jancke, L., Wustenberg, T., Scheich, H., & Heinze, H. J. (2002). Phonetic perception and the temporal cortex. *Neuroimage*, 15(4), 733-746.
- Jaskowski, P., Jaroszyk, F., & Hojan-Jeziarska, D. (1990). Temporal-order judgments and reaction time for stimuli of different modalities. *Psychol Res*, 52(1), 35-38.
- Jiang, W., Jiang, H., Rowland, B. A., & Stein, B. E. (2007). Multisensory orientation behavior is disrupted by neonatal cortical ablation. *J Neurophysiol*, 97(1), 557-562.

- Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., & Stein, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *J Neurophysiol*, *85*(2), 506-522.
- Johnson, J. A., & Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage*, *31*(4), 1673-1681.
- Jones, E. G., & Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, *93*(4), 793-820.
- Jones, J. A., & Callan, D. E. (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport*, *14*(8), 1129-1133.
- Joseph, J. E., Cerullo, M. A., Farley, A. B., Steinmetz, N. A., & Mier, C. R. (2006). fMRI correlates of cortical specialization and generalization for letter processing. *Neuroimage*, *32*(2), 806-820.
- Kaiser, J., Hertrich, I., Ackermann, H., Mathiak, K., & Lutzenberger, W. (2005). Hearing lips: gamma-band activity during audiovisual speech perception. *Cereb Cortex*, *15*(5), 646-653.
- Kanowski, M., Rieger, J. W., Noesselt, T., Tempelmann, C., & Hinrichs, H. (2007). Endoscopic eye tracking system for fMRI. *J Neurosci Methods*, *160*(1), 10-15.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, *17*(11), 4302-4311.
- Kawashima, R., Imaizumi, S., Mori, K., Okada, K., Goto, R., Kiritani, S., et al. (1999). Selective visual and auditory attention toward utterances—a PET study. *Neuroimage*, *10*(2), 209-215.
- Kawashima, R., Okuda, J., Umetsu, A., Sugiura, M., Inoue, K., Suzuki, K., et al. (2000). Human cerebellum plays an important role in memory-timed finger movement: an fMRI study. *J Neurophysiol*, *83*(2), 1079-1087.
- Kayser, C., & Logothetis, N. K. (2007). Do early sensory cortices integrate cross-modal information? *Brain Struct Funct*, *212*(2), 121-132.
- Kayser, C., & Logothetis, N. K. (2009). Directed Interactions Between Auditory and Superior Temporal Cortices and their Role in Sensory Integration. *Front Integr Neurosci*, *3*, 7.
- Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron*, *48*(2), 373-384.
- Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *J Neurosci*, *27*(8), 1824-1835.
- Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb Cortex*, *18*(7), 1560-1574.

- Kelso, J. A. (1984). Phase transitions and critical behavior in human bimanual coordination. *Am J Physiol*, 246(6 Pt 2), R1000-1004.
- Kelso, J. A. S. (1997). *Dynamic patterns - the self-organization of brain and behavior*. Cambridge, Massachusetts: MIT Press.
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303(5660), 1023-1026.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annu Rev Psychol*, 55, 271-304.
- King, A. J. (2005). Multisensory integration: strategies for synchronization. *Curr Biol*, 15(9), R339-341.
- King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res*, 60(3), 492-500.
- King, D. P., & Takahashi, J. S. (2000). Molecular genetics of circadian rhythms in mammals. *Annu Rev Neurosci*, 23, 713-742.
- Knudsen, E. I., & Knudsen, P. F. (1985). Vision guides the adjustment of auditory localization in young barn owls. *Science*, 230(4725), 545-548.
- Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244-247.
- Kriegstein, K. V., & Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, 22(2), 948-955.
- Lagarde, J., & Kelso, J. A. (2006). Binding of movement, sound and touch: multimodal coordination dynamics. *Exp Brain Res*, 173(4), 673-688.
- Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53(2), 279-292.
- Lalonde, R., & Hannequin, D. (1999). The neurobiological basis of time estimation and temporal order. *Rev Neurosci*, 10(2), 151-173.
- Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y. F., Field, A. S., & Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *J Cogn Neurosci*, 14(3), 420-429.
- Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., & Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp Brain Res*, 166(3-4), 289-297.
- Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, 38(2), 317-327.

- Leung, H. C., Gore, J. C., & Goldman-Rakic, P. S. (2002). Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *J Cogn Neurosci*, *14*(4), 659-671.
- Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex*, *10*(9), 873-888.
- Lewis, J. W., & Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J Comp Neurol*, *428*(1), 112-137.
- Liègeois-Chauvel, C., de Graaf, J. B., Laguitton, V., & Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex*, *9*(5), 484-496.
- Loe, P. R., & Benevento, L. A. (1969). Auditory-visual interaction in single units in the orbito-insular cortex of the cat. *Electroencephalogr Clin Neurophysiol*, *26*(4), 395-398.
- Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci*, *23*(10), 3963-3971.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150-157.
- Loose, R., Kaufmann, C., Auer, D. P., & Lange, K. W. (2003). Human prefrontal and sensory cortical activity during divided attention tasks. *Hum Brain Mapp*, *18*(4), 249-259.
- Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat Neurosci*, *9*(11), 1432-1438.
- Ma, W. J., & Pouget, A. (2008). Linking neurons to behavior in multisensory perception: a computational review. *Brain Res*, *1242*, 4-12.
- Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, *39*(12), 1304-1316.
- Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci*, *28*(5), 264-271.
- Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(5482), 1206-1208.
- Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage*, *21*(2), 725-732.
- MacDonald, A. W., 3rd, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*(5472), 1835-1838.

- Mangels, J. A., Ivry, R. B., & Shimizu, N. (1998). Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Brain Res Cogn Brain Res*, 7(1), 15-39.
- Manoach, D. S., White, N. S., Lindgren, K. A., Heckers, S., Coleman, M. J., Dubal, S., et al. (2004). Hemispheric specialization of the lateral prefrontal cortex for strategic processing during spatial and shape working memory. *Neuroimage*, 21(3), 894-903.
- Maravita, A., Spence, C., & Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Curr Biol*, 13(13), R531-539.
- Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., et al. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb Cortex*, 17(7), 1672-1679.
- Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays*, 22(1), 94-103.
- Mathiak, K., Hertrich, I., Grodd, W., & Ackermann, H. (2004). Discrimination of temporal information at the cerebellum: functional magnetic resonance imaging of nonverbal auditory memory. *Neuroimage*, 21(1), 154-162.
- Matlin, M. W., & Foley, H. J. (1997). *Sensation and perception* (4 ed.). Boston u.a.: Allyn and Bacon.
- Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature*, 410(6828), 580-582.
- Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu Rev Neurosci*, 27, 307-340.
- McDonald, J. J., Teder-Salejarvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, 407(6806), 906-908.
- McGrath, M., & Summerfield, Q. (1985). Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *J Acoust Soc Am*, 77(2), 678-685.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.
- Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *J Exp Psychol Anim Behav Process*, 9(2), 171-201.
- Menon, V., Adelman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Hum Brain Mapp*, 12(3), 131-143.
- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci*, 7(10), 3215-3229.
- Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221(4608), 389-391.

- Meredith, M. A., & Stein, B. E. (1985). Descending efferents from the superior colliculus relay integrated multisensory information. *Science*, 227(4687), 657-659.
- Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. *Ann Neurol*, 10(4), 309-325.
- Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann Neurol*, 28(5), 597-613.
- Mesulam, M. M. (1998). From sensation to cognition. *Brain*, 121 ( Pt 6), 1013-1052.
- Mesulam, M. M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philos Trans R Soc Lond B Biol Sci*, 354(1387), 1325-1346.
- Meyer, M., Baumann, S., Marchina, S., & Jancke, L. (2007). Hemodynamic responses in human multisensory and auditory association cortex to purely visual stimulation. *BMC Neurosci*, 8, 14.
- Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Comput*, 1, 359-371.
- Middlebrooks, J. C., & Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *J Neurosci*, 4(10), 2621-2634.
- Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nat Rev Neurosci*, 1(1), 59-65.
- Miller, L. M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *J Neurosci*, 25(25), 5884-5893.
- Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav Brain Res*, 6(1), 57-77.
- Mishra, J., Martinez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Res*, 1242, 102-115.
- Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J Neurosci*, 27(15), 4120-4131.
- Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat Neurosci*, 9(7), 875-877.
- Mollon, J. D., & Perkins, A. J. (1996). Errors of judgement at Greenwich in 1796. *Nature*, 380(6570), 101-102.
- Mottronen, R., Calvert, G. A., Jaaskelainen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., et al. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage*, 30(2), 563-569.
- Mottronen, R., Krause, C. M., Tiippana, K., & Sams, M. (2002). Processing of changes in visual speech in the human auditory cortex. *Brain Res Cogn Brain Res*, 13(3), 417-425.

- Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Percept Psychophys*, 58(3), 351-362.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol*, 83(5), 2580-2601.
- Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., et al. (2005). Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex*, 15(7), 963-974.
- Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47, 90-100.
- Nagy, A., Eordeghe, G., Paroczky, Z., Markus, Z., & Benedek, G. (2006). Multisensory integration in the basal ganglia. *Eur J Neurosci*, 24(3), 917-924.
- Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neurosci Lett*, 413(1), 72-76.
- Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Brain Res Cogn Brain Res*, 25(2), 499-507.
- Noesselt, T., Bergmann, D., Hake, M., Heinze, H. J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Res*, 1220, 157-163.
- Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H. J., et al. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *J Neurosci*, 27(42), 11431-11441.
- Noppeney, U., Josephs, O., Hocking, J., Price, C. J., & Friston, K. J. (2008). The effect of prior visual information on recognition of speech and sounds. *Cereb Cortex*, 18(3), 598-609.
- Ochiai, T., Grimault, S., Scavarda, D., Roch, G., Hori, T., Riviere, D., et al. (2004). Sulcal pattern and morphology of the superior temporal sulcus. *Neuroimage*, 22(2), 706-719.
- Odgaard, E. C., Arieh, Y., & Marks, L. E. (2004). Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Cogn Affect Behav Neurosci*, 4(2), 127-132.
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A*, 87(24), 9868-9872.
- Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., et al. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophys J*, 64(3), 803-812.

