

Beeinflussung von auditorischer Objektbildung durch visuelle Stimulation

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.-Phys. Torsten Rahne**
geb. am 2. August 1979 in Lutherstadt Eisleben

Gutachter: Prof. Dr. Hellmut von Specht
Prof. Dr. Jürgen Kießling

eingereicht am: 24. September 2007

verteidigt am: 25. März 2008

Danksagung

Diese Arbeit ist von vielen Personen durch ihre Ratschläge, Impulse und Hilfe sehr gefördert worden. All ihnen danke ich gerne!

Zunächst danke ich Herrn Prof. Hellmut von Specht, der als Leiter der Abteilung für Experimentelle Audiologie und Medizinische Physik die Rahmenbedingungen zum Durchführen von Experimenten, vielen Konferenzbesuchen und dem Erstellen der Arbeit auch über seine Emeritierung hinaus bereitgestellt hat. Damit unmittelbar verbunden ist der Dank an die Kollegen Martin Böckmann-Barthel, Roland Mühler, Sabine Haumann, Andreas Tiefenau, Thomas Wesarg, die stets kompetente Partner für anregende Diskussionen waren und mir oft den Rücken für die Forschung freigehalten haben. Auch Jacqueline Habisch, Katrin Mentzel, Bettina Schröder und Helga Wiemann – den MTA der Abteilung – sowie Marion Janda und Michael Ziese danke ich herzlich für die technische Unterstützung bei den EEG-Messungen und dem Erstellen von Audiogrammen.

Prof. Stefan C. Müller und Prof. Marcus Hauser als den Leitern der Abteilung Biophysik danke ich für die Möglichkeit, an der Fakultät für Naturwissenschaften zu promovieren sowie für vielfältige Ratschläge zum Inhalt der Arbeit. Zudem danke ich Reinhard König, Norman Zacharias und Nicole Behne als Mitarbeiter des Speziallabors „Nicht-Invasive Bildgebung“ des Leibniz-Instituts für Neurobiologie in Magdeburg. Durch die Bereitstellung des MEG-Labors und mit ihrer methodischen Unterstützung konnte ich die MEG-Studie erfolgreich durchführen und auswerten.

Ein besonderer Dank gilt Frau Prof. Elyse Sussman und ihrem Team des Department of Neuroscience des Albert Einstein College of Medicine in New York. Während meines Forschungsaufenthaltes in ihrem Labor hat diese Arbeit wertvolle Impulse erhalten. Vielen Dank für die schöne Zeit! Auch danke ich den vielen Probanden – in New York und Magdeburg – die mit ihrer Geduld die Messungen ermöglicht haben.

Schließlich danke ich meiner Frau Ruth, die von Anfang an durch ihre Motivierung und manchen Verzicht zum Gelingen der Arbeit beigetragen hat.

Magdeburg, im September 2007



Inhaltsverzeichnis

1	Einleitung	1
2	Auditorische Szenenanalyse	5
2.1	Einführung in Objekte, Streaming und Grouping	5
2.2	Streaming	10
2.2.1	Ambiguität	10
2.2.2	Mechanismen von Streaming	12
2.2.3	Modelle für Streaming	18
2.3	Multimodale Reizverarbeitung	21
3	Methodik zur Messung von Streaming	27
3.1	Psychometrie	28
3.2	Messung von Hirnaktivität	29
3.2.1	Elektroenzephalographie	29
3.2.2	Magnetenzephalographie	31
3.3	Evozierte Potentiale und Felder	35
3.4	Mismatch-Negativity	38
3.4.1	Grundlegendes	38
3.4.2	Neuronale Generatoren der MMN	40

3.4.3	Anwendungsmöglichkeiten – Messparadigmen	42
3.5	Paradigmen mit MMN bei Streaming	44
3.6	Visuelle Mismatch-Negativity	46
3.7	Problemstellung	48
4	Versuchsdurchführung und Auswertung	51
4.1	Stimulationsreize und Präsentation	52
4.2	EEG-Studie im ambigen Bereich	54
4.2.1	Probanden	54
4.2.2	Stimuli	54
4.2.3	Bedingungen	56
4.2.4	Messablauf	58
4.2.5	EEG-Messungen	59
4.2.6	Analyse der EEG-Daten	60
4.3	Psychoakustische Studie	63
4.3.1	Probanden	63
4.3.2	Stimuli	63
4.3.3	Bedingungen	63
4.3.4	Messablauf	65
4.3.5	Datenverarbeitung und -analyse	66
4.4	MEG-Studie im ambigen Bereich	67
4.4.1	Probanden	67
4.4.2	Stimuli	69
4.4.3	Bedingungen	69
4.4.4	Messablauf	70
4.4.5	MEG-Messungen	70

4.4.6	Analyse der MEG-Daten	70
4.5	EEG-Studie im nicht-ambigen Bereich	73
4.5.1	Probanden	73
4.5.2	Stimuli	73
4.5.3	Bedingungen	74
4.5.4	Messablauf	75
4.5.5	EEG-Messungen	75
4.5.6	Analyse der EEG-Daten	76
4.6	Analyse der Aufmerksamkeit	76
5	Experimentelle Ergebnisse	79
5.1	EEG-Studie im ambigen Bereich	79
5.2	Psychoakustische Studie	86
5.3	MEG-Studie im ambigen Bereich	90
5.4	EEG-Studie im nicht-ambigen Bereich	96
5.5	Kontrolle der Aufmerksamkeit	102
6	Diskussion der Ergebnisse	105
6.1	Audiovisuelle Stimulation	105
6.1.1	Auflösung der Ambiguität	105
6.1.2	Nicht-ambiger Bereich	110
6.2	Audiovisuelle Interaktionen	113
6.2.1	Verarbeitungsstufe	113
6.2.2	Einfluss der Aufmerksamkeit	114
6.2.3	Auditorische und visuelle MMN	116
7	Zusammenfassung und Ausblick	121

Abbildungsverzeichnis

2.1	<i>ABA</i> -Tonsequenz zur Erzeugung von Streaming	6
2.2	Visualisierung von auditorischem Streaming	7
2.3	Visualisierung von auditorischem Grouping	7
2.4	Van-Noorden-Diagramm	8
2.5	Beispiel für Streaming in der Barockmusik	9
2.6	Rubin'sche Vase als Beispiel für visuelle Mehrdeutigkeit	11
2.7	Prozesse im afferenten auditorischen Pfad	14
2.8	Charakteristiken des auditorischen und visuellen Systems	22
2.9	Modell der MMN-Prozesse für uni- und bimodale Sinnesreize	23
3.1	Funktionsweise einer Gradiometerspule	32
3.2	Frühe, mittellatente und späte auditorisch evozierte Potentiale	36
3.3	Oddball-Paradigma	37
3.4	Modell zum MMN-Auslösemechanismus	40
3.5	Stimulationsparadigma mit Tripel	45
4.1	Versuchsordnung für EEG und Psychometrie	55
4.2	Stimulationsparadigma der EEG- und MEG-Experimente	57
4.3	Internationales 10-20-System zur Elektrodenanordnung	59
4.4	Stimulationsparadigmata der psychoakustischen Messung	64