- Ojanen, V., Mottonen, R., Pekkola, J., Jaaskelainen, I. P., Joensuu, R., Autti, T., et al. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage*, 25(2), 333-338.
- Okada, M., & Kashino, M. (2003). The role of spectral change detectors in temporal order judgment of tones. *Neuroreport*, 14(2), 261-264.
- Olson, I. R., Gatenby, J. C., & Gore, J. C. (2002). A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Brain Res Cogn Brain Res*, 14(1), 129-138.
- Onoe, H., Komori, M., Onoe, K., Takechi, H., Tsukada, H., & Watanabe, Y. (2001). Cortical networks recruited for time perception: a monkey positron emission tomography (PET) study. *Neuroimage*, 13(1), 37-45.
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J Neurophysiol*, 76(1), 109-129.
- Padberg, J., Seltzer, B., & Cusick, C. G. (2003). Architectonics and cortical connections of the upper bank of the superior temporal sulcus in the rhesus monkey: an analysis in the tangential plane. *J Comp Neurol*, 467(3), 418-434.
- Pandya, D. N. (1995). Anatomy of the auditory cortex. *Rev Neurol (Paris)*, 151(8-9), 486-494.
- Patton, P. E., & Anastasio, T. J. (2003). Modeling cross-modal enhancement and modality-specific suppression in multisensory neurons. *Neural Comput*, 15(4), 783-810.
- Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D., Goldstein, L., Heather, J., et al. (1995). The physiology of coloured hearing. A PET activation study of colour-word synaesthesia. *Brain*, 118 ( Pt 3), 661-676.
- Pekkola, J., Laasonen, M., Ojanen, V., Autti, T., Jaaskelainen, I. P., Kujala, T., et al. (2006). Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: an fMRI study at 3 T. *Neuroimage*, 29(3), 797-807.
- Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *J Neurosci*, 23(17), 6819-6825.
- Penney, T. B., Gibbon, J., & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *J Exp Psychol Hum Percept Perform*, 26(6), 1770-1787.
- Penney, T. B., & Turrett, S. (2005). Les effets de la modalité sensorielle sur la perception du temps. *Psychologie Française*, 50(1), 131-143.
- Pernet, C., Celsis, P., & Demonet, J. F. (2005). Selective response to letter categorization within the left fusiform gyrus. *Neuroimage*, 28(3), 738-744.
- Petrides, M., & Pandya, D. N. (1988). Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J Comp Neurol*, 273(1), 52-66.
- Piaget, J. (1974). *Der Aufbau der Wirklichkeit beim Kinde*. Stuttgart: Klett.

- Pick, H. L., Jr., Warren, D. H., & Hay, J. C. (1969). Sensory conflicts in judgments of spatial direction. *Perception & Psychophysics*, 6, 203-205.
- Ploghaus, A., Tracey, I., Gati, J. S., Clare, S., Menon, R. S., Matthews, P. M., et al. (1999). Dissociating pain from its anticipation in the human brain. *Science*, 284(5422), 1979-1981.
- Pöppel, E. (1985). *Grenzen des Bewusstseins. Wie kommen wir zur Zeit und wie entsteht Wirklichkeit*. Stuttgart: Deutsche Verlags-Anstalt.
- Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L., & Mishkin, M. (2003). Functional mapping of the primate auditory system. *Science*, 299(5606), 568-572.
- Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *J Neurosci*, 28(11), 2745-2752.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci*, 18(6), 2188-2199.
- Rajj, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, 28(2), 617-625.
- Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nat Neurosci*, 4(3), 317-323.
- Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, 268(5207), 111-114.
- Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc Natl Acad Sci U S A*, 95(3), 869-875.
- Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *J Neurophysiol*, 89(2), 1078-1093.
- Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition*, 98(1), 53-84.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp Brain Res*, 111(2), 246-252.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (1997). Parietal cortex: from sight to action. *Curr Opin Neurobiol*, 7(4), 562-567.
- Rizzuto, D. S., Mamelak, A. N., Sutherling, W. W., Fineman, I., & Andersen, R. A. (2005). Spatial selectivity in human ventrolateral prefrontal cortex. *Nat Neurosci*, 8(4), 415-417.
- Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int J Psychophysiol*, 50(1-2), 19-26.

- Rockland, K. S., & Van Hoesen, G. W. (1994). Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cereb Cortex*, 4(3), 300-313.
- Romanski, L. M., Bates, J. F., & Goldman-Rakic, P. S. (1999). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J Comp Neurol*, 403(2), 141-157.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb Cortex*, 17(5), 1147-1153.
- Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. *J Cogn Neurosci*, 19(9), 1435-1452.
- Rousseau, R., Poirier, J., & Lemyre, L. (1983). Duration discrimination of empty time intervals marked by intermodal pulses. *Percept Psychophys*, 34(6), 541-548.
- Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nat Neurosci*, 10(6), 792-797.
- Rubinstein, L., & Gruenberg, E. M. (1971). Intramodal and crossmodal sensory transfer of visual and auditory temporal patterns. *Perception and Psychophysics*, 9, 385-390.
- Sagiv, N., & Ward, J. (2006). Crossmodal interactions: lessons from synesthesia. *Prog Brain Res*, 155, 259-271.
- Saito, D. N., Yoshimura, K., Kochiyama, T., Okada, T., Honda, M., & Sadato, N. (2005). Cross-modal binding and activated attentional networks during audio-visual speech integration: a functional MRI study. *Cereb Cortex*, 15(11), 1750-1760.
- Saunders, J. A., & Knill, D. C. (2001). Perception of 3D surface orientation from skew symmetry. *Vision Res*, 41(24), 3163-3183.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu Rev Psychol*, 55, 87-124.
- Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, 267(5198), 699-701.
- Schmidt, C. F., Zaehle, T., Meyer, M., Geiser, E., Boesiger, P., & Jancke, L. (2008). Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. *Hum Brain Mapp*, 29(1), 46-56.
- Schoenfeld, M. A., Tempelmann, C., Martinez, A., Hopf, J. M., Sattler, C., Heinze, H. J., et al. (2003). Dynamics of feature binding during object-selective attention. *Proc Natl Acad Sci U S A*, 100(20), 11806-11811.

- Schroeder, C. E., & Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res*, 14(1), 187-198.
- Schroeder, C. E., & Foxe, J. J. (2004). Multisensory Convergence in Early Cortical Processing. In G. A. Calvert, C. Spence & B. E. Stein (Eds.), *The Handbook of Multisensory Processes*. Cambridge, Massachusetts; London: The MIT Press.
- Schroeder, C. E., Mehta, A. D., & Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb Cortex*, 8(7), 575-592.
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123 Pt 12, 2400-2406.
- Sekiyama, K., Kanno, I., Miura, S., & Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neurosci Res*, 47(3), 277-287.
- Seltzer, B., & Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *J Comp Neurol*, 281(1), 97-113.
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., et al. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Curr Biol*, 10(1), 16-22.
- Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., & Woldorff, M. G. (2007). Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia*, 45(3), 561-571.
- Sestieri, C., Di Matteo, R., Ferretti, A., Del Gratta, C., Caulo, M., Tartaro, A., et al. (2006). "What" versus "where" in the audiovisual domain: an fMRI study. *Neuroimage*, 33(2), 672-680.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, 408(6814), 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Brain Res Cogn Brain Res*, 14(1), 147-152.
- Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport*, 16(17), 1923-1927.
- Shapleske, J., Rossell, S. L., Woodruff, P. W., & David, A. S. (1999). The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. *Brain Res Brain Res Rev*, 29(1), 26-49.
- Shipley, T. (1964). Auditory Flutter-Driving of Visual Flicker. *Science*, 145, 1328-1330.
- Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *J Neurosci*, 24(47), 10702-10706.

- Shomstein, S., & Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *J Neurosci*, 26(2), 435-439.
- Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb Cortex*, 17(10), 2387-2399.
- Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, 12(1), 7-10.
- Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. *J Cogn Neurosci*, 13(7), 930-936.
- Smith, A. L., Parsons, C. H., Lanyon, R. G., Bizley, J. K., Akerman, C. J., Baker, G. E., et al. (2004). An investigation of the role of auditory cortex in sound localization using muscimol-releasing Elvax. *Eur J Neurosci*, 19(11), 3059-3072.
- Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc Natl Acad Sci U S A*, 95(20), 12061-12068.
- Spelke, E. S. (1979). Perceiving Bimodally Specified Events in Infancy. *Developmental Psychology*, 15(6), 626-686.
- Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *J Exp Psychol Gen*, 130(4), 799-832.
- Spence, C., & Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Curr Biol*, 13(13), R519-521.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J Comp Physiol Psychol*, 43(6), 482-489.
- Staddon, J. E., & Higa, J. J. (1999). Time and memory: towards a pacemaker-free theory of interval timing. *J Exp Anal Behav*, 71(2), 215-251.
- Stein, B. E., Jiang, W., & Stanford, T. R. (2004). Multisensory Integration in Single Neurons of the Midbrain In G. A. Calvert, C. Spence & B. E. Stein (Eds.), *The Handbook of Multisensory Processes* (pp. 243-264). Cambridge, Massachusetts; London: The MIT Press.
- Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J Cogn Neurosci*, 8, 497-506.
- Stein, B. E., & Meredith, M. E. (1993). *The merging of the sense*. Cambridge, MA: The MIT Press.
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci*, 9(4), 255-266.
- Stekelenburg, J. J., Vroomen, J., & de Gelder, B. (2004). Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neurosci Lett*, 357(3), 163-166.

- Stevenson, R. A., Geoghegan, M. L., & James, T. W. (2007). Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Exp Brain Res*, 179(1), 85-95.
- Stevenson, R. A., & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage*, 44(3), 1210-1223.
- Stevenson, R. A., Kim, S., & James, T. W. (2009). An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Exp Brain Res*.
- Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001). When is now? Perception of simultaneity. *Proc Biol Sci*, 268(1462), 31-38.
- Stratton, G. M. (1896). Some preliminary experiments on vision without inversion of the retinal image. *Psychol. Rev.*, 3, 611-617.
- Sugihara, T., Diltz, M. D., Averbeck, B. B., & Romanski, L. M. (2006). Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *J Neurosci*, 26(43), 11138-11147.
- Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, 421(6926), 911.
- Sukhbinder, K., Klaas, E. S., Warren, J. D., Friston, K. J., & Griffiths, T. D. (2007). Hierarchical Processing of Auditory Objects in Humans. *PLOS Computational Biology*, 3(6), 977-985.
- Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in noise. *J. Acoust. Soc. Am.*, 26, 212-215.
- Suzuki, W. A. (1996). The anatomy, physiology and functions of the perirhinal cortex. *Curr Opin Neurobiol*, 6(2), 179-186.
- Suzuki, W. A., & Amaral, D. G. (1994). Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J Comp Neurol*, 350(4), 497-533.
- Szycik, G. R., Jansma, H., & Munte, T. F. (2008a). Audiovisual integration during speech comprehension: An fMRI study comparing ROI-based and whole brain analyses. *Hum Brain Mapp*.
- Szycik, G. R., Tausche, P., & Munte, T. F. (2008b). A novel approach to study audiovisual integration in speech perception: localizer fMRI and sparse sampling. *Brain Res*, 1220, 142-149.
- Tanabe, H. C., Honda, M., & Sadato, N. (2005). Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. *J Neurosci*, 25(27), 6409-6418.
- Taylor, K. I., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2006). Binding crossmodal object features in perirhinal cortex. *Proc Natl Acad Sci U S A*, 103(21), 8239-8244.

- Templeton, W. B., Howard, I. P., & Lowman, A. E. (1966). Passively generated adaptation to prismatic distortion. *Percept Mot Skills*, 22(1), 140-142.
- Thompson, J. C., Clarke, M., Stewart, T., & Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *J Neurosci*, 25(39), 9059-9066.
- Timmann, D., Watts, S., & Hore, J. (1999). Failure of cerebellar patients to time finger opening precisely causes ball high-low inaccuracy in overarm throws. *J Neurophysiol*, 82(1), 103-114.
- Tootell, R. B., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., et al. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature*, 375(6527), 139-141.
- Topka, H., Valls-Sole, J., Massaquoi, S. G., & Hallett, M. (1993). Deficit in classical conditioning in patients with cerebellar degeneration. *Brain*, 116 ( Pt 4), 961-969.
- Toplak, M. E., & Tannock, R. (2005). Time perception: modality and duration effects in attention-deficit/hyperactivity disorder (ADHD). *J Abnorm Child Psychol*, 33(5), 639-654.
- Tuomainen, J., Andersen, T. S., Tiippana, K., & Sams, M. (2005). Audio-visual speech perception is special. *Cognition*, 96(1), B13-22.
- van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, 43(2), 271-282.
- van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb Cortex*, 17(4), 962-974.
- van de Par, S., & Kohlrausch, A. (1999). Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *J Acoust Soc Am*, 106(4 Pt 1), 1940-1947.
- van Wassenhove, V., Buonomano, D. V., Shimojo, S., & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE*, 3(1), e1437.
- van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc Natl Acad Sci U S A*, 102(4), 1181-1186.
- van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, 45(3), 598-607.
- Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Temporal recalibration during asynchronous audiovisual speech perception. *Exp Brain Res*, 181(1), 173-181.
- Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2008). Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments. *Exp Brain Res*, 185(3), 521-529.

- Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain Res*, 1111(1), 134-142.
- Vatakis, A., & Spence, C. (2007). Crossmodal binding: evaluating the "unity assumption" using audiovisual speech stimuli. *Percept Psychophys*, 69(5), 744-756.
- Verstichel, P. (2003). Aphasie thalamique. *Revue Neurologique*, 159(10), 947-957.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143-156.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P., et al. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, 14(1 Pt 1), 170-181.
- Vohn, R., Fimm, B., Weber, J., Schnitker, R., Thron, A., Spijkers, W., et al. (2007). Management of attentional resources in within-modal and cross-modal divided attention tasks: an fMRI study. *Hum Brain Mapp*, 28(12), 1267-1275.
- Voisin, J., Bidet-Caulet, A., Bertrand, O., & Fonlupt, P. (2006). Listening in silence activates auditory areas: a functional magnetic resonance imaging study. *J Neurosci*, 26(1), 273-278.
- Vroomen, J., & De Gelder, B. (2004). Perceptual Effects of Cross-Modal Stimulation: Ventriloquism and the Freezing Phenomenon In G. A. Calvert, C. Spence & B. E. Stein (Eds.), *The Handbook of Multisensory Processes*. Cambridge, Massachusetts, London: The MIT Press.
- Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res Cogn Brain Res*, 22(1), 32-35.
- Wallace, M. T., Carriere, B. N., Perrault, T. J., Jr., Vaughan, J. W., & Stein, B. E. (2006). The development of cortical multisensory integration. *J Neurosci*, 26(46), 11844-11849.
- Wallace, M. T., Perrault, T. J., Jr., Hairston, W. D., & Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *J Neurosci*, 24(43), 9580-9584.
- Wang, Q., & Burkhalter, A. (2007). Area map of mouse visual cortex. *J Comp Neurol*, 502(3), 339-357.
- Warren, D. H. (1979). Spatial localization under conflict conditions: is there a single explanation? *Perception*, 8(3), 323-337.
- Watkins, S., Dalton, P., Lavie, N., & Rees, G. (2007). Brain mechanisms mediating auditory attentional capture in humans. *Cereb Cortex*, 17(7), 1694-1700.
- Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., et al. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex*, 3(2), 79-94.

- Watson, R. T., Heilman, K. M., Cauthen, J. C., & King, F. A. (1973). Neglect after cinglectomy. *Neurology*, *23*(9), 1003-1007.
- Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": application of a model of the internal clock in humans. *Q J Exp Psychol B*, *51*(2), 97-120.
- Weiss, P. H., Zilles, K., & Fink, G. R. (2005). When visual perception causes feeling: enhanced cross-modal processing in grapheme-color synesthesia. *Neuroimage*, *28*(4), 859-868.
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nat Neurosci*, *5*(6), 598-604.
- Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *J Neurosci*, *24*(48), 10941-10949.
- Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Percept Psychophys*, *39*(4), 294-300.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychol Bull*, *88*(3), 638-667.
- Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cereb Cortex*, *13*(10), 1034-1043.
- Wundt, W. (1887). *Grundzüge der physiologischen Psychologie* (3 ed.). Leipzig: Engelmann.
- Xing, J., & Andersen, R. A. (2000). Memory activity of LIP neurons for sequential eye movements simulated with neural networks. *J Neurophysiol*, *84*(2), 651-665.
- Xu, J., Bradley, J. S., & Gover, B. N. (2005). An artificial neural network approach for predicting architectural speech security. *J Acoust Soc Am*, *117*(4 Pt 1), 1709-1712.
- Xu, Y., Turk-Browne, N. B., & Chun, M. M. (2007). Dissociating task performance from fMRI repetition attenuation in ventral visual cortex. *J Neurosci*, *27*(22), 5981-5985.
- Yoo, S. S., Paralkar, G., & Panych, L. P. (2004). Neural substrates associated with the concurrent performance of dual working memory tasks. *Int J Neurosci*, *114*(6), 613-631.
- Yuille, A., & Bülthoff, H. H. (1996). Bayesian decision theory and psychophysics. In D. Knill & W. Richards (Eds.), *Perception as Bayesian Inference* (pp. 123-161). Cambridge: Cambridge University Press.
- Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *J Exp Psychol Gen*, *130*(1), 29-58.
- Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. *Percept Psychophys*, *67*(3), 531-544.
- Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Exp Brain Res*, *152*(2), 198-210.

- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*, 6(1), 37-46.
- Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *J Neurosci*, 11(3), 641-649.
- Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nat Neurosci*, 6(2), 175-181.

## 6 Anhang

**Tabelle 7.** Lokale Maxima der uni- und potentiell bimodalen (Überlappung von unisensorischer Aktivität) Masken zu Experiment 1.

<i>Anatomische Struktur</i>	<i>Hemi- sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<b>MNI Koordinaten</b> <b>x y z</b>		
<i>(a) auditorisch &gt; baseline &amp; visuell &gt; baseline (potentiell bimodal) (p &lt; 0,05; k &gt; 100)</i>						
<i>Temporale Gebiete</i>						
<i>pSTS/STG/MTG/SMG/Insula/ Gyrus praecentralis/PT/IPS</i>	R	3426	3,9 (0,000)	56	-40	16
<i>aSTS/HG/Insula</i>	L	183	2,76 (0,004)	-40	-18	-14
<i>pSTS/STG/MTG/SMG/MFG/Gyrus prae-, postcentralis/PT</i>	L	4959	4,9 (0,000)	-48	-40	10
<i>Frontale Gebiete</i>						
<i>Gyrus praecentralis/IFG/MFG</i>	R	1797	3,48 (0,001)	46	26	16
<i>MFG/SFG</i>	R	597	3,27 (0,001)	14	14	60
<i>MFG/IFG</i>	L	363	2,79 (0,004)	-48	50	-2
<i>Gyrus cinguli</i>	L	477	3,53 (0,000)	-8	-22	38
<i>Parietale Gebiete</i>						
<i>Gyrus postcentralis /Lobulus paracentralis</i>	R	294	2,75 (0,004)	22	-26	52
<i>IPS/Lobulus parietalis superior</i>	L	287	3,45 (0,001)	-24	-44	44
<i>Okzipitale Gebiete</i>						
<i>Cerebellum/IOG</i>	R	179	3,78 (0,000)	42	-64	-22
<i>Cerebellum/MOG</i>	R	536	3,75 (0,000)	16	-78	-20
<i>Cerebellum/MOG</i>	L	540	3,27 (0,001)	-44	-66	-22

**(b) auditorisch > baseline (exkl. 'auditorisch > baseline & visuell > baseline')**  
( $p < 0,05$ ;  $k > 100$ )

<i>Temporale Gebiete</i>						
STG/HG/PT/Insula/STS/ Lobulus parietalis superior/Gyrus supramarginalis/angularis/Gyrus cinguli	R	16141	6,71 (0,000)	62	-22	6
STG/HG/PT/Insula/STS/ Lobulus parietalis superior/Gyrus supramarginalis/angularis/MFG/IFG	L	10938	9,25 (0,000)	-40	-30	8
<i>Frontale Gebiete</i>						
MFG	R	242	3,77 (0,000)	36	18	48
anteriore Insula	L	108	3,06 (0,002)	-32	20	-8
<i>Okzipitale Gebiete</i>						
SOG	R	256	3,69 (0,000)	26	-74	8

**(c) visuell > baseline (exkl. 'auditorisch > baseline & visuell > baseline')** ( $p < 0,05$ ;  $k > 100$ )

<i>Okzipitale Gebiete</i>						
ITG/FG	R	106	3,09 (0,002)	26	-8	-38
<i>Frontale Gebiete</i>						
Gyrus praecentralis /IOG/ITG/MTG/IPS/ Gyrus angularis/supramarginalis /FG/LG/Cerebellum/IFG/ MFG	L/R	32881	6,57 (0,000)	-38	-8	46
MFG	L	734	3,63 (0,000)	-44	42	22
MFG/IFG	L	158	2,92 (0,003)	-48	48	2

<i>Insula/Planum polare</i>	L	118	2,61 (0,006)	-38	-4	-24
<i>Subkortikale Gebiete</i>						
<i>Putamen</i>	R	158	3,14 (0,001)	24	-10	2

**Tabelle 8.** Lokale Maxima der experimentellen Kontraste von Experiment 1. Physikalische Effekte (bloße Stimuluspräsentation) wurden überprüft.