4.5	Antwortschema der psychoakustischen Messung	66
4.6	Versuchsordnung für MEG	68
4.7	Sensoren der MEG-Messung	72
5.1	Gruppenmittelwert der EKP an Fz, Cz, Pz, Oz und den Mastoiden	80
5.2	Gruppenmittelwert der Differenzpotentiale aller Stimulationsbedingungen	83
5.3	MMN-Amplituden an Fz, Cz und Oz für alle Stimulationsbedingungen	84
5.4	Spannungsverteilung auf der Schädeloberfläche und SCD-Karten	85
5.5	Lage und Orientierung der Generatoren	85
5.6	Absolute Initialzeiten für alle Stimulationsbedingungen	87
5.7	Absolute Zeiten für die Perzepte <i>integriert</i> und <i>segregiert</i>	88
5.8	Gruppenmittelwerte der relativen Zeiten für alle Stimulationen	89
5.9	Gruppenmittelwerte der durch Standard- und Deviantstimuli evozierten Magnetfelder	91
5.10	Gruppenmittelwerte der Differenzkurven	92
5.11	Gruppenmittelwerte der MEG-Daten	93
5.12	Generatoren der MMN	95
5.13	Gruppenmittelwerte der EKP an Fz, Cz, Pz und den Mastoiden	96
5.14	Gruppenmittelwert der Differenzpotentiale aller Stimulationsbedingungen	98
5.15	MMN-Amplituden an den Elektrodenpositionen Fz, Cz und Pz	99
5.16	Spannungsverteilung auf der Schädeloberfläche und SCD-Karten	100
5.17	Lage und Orientierung der Generatoren	100
5.18	Sensitivität für die Bedingungen mit visueller Stimulation	102
5.19	Reaktionszeiten für alle Bedingungen mit visueller Stimulation	103
6.1	Van-Noorden-Diagramm mit Parameterkonfigurationen dieser Arbeit	110

Tabellenverzeichnis

4.1	Auditorische Stimuli	53
4.2	Visuelle Stimuli	53
4.3	Bedingungen des EEG-Experiments im ambigen Bereich	57
4.4	Bedingungen des EEG-Experiments außerhalb des ambigen Bereichs	74
5.1	Gruppenmittelwerte und Standardabweichung für Standard-, Deviant- und Differenzpotential	81
5.2	t -Tests zwischen Standard- und Deviantpotentialen	82
5.3	<i>Post-hoc</i> -Test der Differenzpotentiale	83
5.4	Vergleiche der Antwortzeiten zwischen den Stimulationsbedingungen im psychoakustischen Experiment	94
5.5	Gruppenmittelwerte mit Standardabweichung für Standard-, Deviant- und Differenzpotential	97

Kapitel 1

Einleitung

In vielen alltäglichen Hörsituationen treffen Schallreize von unterschiedlichen Quellen auf das menschliche Ohr. Im Innenohr entstehen daraus elektrische Impulse, die von der Cochlea als neuronale Signale über unterschiedliche Stationen in beide Hirnhälften zum auditorischen Kortex geleitet werden. Durch die *auditorische Szenenanalyse*, einem psychologischen Prozess, werden aus einem Gemisch komplexer Schallreize einzelne Quellen segregiert. Diese Quellen werden im auditorischen System durch verschiedene Streams repräsentiert, innerhalb derer die Schallereignisse gruppiert sind. Einen intuitiven Zugang zur auditorischen Szenenanalyse und zum Streaming findet man in der Vorstellung, den Gesang von Solisten zu verfolgen, die gleichzeitig in ihrer jeweiligen Stimmlage singen. Obwohl sich die jeweiligen Schallwellen überlagern, entstehen klar trennbare Perzepte (Almonte *et al.* 2005).

Im Zusammenhang mit Streaming betrachtet man oft eine auditorische Stimulation mit alternierenden Tonfolgen (*ABAB*), die je nach Frequenzabstand zwischen den Tönen *A* und *B* und dem zeitlichen Abstand aufeinander folgender Töne zu unterschiedlichen Perzepten führen kann. Bei engem Frequenzabstand oder schneller Stimulationsrate wird diese Folge als ein zusammenhängender Stream wahrgenommen. Diese Integration kann nicht mehr bestehen, wenn der Frequenzabstand deutlich vergrößert oder die Stimulationsrate verringert wird. Dann wechselt die Wahrnehmung zu zwei segregierten Streams (*A – A – A . . .* und *B – B – B . . .*). Zwischen diesen beiden Perzepten existiert ein Parameterraum, in dem beide Perzepte, Integration und Segregation, möglich sind.