<i>Anatomische Struktur</i>	<i>Hemi- sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<b>MNI Koordinaten</b> <b>x y z</b>		
<i>(a) synchrone &gt; asynchrone Stimuli (p &lt; 0,05; k &gt; 30)</i>						
<i>bimodale Maske</i>						
<i>Temporale Gebiete</i>						
<i>MTG</i>	L	77	2,86 (0,002)	-64	-52	2
<i>STG/STS</i>	L	62	2,69 (0,004)	-42	-40	24
<i>Frontale Gebiete</i>						
<i>IFG</i>	R	332	3,07 (0,001)	50	18	4
<i>IFG</i>	R	158	2,51 (0,007)	44	30	-2
<i>Insula</i>	L	56	2,65 (0,004)	-42	-2	4
<i>MFG</i>	L	46	2,49 (0,007)	-32	4	34
<i>Parietale Gebiete</i>						
<i>SMG</i>	L	30	2,41 (0,009)	-58	-44	48
<i>auditorische Maske</i>						
<i>MFG</i>	L	31	2,41 (0,009)	-32	0	36

**(b) asynchrone > synchrone Stimuli ( $p < 0,05$ ;  $k > 30$ )****bimodale Maske***Temporale Gebiete*

<i>pSTS</i>	L	30	2,33 (0,01)	-54	-38	0
<i>mSTS</i>	L	58	2,27 (0,01)	-52	-56	24

**auditorische Maske***Frontale Gebiete*

<i>Operculum Frontale</i>	R	53	2,72 (0,004)	42	-16	26
---------------------------	---	----	-----------------	----	-----	----

**(c) Stimuli Ton zuerst (-125 ms/-63 ms) > synchrone Stimuli ( $p < 0,05$ ;  $k > 30$ )****visuelle Maske***Frontale Gebiete*

<i>Gyrus praecentralis</i>	30	R	2,41 (0,009)	42	-6	48
----------------------------	----	---	-----------------	----	----	----

**(d) Stimuli Licht zuerst (125 ms/ 63 ms) > synchrone Stimuli ( $p < 0,05$ ;  $k > 30$ )****bimodale Maske***Temporale Gebiete*

<i>pSTS</i>	76	L	2,62 (0,005)	-52	-56	24
<i>mSTS</i>	40	L	2,57 (0,006)	-52	-36	2
<i>STG</i>	42	R	2,46 (0,008)	56	8	-10

<i>Cerebellum</i>						
<i>Cerebellum</i>	31	R	2,79 (0,003)	44	-66	-24
<i>auditorische Maske</i>						
<i>Frontale Gebiete</i>						
<i>Operculum frontale</i>	58	R	3,22 (0,001)	42	-16	24

**Tabelle 9.** Lokale Maxima der experimentellen Kontraste von Experiment 1. Die Tabelle stellt perzeptuelle Effekte (unabhängig vom Stimulustyp) dar. Diese wurden über alle Stimulustypen kollabiert, aber auch getrennt nach SOA des Stimulus analysiert (Interaktion aus physikalischem und perzeptuellem Effekt)

<i>Anatomische Struktur</i>	<i>Hemi-sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<b>MNI Koordinaten x y z</b>		
<b>(a) Stimuli Ton zuerst (-125 ms/-63 ms): asynchrone &gt; synchrone Perzeption</b> ( $p < 0,05$ ; $k > 30$ )						
<i>bimodale Maske</i>						
<i>Temporale Gebiete</i>						
<i>mSTS</i>	210	R	2,4 (0,009)	54	-30	-8
<i>Frontale Gebiete</i>						
<i>Insula/IFG/Operculum frontale</i>	475	R	4,31 (0,000)	42	20	-6
<i>Insula/IFG/Operculum frontale</i>	648	L	4,22 (0,000)	-34	22	-2
<i>SF/ACC</i>	445	L/R	3,6 (0,000)	0	18	54
<i>IFG</i>	131	L	2,87 (0,002)	-46	42	-12
<i>IFG</i>	57	L	2,4 (0,009)	-52	-18	30

MFG	210	R	2,71 (0,001)	46	26	30
MFG	221	L	2,82 (0,003)	-40	50	12
<i>Parietale Gebiete</i>						
<hr/>						
IPS/SMG	307	L	3,07 (0,000)	-56	-36	50
IPS	90	R	2,39 (0,009)	38	-50	42
Operculum parietale	46	L	2,31 (0,01)	-48	-20	20
<i>Cerebellum</i>						
<hr/>						
Cerebellum	118	L	3,05 (0,001)	22	-50	-26
Cerebellum	35	R	1,99 (0,02)	-36	-58	-30
<i>auditorische Maske</i>						
<hr/>						
<i>Temporale Gebiete</i>						
<hr/>						
Insula/HG	62	L	2,05 (0,02)	-42	-20	10
<i>visuelle Maske</i>						
<hr/>						
<i>Parietale Gebiete</i>						
<hr/>						
IPS	37	L	2,29 (0,01)	-36	-46	48

**(b) Stimuli Ton zuerst (-125 ms/63 ms): synchrone > asynchrone Perzeption**  
 ( $p < 0,05$ ;  $k > 30$ )

***bimodale Maske***

*Temporale Gebiete*

<i>pSTS/MTG/STG</i>	173	L	3,15 (0,001)	-44	-66	18
<i>mSTS/MTG</i>	71	L	3,18 (0,001)	-60	-24	-10
<i>aSTS</i>	48	L	2,8 (0,003)	-34	-12	-22
<i>aSTS/Planum polare</i>	33	L	2,43 (0,008)	-38	-4	-26
<i>aSTS</i>	43	L	2,36 (0,01)	-56	-6	-12
<i>mSTS</i>	77	R	2,94 (0,002)	38	-20	12
<i>pSTS</i>	39	R	2,55 (0,08)	44	-42	12

*Frontale Gebiete*

<i>MFG</i>	42	L	2,97 (0,002)	-32	16	50
<i>IFG</i>	31	R	2,71 (0,004)	46	30	10

***visuelle Maske***

*Okzipitale Gebiete*

<i>FG</i>	47	R	2,93 (0,002)	32	-44	-18
<i>FG</i>	42	R	2,91 (0,002)	28	-60	-8
<i>MTG/ITG</i>	128	R	2,47 (0,007)	46	-60	-2

(c) *Stimuli synchron: synchrone > asynchrone Perzeption* ( $p < 0,05$ ;  $k > 30$ )

<i>bimodale Maske</i>						
<i>Temporale Gebiete</i>						
MTG/pSTS	140	R	3,12 (0,001)	48	-54	0
pSTS/MTG	171	L	3,2 (0,001)	-50	-66	18
mSTS	336	L	3,48 (0,000)	-52	-24	-6
STG	71	R	2,47 (0,007)	56	8	-10
STG	33	L	2,27 (0,01)	-56	4	-6
Insula/Planum polare	39	L	2,66 (0,004)	-36	0	-22
Gyrus parahippocampalis	92	R	2,97 (0,002)	14	-32	-10
Gyrus parahippocampalis/Cerebellum	86	L	3,1 (0,001)	-22	-32	-22
HG	71	R	2,34 (0,01)	62	-6	4
<i>Frontale Gebiete</i>						
Gyrus praecentralis	159	R	3,57 (0,000)	48	-8	44
Gyrus Post-/praecentralis/ Lobulus paracentralis	117	L	3,12 (0,001)	-10	-32	56
IFG/Gyrus praecentralis	95	R	2,93 (0,002)	62	14	24
IFG	42	R	2,75 (0,003)	34	42	0
Gyrus Orbitalis laterale/IFG	45	L	3,24 (0,001)	-32	42	-10
MFG	44	R	2,88 (0,002)	40	22	50
MFG	264	L	3,18 (0,001)	-32	24	50
MFG/SFG/Sulcus frontalis superior	58	L	2,47 (0,007)	-22	12	58

<i>Okzipitale Gebiete</i>							
<i>Gyrus lingualis/fusifformis</i>	110	R	2,5 (0,007)	14	-68	-8	
MOG	31	R	2,24 (0,01)	28	-78	-22	
<i>auditorische Maske</i>							
<i>Temporale Gebiete</i>							
<i>mSTS/Insula/Planum polare/HG</i>	137	L	3,26 (0,001)	-52	-22	-8	
<i>aSTS/HG/PT</i>	86	R	2,9 (0,002)	62	-2	-12	
<i>Insula</i>	60	R	2,38 (0,009)	38	-16	4	
<i>Frontale Gebiete</i>							
<i>IFG/Gyrus orbitalis laterale</i>	68	L	3,24 (0,001)	-32	42	-10	
<i>visuelle Maske</i>							
<i>Okzipitale Gebiete</i>							
<i>ITG/IOG/FG</i>	415	R	3,74 (0,000)	44	-52	-10	
<i>Frontale Gebiete</i>							
<i>Gyrus praecentralis</i>	158	L	3,64 (0,000)	-44	-14	44	
<i>Gyrus praecentralis</i>	142	R	3,57 (0,000)	48	-8	44	

*(d) Stimuli synchron: asynchrone > synchrone Perzeption (p < 0,05; k > 30)**bimodale Maske**Frontale Gebiete*

<i>Insula/Operculum frontale/IFG/MFG</i>	390	R	3,34 (0,001)	44	20	18
<i>Insula</i>	96	L	3,09 (0,001)	-36	16	-10
<i>Gyrus praecentralis</i>	67	R	2,54 (0,006)	44	6	24
<i>IFG</i>	172	L	3,21 (0,001)	-44	42	-10
<i>MFG</i>	149	L	2,55 (0,006)	-34	38	8
<i>SFG/anteriores Cingulum</i>	156	R	2,54 (0,006)	6	16	56
<i>Operculum frontale</i>	55	L	2,37 (0,01)	-46	18	8

*Parietale Gebiete*

<i>IPS/Gyrus supramarginalis/angularis</i>	156	L	3,58 (0,000)	-52	-38	46
<i>IPS/SMG</i>	62	R	3,29 (0,001)	36	-54	36
<i>Operculum parietale</i>	70	R	2,59 (0,005)	60	-16	22
<i>Operculum parietale</i>	242	L	2,95 (0,002)	-56	-20	22

*Cerebellum*

<i>Cerebellum</i>	179	R	3,20 (0,000)	20	-52	-24
<i>Cerebellum</i>	43	L	2,31 (0,01)	-42	-68	-28

(e) *Stimuli Licht zuerst (125 ms/63 ms): asynchrone > synchrone Perzeption*  
( $p < 0,05$ ;  $k > 30$ )

---

***bimodale Maske***

---

*Frontale Gebiete*

---

<i>Insula/IFG/Operculum frontale</i>	392	L	3,87 (0,000)	-34	20	0
<i>Insula/IFG/Operculum frontale</i>	175	R	3,36 (0,001)	42	20	-6
<i>IFG</i>	103	L	3,29 (0,001)	-44	40	-8
<i>SFG/ACC</i>	217	R	3,27 (0,001)	0	18	48

*Parietale Gebiete*

---

<i>Gyrus supramarginalis/IPS</i>	244	L	3,32 (0,001)	-50	-40	44
<i>Operculum parietale</i>	72	L	2,85 (0,002)	-46	-22	22

*Cerebellum Gebiete*

---

<i>Cerebellum</i>	79	R	3,03 (0,001)	22	-50	-26
-------------------	----	---	-----------------	----	-----	-----

(f) *Stimuli Licht zuerst (125 ms/63 ms): synchrone > asynchrone Perzeption*  
( $p < 0,05$ ;  $k > 30$ )

---

***bimodale Maske***

---

*Temporale Gebiete*

---

<i>mSTS/MTG</i>	1018	R	4,08 (0,000)	46	-46	-4
<i>aSTS/Insula/Planum polare/FG</i>	273	R	3,88 (0,000)	50	-10	-16
<i>aSTS/STG/Gyrus separans</i>	107	R	3,11 (0,001)	52	2	-14

---

<i>mSTS/STG/PT/HG</i>	98	L	2,93 (0,002)	-60	-28	6
<i>pSTS/MTG/STG</i>	1359	L	4,58 (0,000)	-50	-68	14
<i>aSTS/STG/MTG/Insula/Planum polare/PT/HG</i>	1196	L	3,9 (0,000)	-60	-18	-8
<i>PT/STG/Operculum Parietale/ frontale</i>	132	R	3,39 (0,000)	56	-8	10
<i>Gyrus parahippocampalis/ perirhinaler Kortex/Cerebellum</i>	538	R	3,88 (0,000)	20	-30	-20
<i>Cerebellum/Gyrus parahippocampalis/ perirhinaler Kortex</i>	589	L	4,06 (0,000)	-16	-40	-36
<i>Frontale Gebiete</i>						
<i>MFG</i>	85	R	3,96 (0,000)	32	20	44
<i>MFG</i>	59	L	3,0 (0,002)	-22	10	54
<i>Gyrus praecentralis</i>	116	R	3,13 (0,001)	48	-6	46
<i>MFG/Gyrus praecentralis</i>	235	L	3,65 (0,000)	-30	18	44
<i>Gyrus orbitalis laterale</i>	81	R	3,17 (0,001)	32	40	-8
<i>Lobulus paracentralis/ACC</i>	384	L	2,76 (0,003)	-8	-32	56
<i>IFG</i>	65	R	2,62 (0,005)	46	28	8
<i>Gyrus frontomarginalis</i>	34	R	2,14 (0,01)	42	54	-10
<i>Okzipitale Gebiete</i>						
<i>MOG/LG</i>	351	R	3,17 (0,001)	24	-86	-8
<i>MOG/FG</i>	37	R	2,57 (0,006)	50	-68	-10

*auditorische Maske**Temporale Gebiete*

<i>aSTS/HG/STG/Planum polare</i>	684	R	5,04 (0,000)	60	-4	-14
<i>aSTS/HG/STG</i>	115	L	3,51 (0,000)	-58	-18	-8
<i>Planum polare</i>	34	L	2,5 (0,007)	-40	2	-18

*visuelle Maske**Okzipitale Gebiete*

<i>MTG/SOG/pSTS</i>	377	L	4,54 (0,000)	-44	-68	14
<i>FG/MTG/ITG/pSTS</i>	1384	R	4,21 (0,000)	32	-46	-10

*Parietale Gebiete*

<i>IPS/Gyrus angularis/supramarginalis</i>	95	L	3,35 (0,001)	-24	-46	50
--	----	---	-----------------	-----	-----	----

*Frontale Gebiete*

<i>Gyrus praecentralis</i>	142	R	3,13 (0,001)	48	-6	46
<i>Gyrus praecentralis</i>	250	L	3,15 (0,001)	-38	-14	44
<i>Gyrus postcentralis</i>	41	R	3,05 (0,001)	24	-34	54

(g) *alle Stimuli: synchrone > asynchrone Perzeption* ( $p < 0,05$ ;  $k > 30$ )

*bimodale Maske**Temporale Gebiete*

<i>MTG/pSTS</i>	482	R	3,35 (0,001)	48	-50	-6
-----------------	-----	---	-----------------	----	-----	----

<i>mSTS</i>	179	R	3,49 (0,000)	50	-10	-16	
<i>aSTG</i>	61	R	2,75 (0,004)	54	-2	-6	
<i>MTG/pSTS</i>	767	L	4,5 (0,000)	-44	-66	18	
<i>aSTS</i>	778	L	4,06 (0,000)	-34	-12	-22	
<i>Gyrus parahippocampalis</i>	104	R	3,25 (0,001)	20	-28	-18	
<i>Gyrus parahippocampalis/ Cerebellum</i>	339	L	3,89 (0,000)	-22	-32	-22	
<i>Frontale Gebiete</i>							
<i>MFG</i>	132	L	3,91 (0,000)	-30	18	46	
<i>MFG</i>	56	R	3,44 (0,000)	34	22	46	
<i>MFG</i>	36	R	3,15 (0,001)	48	30	10	
<i>MFG</i>	54	L	2,67 (0,004)	-26	12	56	
<i>Lobulus paracentralis</i>	93	L	2,79 (0,003)	-10	-34	58	
<i>Gyrus praecentralis</i>	55	R	2,62 (0,005)	48	-8	46	
<i>Lateraler Gyrus orbitofrontalis</i>	39	R	2,56 (0,006)	34	36	-10	
<i>Lateraler Gyrus orbitofrontalis</i>	34	L	2,43 (0,008)	-34	38	-10	
<i>Okipitale Gebiete</i>							
<i>IOG</i>	50	R	2,51 (0,007)	24	-82	-10	
<i>LG</i>	69	R	2,41 (0,009)	10	-60	-8	
<i>Parietale Gebiete</i>							
<i>Operculum parietale</i>	40	R	2,57 (0,006)	56	-6	10	
<i>Operculum parietale</i>	37	R	2,37 (0,01)	44	-32	22	

<i>Cerebellum</i>						
<hr/>						
<i>Cerebellum</i>	78	L	2,69 (0,004)	-2	-66	-18
<i>Cerebellum</i>	40	R	2,09 (0,02)	8	-46	-26
<b><i>auditorische Maske</i></b>						
<hr/>						
<i>Temporale Gebiete</i>						
<hr/>						
<i>aSTS/HG/PT</i>	360	R	4,26 (0,000)	60	-4	-14
<i>mSTS</i>	64	L	3,35 (0,001)	-52	-22	-8
<i>Insula/HG</i>	73	R	2,61 (0,005)	38	-30	20
<i>Frontale Gebiete</i>						
<hr/>						
<i>IFG</i>	47	L	2,61 (0,005)	-30	38	-10
<b><i>visuelle Maske</i></b>						
<hr/>						
<i>Okipitale Gebiete</i>						
<hr/>						
<i>MOG/IOG/FG</i>	1006	R	3,41 (0,000)	30	-44	-16
<i>IOG</i>	275	L	4,0 (0,000)	-46	-68	14
<i>Frontale Gebiete</i>						
<hr/>						
<i>Gyrus praecentralis</i>	66	R	2,62 (0,005)	48	-8	46
<i>Gyrus praecentralis</i>	72	L	3,06 (0,001)	-38	-14	44
<i>Lobulus paracentralis</i>	41	R	2,78 (0,003)	18	-34	52

*(h) alle Stimuli: asynchrone > synchrone Perzeption (p < 0,05; k > 30)**bimodale Maske**Temporale Gebiete**pSTS*

36	L	2,54 (0,006)	-62	-52	30
----	---	-----------------	-----	-----	----

*Frontale Gebiete**MFG*

161	R	2,92 (0,002)	48	26	28
-----	---	-----------------	----	----	----

*MFG*

52	L	2,62 (0,005)	-50	22	28
----	---	-----------------	-----	----	----

*IFG*

450	L	3,71 (0,000)	-44	42	-10
-----	---	-----------------	-----	----	-----

*Insula/Operculum frontale/IFG/MFG*

415	R	4,62 (0,000)	42	20	-4
-----	---	-----------------	----	----	----

*Insula/Operculum frontale/IFG*

572	L	4,77 (0,000)	-34	20	0
-----	---	-----------------	-----	----	---

*SFG/ACC*

395	R/L	3,84 (0,000)	0	18	48
-----	-----	-----------------	---	----	----

*Parietale Gebiete**IPS/Gyrus  
supramarginalis/angularis*

354	L	3,35 (0,001)	-56	-36	50
-----	---	-----------------	-----	-----	----

*Operculum parietale*

197	L	3,11 (0,003)	-46	-22	20
-----	---	-----------------	-----	-----	----

*IPS/Gyrus angularis*

37	L	2,05 (0,02)	40	-50	44
----	---	----------------	----	-----	----

*Cerebellum**Cerebellum*

142	R	3,94 (0,000)	22	-50	-26
-----	---	-----------------	----	-----	-----

*Cerebellum*

42	L	2,0 (0,02)	-42	-58	-30
----	---	---------------	-----	-----	-----

<i>auditorische Maske</i>						
<i>Frontale Gebiete</i>						
<i>IFG</i>	52	L	3,3 (0,001)	-42	42	-10
<i>visuelle Maske</i>						
<i>Parietale Gebiete</i>						
<i>IPS</i>	43	L	2,29 (0,01)	-40	-46	42

**Tabelle 10.** Lokale Maxima der experimentellen Kontraste von Experiment 1. Die Kontraste stellen das Antwortverhalten gegenüber asynchronen und synchronen Sequenzen dar. Die Häufigkeit asynchroner Antworten ist gegenüber asynchronen Stimuli hoch und nimmt mit dem SOA der Reize zu, die Häufigkeit synchroner Antworten nimmt bei großen SOAs ab und steigt bei phasengleichen Reizen. Diese Variation wird durch die hier aufgeführten Kontraste dargestellt (s. Haupttext).

<i>Anatomische Struktur</i>	<i>Hemi- sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<b>MNI Koordinaten</b> <i>x y z</i>		
<b>(a) Antwortverhalten synchron (<math>p &lt; 0,05</math>; <math>k &gt; 30</math>)</b>						
<i>bimodale Maske</i>						
<i>Temporale Gebiete</i>						
<i>pSTS/STG</i>	R	1930	3,52 (0,000)	54	-40	24
<i>pSTS</i>	R	20	3,0 (0,002)	50	-36	6
<i>pSTS/MTG</i>	L	161	3,72 (0,000)	-60	-56	6
<i>MTG/pSTS</i>	L	118	2,51 (0,007)	-48	-70	12
<i>mSTS</i>	L	36	3,65 (0,000)	-46	-24	-8
<i>pSTS/SMG</i>	L	182	2,98 (0,002)	-62	-42	20

<i>MTG</i>	R	30	3,37 (0,000)	56	-56	2
<i>Planum polare/Insula</i>	R	37	3,23 (0,001)	42	-10	-18
<i>Planum polare/Gyrus separans /STG</i>	R	36	3,07 (0,001)	50	16	-8
<i>Insula/Gyrus separans /Operculum frontale</i>	L	54	3,23 (0,001)	-48	6	2
<i>PT/STG</i>	R	23	2,9 (0,002)	68	-34	22
<i>Frontale Gebiete</i>						
<i>MFG/IFG</i>	R	2677	3,73 (0,000)	58	20	10
<i>MFG</i>	R	78	2,86 (0,002)	36	54	14
<i>MFG</i>	L	4071	4,05 (0,000)	-52	28	20
<i>Gyrus praecentralis/MFG</i>	L	751	3,76 (0,000)	-44	4	46
<i>IFG/MFG/Gyrus frontomarginals</i>	L	351	3,02 (0,002)	-42	50	-14
<i>SFG/ACC</i>	L	229	3,95 (0,000)	-2	14	48
<i>Parietale Gebiete</i>						
<i>Lobulus paracentralis</i>	R	160	2,74 (0,004)	0	-24	60
<i>SMG</i>	L	70	3,3 (0,001)	-44	-44	32
<i>IPS</i>	L	334	3,29 (0,001)	-54	-48	50
<i>Okzipitale Gebiete</i>						
<i>LG/IOG/Area striata</i>	R	76	3,33 (0,001)	6	-78	-8
<i>FG</i>	R	20	3,32 (0,001)	38	-60	-20
<i>FG/Cerebellum</i>	L	557	4,14 (0,000)	-40	-60	-20

<i>Cerebellum</i>						
<i>Cerebellum</i>	R	1680	3,62 (0,000)	30	-60	-34
<i>Cerebellum</i>	R	61	3,27 (0,001)	24	-74	-28
<i>auditorische Maske</i>						
<i>Temporale Gebiete</i>						
<i>aSTS/STG</i>	R	70	2,5 (0,007)	64	-4	-4
<i>mSTS/HG/Insula</i>	L	347	3,65 (0,000)	-46	-24	-8
<i>Frontale Gebiete</i>						
<i>MFG/IFG</i>	L	114	2,69 (0,004)	-32	46	-2
<i>Gyrus praecentralis</i>	L	35	2,42 (0,009)	-34	0	36
<i>visuelle Maske</i>						
<i>Okzipitale Gebiete</i>						
<i>FG/MTG</i>	R	738	3,37 (0,000)	36	-60	-18
<i>ITG</i>	L	179	2,52 (0,006)	-46	-72	12
<i>Parietale Gebiete</i>						
<i>IPS</i>	L	85	2,07 (0,02)	-34	-46	38
<i>Frontale Gebiete</i>						
<i>Gyrus praecentralis/MFG</i>	L	287	3,6 (0,000)	-42	0	48

<i>Cerebellum</i>						
<i>Cerebellum</i>	L	114	3,47 (0,000)	-22	-64	-34
<b>(b) Antwortverhalten asynchron (<math>p &lt; 0,05</math>; <math>k &gt; 30</math>)</b>						
<i>bimodale Maske</i>						
<i>Temporale Gebiete</i>						
<i>mSTS/STG</i>	L	44	2,6 (0,005)	-42	-38	2
<i>Frontale Gebiete</i>						
<i>Gyrus Post-, praecentralis</i>	L	41	3,88 (0,000)	-20	-26	58

**Tabelle 11.** Lokale Maxima der experimentellen Kontraste von Experiment 2. Neuronale Aktivität zu Beginn der Expositionsphase wurde mit solcher zu Ende der Exposition verglichen. Des Weiteren wurden die Kontraste innerhalb derjenigen Areale untersucht, die in Experiment durch synchrone bzw. asynchrone Perzeption moduliert wurden.