Das Phänomen des Streaming wurde erstmals von Miller (1947) unter Verwendung einer einfachen auditorischen Situation beschrieben und untersucht. Die Probanden berichte-

ten bei dieser weitgehend rhythmusfreien Präsentation von zwei Tönen unterschiedlicher Frequenz, dass die hohen und tiefen Töne ab einem bestimmten Frequenzabstand zwei separaten Quellen zugeordnet werden können. In der folgenden Zeit wurden weitere Studien auf diesem Gebiet durchgeführt, die einen Höhepunkt in den Arbeiten von van Noorden (1975, 1977) fanden. Diese untersuchten das Phänomen systematisch und klassifizierten Stimulationsparameter, die entweder zur perzeptuellen Integration der Stimuli in einen Stream oder zu deren Segregation in verschiedene Streams führen.

Der ambige Bereich des Streamings, in dem beide Perzepte möglich sind, ist kaum Gegenstand von Studien gewesen, obwohl seit dieser Zeit zahlreiche Arbeiten zu Streaming durchgeführt wurden. Hier stellt sich die Frage, ob und wie das auditorische Perzept, oder zumindest die entsprechende Organisation, im auditorischen Kortex beeinflusst werden kann. Insbesondere der Einfluss anderer Sinnesmodalitäten rückt hier in den Fokus des Interesses, denn in einigen Studien konnte bereits der Einfluss multimodaler Interaktionen auf die Wahrnehmung in einer Modalität gezeigt werden (Duncan *et al.* 1997; Shams *et al.* 2000).

Aus den zahlreichen möglichen Wechselwirkungen greift die vorliegende Arbeit den Einfluss visueller Reize auf die auditorische Wahrnehmung heraus und untersucht explizit, wie die neuronale Organisation von Tönen durch visuelle Einflüsse verändert werden kann. Bei dieser Wechselwirkung ist ein Bezug zu alltäglichen Situationen leicht herzustellen. So kennt man zum Beispiel das Phänomen eines verbesserten Sprachverständnisses, wenn zusätzlich zu den akustischen Reizen auch die synchronen Lippenbewegungen des Sprechers wahrgenommen werden. Weitere bekannte Effekte der audiovisuellen Interaktion sind der McGurk-McDonald-Effekt (McGurk und McDonald 1976) oder auch die Bauchredner-Illusion (Bertelson *et al.* 2000). In beiden Fällen hat das visuelle System einen signifikanten Einfluss auf die auditorische Wahrnehmung.

Diese Arbeit widmet sich mit verschiedenen experimentellen Ansätzen der Untersuchung der beim Streaming-Phänomen auftretenden audiovisuellen Interaktionen. Geeignete physikalisch wohldefinierte Schallreize (Sinustöne) bilden eine ambige Hörsituation. Synchron zu den Tönen werden zusätzliche visuelle Reize präsentiert, um die Ambiguität aufzulösen und das Perzept eines integrierten oder zweier segregierter auditorischer Streams zu verstärken. Sowohl durch die Messung der Hirnaktivität von Probanden mittels Elektro- und Magnetenzephalographie (EEG und MEG) als auch mit psychoakustischen Methoden wird das durch die Reize erzeugte auditorische Perzept beziehungsweise die diesem zu Grunde liegende auditorische Organisation im Kortex untersucht.

Derartige Ansätze sind notwendig, um die audiovisuelle Interaktion und die auditorische Szenenanalyse besser zu verstehen. Gerade bei der Entwicklung technischer Hörhilfen ist man auf intelligente Algorithmen angewiesen, die die akustische Umwelt analysieren. Die Vergleichbarkeit mit dem Hören hörgesunder Menschen könnte z. B. durch ein automatisches Erkennen der aktuellen akustischen Situation (Sprache, Sprache im Störgeräusch, Störgeräusch, Musik, usw.) noch weiter verbessert werden (Dillier 2005).