<i>Anatomische Struktur</i>	<i>Hemi-sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<b>MNI Koordinaten x y z</b>		
<b>(a) asynchrone Exposition (-125 ms/125 ms): 3. &gt; 1. Minute exklusive synchrone Exposition: 3. &gt; 1. Minute (<math>p &lt; 0,05</math>; <math>k &gt; 20</math>)</b>						
<i>Temporale Gebiete</i>						
<i>ITG/MTG</i>	R	507	3,69 (0,000)	58	-60	-6
<i>aSTS/MTG</i>	R	84	2,78 (0,003)	52	-4	14
<i>MTG</i>	R	26	2,13 (0,02)	58	-62	8
<i>aSTS/STG/MTG</i>	L	358	2,96 (0,002)	-52	-2	-16
<i>MTG/temporo-okzipitale Verbindung/pSTS/Operculum parietale</i>	L	1742	3,59 (0,000)	-54	-76	8

<i>Perirhinaler Kortex/Gyrus parahippocampalis</i>	R	28	2,22 (0,01)	14	-30	-8
<i>Perirhinaler Kortex/Gyrus parahippocampalis</i>	L	64	2,51 (0,006)	-22	-28	-6
<i>Operculum frontale</i>	R	198	2,83 (0,002)	56	2	10
<i>Frontale Gebiete</i>						
<i>IFG</i>	R	210	2,79 (0,003)	60	26	6
<i>IFG</i>	L	50	2,71 (0,004)	-54	30	4
<i>SFG</i>	R	66	2,42 (0,008)	24	42	30
<i>MFG</i>	R	49	2,59 (0,005)	22	68	10
<i>MFG</i>	R	22	2,32 (0,01)	50	48	8
<i>SFG/MFG</i>	L	103	2,39 (0,009)	-18	42	34
<i>Insula</i>	R	144	2,47 (0,007)	44	18	-16
<i>Insula</i>	L	72	2,61 (0,005)	-36	-4	-2
<i>Insula</i>	L	32	2,32 (0,01)	-24	14	-10
<i>Insula/IFG</i>	L	93	2,1 (0,02)	-46	26	-10
<i>Gyrus rostralis superior/inferior/ACC</i>	R	322	3,2 (0,001)	14	52	-6
<i>ACC</i>	L	42	2,22 (0,01)	-10	46	-4
<i>Gyrus praecentralis</i>	R	46	1,98 (0,02)	54	10	30
<i>Gyrus praecentralis</i>	L	21	2,1 (0,02)	-38	-6	40
<i>Parietale Gebiete</i>						
<i>SMG/PT/ Operculum parietale</i>	R	358	2,96 (0,002)	64	-28	36
<i>IPS/Lobulus parietalis superior</i>	L	4305	4,32 (0,000)	-20	-52	62
<i>Gyrus postcentralis</i>	R	25	2,17 (0,02)	36	-22	40

<i>Operculum parietale</i>	R	24	1,96 (0,03)	56	-24	20
<i>Okzipitale Gebiete</i>						
<hr/>						
<i>Temporo-okzipitale Verbindung/SOG</i>	L/R	811	4,2 (0,000)	44	-78	26
<i>Area striata/SOG</i>	R	308	3,8 (0,000)	12	-88	10
<i>SOG</i>	R	35	2,57 (0,005)	20	-64	16
<i>IOG</i>	R	48	3,66 (0,000)	14	-92	26
<i>IOG</i>	L	52	2,63 (0,005)	-40	-72	10
<i>IOG/Area striata</i>	L	183	3,32 (0,001)	-22	-54	2
<i>MOG</i>	R	304	3,44 (0,000)	22	-86	-22
<i>MOG</i>	L	302	3,27 (0,001)	-20	-92	16
<i>MOG</i>	L	41	2,81 (0,003)	-40	-80	-16
<i>LG</i>	R	46	3,18 (0,001)	16	-72	-14
<i>LG</i>	L	40	3,24 (0,001)	-8	-68	-8
<i>IOG/LG/Area striata</i>	L	27	2,57 (0,005)	-12	-56	6
<i>Area Striata</i>	L	34	2,39 (0,009)	-14	-74	-20
<i>Cerebellum</i>						
<hr/>						
<i>Cerebellum</i>	R	93	3,18 (0,001)	26	-44	-24
<i>Cerebellum</i>	R	22	2,66 (0,004)	2	-78	-36
<i>Cerebellum</i>	L	95	2,64 (0,004)	-38	-78	-34
<i>Subkortikale Strukturen</i>						
<hr/>						
<i>Medial dorsaler Kern des Thalamus</i>	L	33	2,88 (0,002)	-14	-22	-2

(b) *asynchrone Exposition (-125 ms/125 ms): 1. > 3. Minute exklusive synchrone Exposition:  
1. > 3. Minute* ( $p < 0,05; k > 20$ )

<i>Temporale Gebiete</i>						
STG/PT	L	166	2,39 (0,009)	-60	-30	6
Gyrus parahippocampalis	L	30	2,31 (0,01)	-36	-32	2
<i>Frontale Gebiete</i>						
MFG	L	184	3,1 (0,001)	-46	34	28
MFG/Gyrus praecentralis	L	81	3,1 (0,001)	-60	10	32
MFG	L	73	2,45 (0,007)	-36	20	54
IFG/MFG	L	158	2,58 (0,005)	-44	50	2
MFG	L	35	2,28 (0,01)	-34	0	64
MFG	R	33	2,02 (0,02)	46	34	36
SFG	L	20	2,45 (0,007)	-6	34	58
Gyrus praecentralis	R	70	2,75 (0,003)	58	-10	48
<i>Parietale Gebiete</i>						
IPS/Gyrus angularis/supramarginalis	L	488	2,9 (0,002)	-52	-48	52
<i>Cerebellum</i>						
Cerebellum	L	26	2,17 (0,02)	-16	-56	-20
Cerebellum	L	30	1,97 (0,02)	-26	-50	-28
Cerebellum	R	24	2,0 (0,02)	40	-66	-30

(c) *asynchrone Exposition (-125 ms/125 ms): 3. > 1. Minute exklusive synchrone Exposition:  
3. > 1. Minute innerhalb perzeptueller Synchronizität aus Experiment 1 ( $p < 0,05$ ;  $k > 20$ )*

*bimodale Maske*

*Temporale Gebiete*

<i>aSTS/MTG/STG</i>	R	45	2,35 (0,01)	52	-8	-14
<i>pSTS/STG/MTG</i>	L	243	3,27 (0,001)	-54	-74	8
<i>aSTS/mSTS/STG/MTG</i>	L	186	2,96 (0,002)	-52	-2	-16
MTG	R	97	2,69 (0,004)	54	-58	-2
MTG	L	23	2,15 (0,02)	-52	-56	-4
<i>Perirhinaler Kortex/Gyrus parahippocampalis</i>	R	20	2,22 (0,01)	14	-30	-8

*Frontale Gebiete*

<i>Lobulus paracentralis</i>	L	34	3,25 (0,001)	-12	-34	56
IFG/MFG	R	31	2,02 (0,02)	46	26	14

*auditorische Maske*

<i>aSTS</i>	R	15	2,78 (0,003)	52	-4	-14
-------------	---	----	-----------------	----	----	-----

*visuelle Maske*

MTG/ITG	R	40	2,63 (0,005)	52	-60	-2
---------	---	----	-----------------	----	-----	----

(d) *asynchrone Exposition (-125 ms/125 ms): 1. > 3. Minute exklusive synchrone Exposition: 1. > 3. Minute innerhalb perzeptueller Asynchronizität aus Experiment 1 ( $p < 0,05$ ;  $k > 20$ )*

*bimodale Maske*

<i>Frontale Gebiete</i>						
MFG	L	31	2,45 (0,007)	-42	44	18
MFG	L	87	2,37 (0,009)	-42	48	4
<i>Parietale Gebiete</i>						
IPS/Gyrus angularis/supramarginalis	L	161	2,9 (0,002)	-52	-48	52

**Tabelle 12.** Aufgeführt sind die lokalen Maxima derjenigen Areale, die für Experiment 3 in visuellen und auditorischen Lokalisationsbedingungen mehr Aktivität zeigten als in der Ruhebedingung. Ausmaß der Aktivierungen (T-Werte) in der bimodalen Lokalisationsbedingung sind nicht aufgeführt, da sie durch das gewählte Maskierverfahren (auditorisch inklusive visuell) ihre Aussagekraft verlieren.

<i>Anatomische Struktur</i>	<i>Hemi- sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<i>Koordinaten/MNI x y z</i>		
<i>(a) auditorisch &gt; baseline (exkl. 'auditorisch &gt; baseline &amp; visuell &gt; baseline')</i>						
<i>(<math>p &lt; 0,005</math>; <math>k &gt; 200</math>)</i>						
<i>Temporale Gebiete</i>						
STG/MTG/STS/PT	R	4604	21,22 (0,000)	44	-24	6
STG/MTG/STS/PT	L	5296	19,28 (0,000)	-40	-28	8
<i>Frontale Gebiete</i>						
IFG	L	337	5,18 (0,001)	-58	26	0

<i>Subkortikale Strukturen</i>						
<i>Corpus geniculatum mediale</i>	L	209	14,64 (0,000)	-18	-28	-6
<b>(b) visuell &gt; baseline (exkl. 'auditorisch &gt; baseline &amp; visuell &gt; baseline') (<math>p &lt; 0,005</math>; <math>k &gt; 100</math>)</b>						
<i>Okzipitale Gebiete</i>						
FG	L	1674	10,58 (0,000)	-38	-66	-16
FG/ITG	R	2401	10,16 (0,000)	40	-66	-14
MOG/Sulcus occipitalis transversus	R	279	5,53 (0,001)	34	-76	8
MOG	L	303	6,62 (0,000)	-28	-82	20
<i>Frontale Gebiete</i>						
Gyrus praecentralis/MFG/IFG	R	288	6,26 (0,000)	38	6	34
<i>Parietale Gebiete</i>						
MFG	R	152	5,25 (0,001)	46	6	50
Lobus parietalis superior	R	113	4,15 (0,002)	28	-54	46
<b>(c) auditorisch &gt; baseline &amp; visuell &gt; baseline (potentiell bimodal) (<math>p &lt; 0,05</math>; <math>k &gt; 100</math>)</b>						
<i>Temporale Gebiete</i>						
STS/STG/MTG	L	1708				
STS/STG/MTG	R	2396				
<i>Frontale Gebiete</i>						
MFG/Gyrus praecentralis	L	354				
MFG/Gyrus praecentralis/Insula	R	1946				

*Cerebellum**Cerebellum*

L

373

**Tabelle 13.** Lokale Maxima von Experiment 3 für physikalische Synchronizität > physikalische Asynchronizität, physikalische Asynchronizität (Ton voran) > physikalische Synchronizität und physikalische Asynchronizität (Bild voran) > physikalische Synchronizität innerhalb bimodaler Gebiete.

<i>Anatomische Struktur</i>	<i>Hemi-sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<i>MNI Koordinaten x y z</i>		
<b>(a) physikalische Asynchronizität (Ton zuerst) &gt; physikalische Synchronizität</b> ( $p < 0,05$ ; $k > 5$ )						
<b><i>bimodale Maske</i></b>						
<i>Temporale Gebiete</i>						
<i>pSTS</i>	L	57	2,11 (0,009)	-54	-54	8
<i>pSTS</i>	R	122	3,22 (0,002)	54	-46	-2
<i>pSTS</i>	L	12	3,4 (0,001)	-68	-38	16
<i>aSTS</i>	R	177	3,19 (0,002)	62	-14	-8
<i>Frontale Gebiete</i>						
<i>Gyrus praecentralis</i>	R	17	3,05 (0,003)	44	0	40
<i>Gyrus praecentralis</i>	R	7	1,89 (0,02)	48	6	44
<i>IFG/MFG</i>	L	19	1,93 (0,02)	-46	22	24
<i>anteriore Insula/IFG</i>	L	6	3,19 (0,002)	-36	40	-18

**(b) physikalische Synchronizität > physikalische Asynchronizität ( $p < 0,05$ ;  $k > 5$ )****bimodale Maske***Temporale Gebiete*

<i>pSTS</i>	R	8	1,73 (0,04)	62	-50	10
<i>pSTS</i>	R	9	1,79 (0,03)	54	-46	14

*Frontale Gebiete*

<i>IFG/MFG</i>	R	14	2,24 (0,008)	50	36	12
----------------	---	----	-----------------	----	----	----

**(c) physikalische Asynchronizität (Bild zuerst) > physikalische Synchronizität ( $p < 0,05$ ;  $k > 5$ )****bimodale Maske***Temporale Gebiete*

<i>pSTS/mSTS</i>	L	9	2,79 (0,005)	-68	-38	14
<i>pSTS/mSTS</i>	L	14	2,49 (0,007)	-60	-30	8
<i>aSTS</i>	R	370	3,21 (0,001)	64	-20	-12

*Frontale Gebiete*

<i>IFG/MFG</i>	R	41	1,99 (0,01)	38	18	26
<i>IFG/MFG</i>	L	11	2,0 (0,01)	-46	20	28
<i>anteriore Insula/IFG</i>	L	41	3,35 (0,001)	-36	38	-16
<i>anteriore Insula/IFG</i>	R	8	2,59 (0,006)	50	42	2

**Tabelle 14.** Lokale Maxima von Experiment 3 für synchrone Wahrnehmung verglichen mit asynchroner Wahrnehmung in beide Richtungen (Ton voran und Bild voran) innerhalb bimodaler Gebiete über alle Stimulusbedingungen zusammengefasst.

<i>Anatomische Struktur</i>	<i>Hemi-sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<i>MNI Koordinaten x y z</i>		
<b>(a) synchrones Perzept &gt; asynchrones Perzept (<math>p &lt; 0,05</math>; <math>k &gt; 5</math>)</b>						
<b><i>bimodale Maske</i></b>						
<b><i>Temporale Gebiete</i></b>						
<i>pSTS/mSTS</i>	R	32	2,01 (0,01)	48	-38	8
<i>pSTS/mSTS</i>	L	13	2,39 (0,007)	-58	-34	-4
<i>aSTS</i>	R	96	3,31 (0,001)	60	-22	-2
<i>aSTS</i>	R	22	2,73 (0,005)	62	-10	-10
<b>(b) asynchrones Perzept &gt; synchrones Perzept (<math>p &lt; 0,05</math>; <math>k &gt; 5</math>)</b>						
<b><i>bimodale Maske</i></b>						
<b><i>Temporale Gebiete</i></b>						
<i>pSTS</i>	L	17	2,61 (0,006)	-66	-50	2
<i>pSTS</i>	R	10	2,47 (0,007)	58	-40	-8
<i>pSTS</i>	R	14	2,21 (0,008)	54	-44	18
<i>pSTS/mSTS</i>	R	206	3,29 (0,001)	66	-34	-6
<b><i>Frontale Gebiete</i></b>						
<i>anteriore Insula</i>	R	14	1,98 (0,01)	42	36	-10
<i>IFG/MFG</i>	R	644	3,33 (0,001)	56	24	22

<i>IFG/MFG</i>	L	9	1,96 (0,02)	-54	30	14
----------------	---	---	----------------	-----	----	----

**Tabelle 15.** Perzeptuelle Kontraste von Experiment 3: perzeptuelle Asynchronizität (Ton voran) > synchrone Perzeption, wenn Stimuli gezeigt wurden, bei denen der Ton vorangeht, perzeptuelle Synchronizität > asynchrone Perzepte, wenn synchrone Stimuli gezeigt wurden und perzeptuelle Asynchronizität (Bild voran) > synchrone Perzeption wenn Stimuli gezeigt wurden, bei denen das Bild vorangeht innerhalb bimodaler Gebiete.

<i>Anatomische Struktur</i>	<i>Hemi- sphäre</i>	<i>Clustergröße (Voxel)</i>	<i>t-Wert (p-Wert)</i>	<b>MNI Koordinaten</b> <b>x y z</b>		
-----------------------------	-------------------------	---------------------------------	----------------------------	--	--	--

**(a) kohärentes Perzept (Ton voran) > inkohärente perzeptuelle Synchronizität ( $p < 0,05$ ;  $k > 5$ )**

***bimodale Maske***

*Temporale Gebiete*

<i>pSTS</i>	R	15	2,48 (0,007)	64	-42	12
<i>pSTS</i>	L	57	2,01 (0,01)	-52	-42	4
<i>mSTS/pSTS</i>	L	18	1,99 (0,01)	-64	-38	10
<i>mSTS</i>	L	6	2,06 (0,01)	-64	-36	-8
<i>mSTS</i>	R	8	2,7 (0,005)	54	-30	-14

*Frontale Gebiete*

<i>Gyrus praecentralis</i>	L	12	2,05 (0,01)	-50	6	44
<i>Gyrus praecentralis</i>	L	11	2,59 (0,006)	-40	8	38
<i>Gyrus praecentralis</i>	R	86	3,17 (0,002)	48	10	40
<i>IFG/MFG</i>	L	28	3,99 (0,000)	-36	8	60
<i>IFG/MFG</i>	L	121	4,01 (0,000)	-40	20	24
<i>IFG/MFG</i>	R	191	3,1 (0,003)	54	28	12
<i>anteriore Insula</i>	L	20	3,29 (0,001)	-32	28	-6

(b) *kohärente perzeptuelle Synchronizität > inkohärentes asynchrones Perzept*  
( $p < 0,05$ ;  $k > 5$ )

*bimodale Maske*

*Temporale Gebiete*

<i>pSTS</i>	R	5	2,01 (0,01)	48	-52	10
<i>pSTS</i>	L	29	2,42 (0,007)	-50	-46	8
<i>mSTS</i>	L	62	3,14 (0,002)	-62	-24	4
<i>aSTS</i>	R	447	3,37 (0,001)	62	-16	-4

*Frontale Gebiete*

<i>anteriore Insula</i>	L	14	2,04 (0,01)	-34	40	-14
-------------------------	---	----	----------------	-----	----	-----

(c) *kohärentes Perzept (Bild voran) > inkohärente perzeptuelle Synchronizität*  
( $p < 0,05$ ;  $k > 5$ )

*bimodale Maske*

*Temporale Gebiete*

<i>pSTS</i>	R	18	3,21 (0,002)	64	-50	2
<i>pSTS</i>	R	5	2,01 (0,01)	62	-50	14
<i>pSTS</i>	L	29	3,17 (0,002)	-58	-50	8

*Frontale Gebiete*

<i>Gyrus praecentralis</i>	R	21	3,2 (0,002)	42	8	46
<i>Gyrus praecentralis</i>	L	134	3,2 (0,002)	-42	16	26

IFG/MFG	R	451	3,31 (0,001)	50	24	24
anteriore Insula/IFG	L	20	2,37 (0,008)	-34	30	-2
anteriore Insula/IFG	R	72	4,1 (0,000)	44	40	-10

**Tabelle 16.** Lokale Maxima der interregionalen Konnektivitätsanalyse von Experiment 3 für die Vergleiche ‚kohärente Perzepte bei > asynchrone Perzepte synchroner Stimulation‘, ‚kohärente Perzepte > synchrone Perzepte bei asynchroner Stimulation (Ton voran)‘ und ‚kohärente Perzepte > synchrone Perzepte bei asynchroner Stimulation (Bild voran)‘.