Zunächst führt die Arbeit in die auditorische Szenenanalyse ein und beleuchtet die grundlegenden Begriffe *Objekte*, *Streaming* und *Grouping*, die der Wahrnehmungspsychologie entlehnt sind. Im weiteren Verlauf werden der ambige Bereich sowie die dem Streaming zu Grunde liegenden Mechanismen und Modelle näher betrachtet. Ausführungen zu multimodaler Reizverarbeitung beschließen das Kapitel.

Das dritte Kapitel gibt einen Überblick über die zum Teil auch hier verwendeten Messmethoden im Kontext von auditorischem Streaming. Neben psychometrischen Methoden haben sich in letzter Zeit auch einige nicht von der Wahrnehmung des Probanden abgeleitete Methoden etabliert. Insbesondere die Messung der Hirnaktivität mittels EEG, MEG sowie funktioneller Magnetresonanztomographie (fMRI) sind hierbei *state of the art*. Mit Ausnahme der fMRI werden diese Methoden in der vorliegenden Arbeit verwendet. Durch die auditorischen und visuellen Stimuli werden elektrische Potentiale und Magnetfelder im Kortex evoziert. Besonders das späte, ereigniskorrelierte Potential *Mismatch-Negativity* (MMN) wird in dieser Arbeit im Kontext des durch Streaming veränderten auditorischen Perzepts untersucht. Ihre Morphologie, neuronalen Generatoren und Anwendungsmöglichkeiten werden zusammengefasst und schließlich ihre konkrete Anwendbarkeit auf Streaming-Fragestellungen sowie ihr Pendant in der visuellen Modalität (visuelle MMN) beschrieben.

Im vierten Kapitel wird die Methodik beschrieben. Zwei EEG-Studien, jeweils eine im nicht-ambigen und eine im ambigen Bereich von Streaming, versuchen durch synchrone auditorische und visuelle Stimulation das auditorische Perzept zu beeinflussen. Weiterhin dient eine MEG-Studie vor allem zur genauen Detektion der neuronalen Quellen sowie eine psychoakustische Studie zur Verifizierung der elektrophysiologisch gewonnenen Ergebnisse.

Alle Ergebnisse, die durch die verschiedenen Experimente gewonnen wurden, sind im fünften Kapitel zusammengefasst. Beginnend mit Ergebnissen des EEG-Experiments im ambigen Bereich, werden im weiteren Verlauf die Resultate der psychoakustischen Studie

und der MEG-Studie dargestellt. Daten, die durch das EEG-Experiment in den nicht-ambigen Bereichen gewonnen wurden, schließen die Darstellung ab.

Weiterhin, im sechsten Kapitel, werden die Ergebnisse unter verschiedenen Aspekten diskutiert. Zunächst wird der Einfluss visueller Stimulation bei unterschiedlicher Stimuluskonfigurationen betrachtet. Besonders die Unterschiede zwischen der ambigen und nicht-ambigen Stimulation werden hier herausgestellt. Zudem werden Aspekte der audiovisuellen Interaktion analysiert. Hier werden die Ergebnisse im Hinblick auf die Verarbeitungsstufe der Interaktion, dem Einfluss der Aufmerksamkeit und schließlich das Auftreten der visuellen MMN durch Betrachtung der neuronalen Quellen diskutiert. Mit einer Zusammenfassung der Ergebnisse und den daraus gewonnenen Erkenntnissen schließt diese Arbeit, nicht jedoch ohne einen Ausblick auf Perspektiven zu geben, die daraus abgeleitet werden können.