Anatomische Struktur	Hemi- sphäre	Clustergröße (Voxel)	t-Wert (p-Wert)	MNI Koordinaten x y z		
----------------------	-----------------	-------------------------	--------------------	--------------------------	--	--

**(a) psychophysiologische Interaktion asynchroner Perzepte (Ton voran)**  
( $p < 0,01$ ;  $k > 20$ )

<i>Temporale Gebiete</i>						
aSTS	R	73	3,57 (0,001)	64	-12	-8
pSTS	L	22	3,32 (0,001)	-50	-46	16
<i>Frontale Gebiete</i>						
MFG/IFG	R	168	2,94 (0,002)	40	22	16
Gyrus praecentralis	L	54	2,68 (0,005)	-48	-2	46

**(b) psychophysiologische Interaktion synchroner Perzepte** ( $p < 0,05$ ;  $k > 5$ )

<i>Temporale Gebiete</i>						
pSTS	L	31	3,23 (0,001)	-54	-54	12

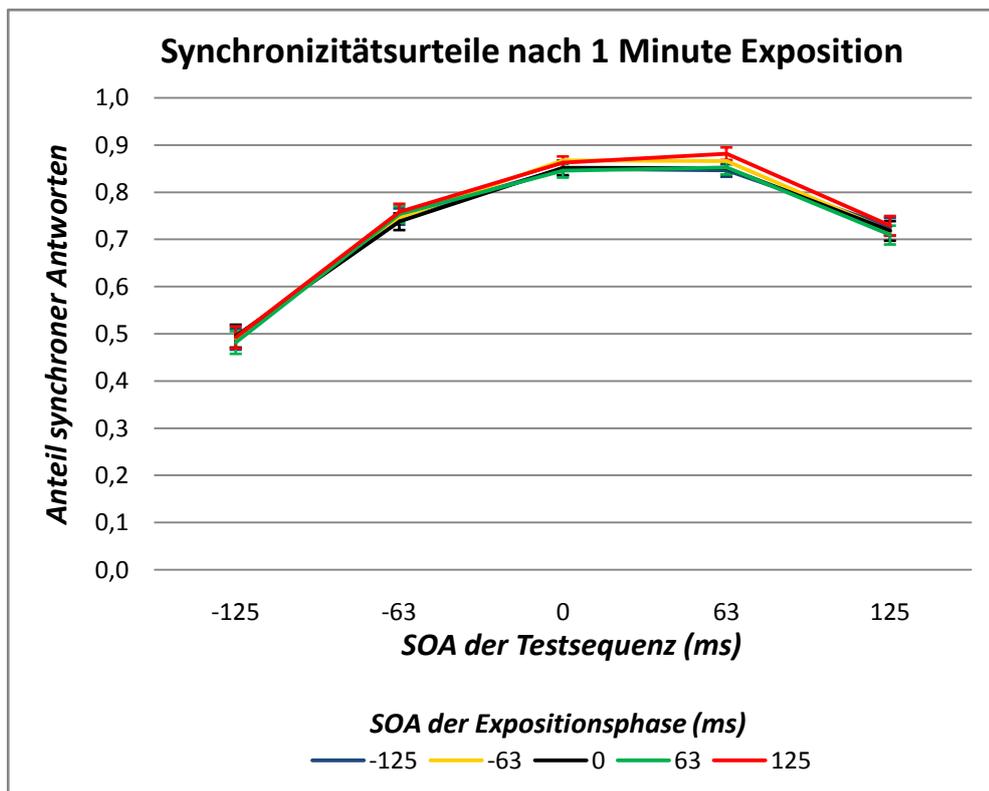
(c) *psychophysiologische Interaktion asynchroner Perzepte (Bild voran)*  
 ( $p < 0,05$ ;  $k > 5$ )

*Temporale Gebiete*

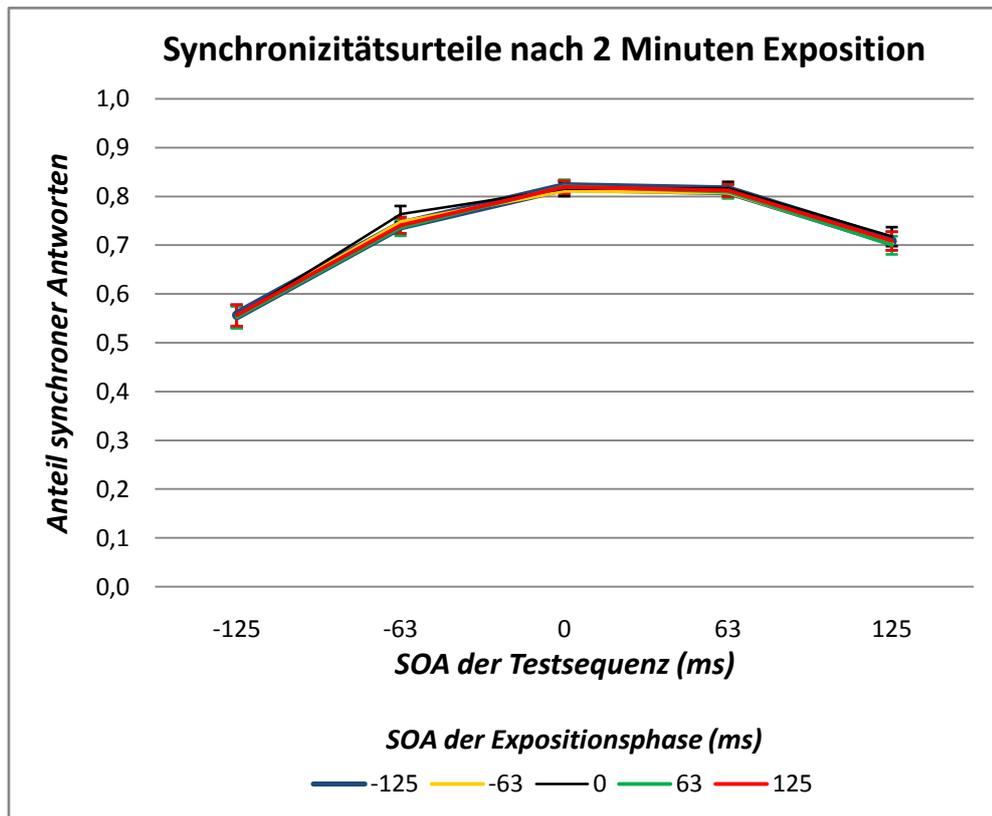
<i>pSTS/STG</i>	R	44	2,42 (0,009)	52	-46	0
-----------------	---	----	-----------------	----	-----	---

*Frontale Gebiete*

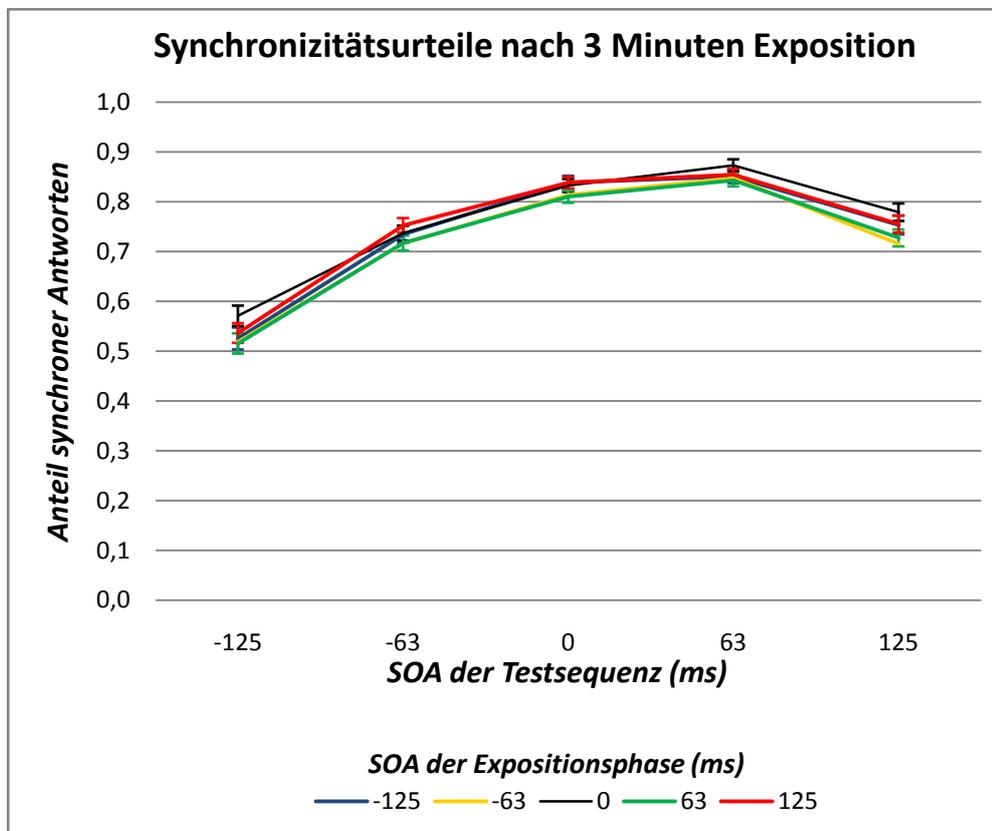
<i>IFG</i>	R	21	2,51 (0,008)	58	22	14
------------	---	----	-----------------	----	----	----



**Abbildung 43.** Anteil synchroner Antworten von Vorexperiment 2.1 nach einminütiger Exposition in Abhängigkeit der SOAs der Expositionsphase. Fehlerbalken stellen Standardfehler dar.



**Abbildung 44.** Anteil synchroner Antworten von Vorexperiment 2.1 nach zweiminütiger Exposition in Abhängigkeit der SOAs der Expositionsphase. Fehlerbalken stellen Standardfehler dar.



**Abbildung 45.** Anteil synchroner Antworten von Vorexperiment 2.1 nach dreiminütiger Exposition in Abhängigkeit der SOAs der Expositionsphase. Fehlerbalken stellen Standardfehler dar.

# LEBENS LAUF

## *Persönliche Angaben*

---

<i>Name</i>	Daniel Bergmann
<i>Geburtsort</i>	Leipzig
<i>Familienstand</i>	ledig
<i>Geburtsdatum</i>	22. Februar 1981
<i>Staatsangehörigkeit</i>	deutsch

## *Ausbildung*

---

<i>Februar –April 2007</i>	Auslandsaufenthalt am Department of Experimental Psychology der University of Oxford
<i>April 2006 – Juli 2009</i>	Wissenschaftlicher Mitarbeiter an der Neurologischen Universitätsklinik der Otto-von-Guericke-Universität Magdeburg  Promotionsvorhaben an der Fakultät für Naturwissenschaften der Otto-von-Guericke-Universität
<i>März 2006</i>	Diplom (Psychologie), Note: sehr gut
<i>Januar 2003 – März 2006</i>	Wissenschaftliche Hilfskraft an der Klinik für Neurologie II, Otto-von-Guericke-Universität, Magdeburg  Tutor am Lehrstuhl für Methodenlehre, Institut für Psychologie I, Otto-von-Guericke-Universität, Magdeburg  Wissenschaftliche Hilfskraft am Lehrstuhl für Methodenlehre, Institut für Psychologie I, Otto-von-Guericke-Universität, Magdeburg  Wissenschaftliche Hilfskraft am Institut für Sportwissenschaften, Otto-von-Guericke-Universität, Magdeburg
<i>September 2003</i>	Vordiplom Psychologie
<i>Oktober 2001 - März 2006</i>	Ordentlicher Student der Psychologie an der Otto-von-Guericke-Universität, Magdeburg, Abschluss: Diplom

*August 2000 – September 2001*    Anderer Dienst im Ausland in Marseille, Frankreich

*August 1991 – Juli 2000*        Georg-Büchner-Gymnasium    Rheinfelden    (Baden),  
Abschluss: Abitur

Magdeburg, den 28. Januar 2009

---

Dipl.-Psych. Daniel Bergmann

Daniel Bergmann  
Lessingstraße 28  
39108 Magdeburg

## Erklärung

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation zu dem Thema:

*„Neuronale Mechanismen zeitabhängiger Prozesse bei audiovisueller Verarbeitung“*

selbständig verfasst, nicht bereits als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, den 28.01.2010

---

Dipl.-Psych. Daniel Bergmann