Kapitel 2

Auditorische Szenenanalyse

2.1 Einführung in Objekte, Streaming und Grouping

Typischerweise besteht der Schall, der das menschliche Ohr erreicht, aus überlagerten akustischen Komponenten. Diese entstammen unterschiedlichen Schallquellen und stellen als akustische Objekte die auditorische Szene dar. Auf verschiedenen Verarbeitungsstufen des auditorischen Systems wird dieses Gemisch auditorischer Informationen analysiert, getrennt und die segregierten Komponenten ihren ursprünglichen Quellen zugeordnet. Dabei sind die objektunterscheidenden physikalischen Parameter die Grundlage für die Segregation. Beispiele dieser Parameter sind Frequenz, spektraler Gehalt, Schalldruckpegel und räumliche Orientierung, die den psychologischen Kategorien Tonhöhe, Klangfarbe, Lautstärke und Richtung entsprechen. Der Prozess der Kategorisierung in separate neuronale und schließlich auch perzeptuelle Repräsentationen des auditorischen Inputs wird *auditorische Objektbildung* genannt (Bregman 1999, S. 47ff).

Miller (1947) beschrieb eine Auffälligkeit, die er während seiner psychologischen Experimente zu Tonmustern durchführte. Er präsentierte hohe und tiefe Töne alternierend mit einem zeitlichen Abstand (Stimulus Onset Asynchrony, SOA) von 100 ms (vgl. Abb. 2.1). Die Probanden berichteten, dass bei geringem Frequenzabstand die Tonhöhe auf- und abwärts zu springen schien und bei Zunahme des Frequenzabstands ein kritischer Wert erreicht wird, ab dem diese Ordnung scheinbar aufgelöst wird, und verschiedene, nicht zusammenhängende (hohe und tiefe) Tonfolgen wahrgenommen werden. Diese auditorische Objektbildung wird *Streaming* genannt, die Streams stellen hierbei die perzeptuell entstandenen Objekte dar. Miller und Heise (1950) bestimmten die Schwelle, ab der der

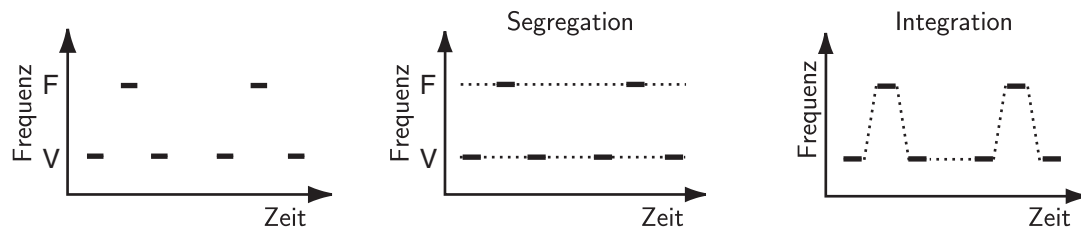


Abb. 2.1: Akustische Tonsequenz, bei der hohe und tiefe Töne alterniert präsentiert werden. Das auditorische Perzept, dargestellt durch die gepunkteten Linien, ist abhängig von den Parametern *Frequenzabstand* und *Präsentationsrate* ein integrierter Stream oder zwei segregierte Streams.

Zusammenhang zwischen den Frequenzbereichen verloren geht und fanden, dass hierzu etwa 15 % Abweichung zur Frequenz des Referenztons notwendig sind.

Sie erkannten auch, dass dieses Phänomen mit der Gestaltpsychologie der visuellen Wahrnehmung (Bodenmann und Schaer 2006) vergleichbar sei, eine Theorie, die sich vorwiegend mit Aspekten der Wahrnehmung im Sinne des Erkennens einer Gestalt (z. B. eine Melodie oder visuelle Struktur) beschäftigt. Dabei spielen sowohl Eigenschaften des Reizmaterials als auch universelle Gesetzmäßigkeiten eine Rolle. Einige der wichtigsten dieser Gesetze seien hier angeführt:

- *Prägnanz.* Jeweils die einfachste und einheitlichste Gestalt der möglichen erkennbaren Strukturen wird wahrgenommen.
- *Nähe.* Sowohl die räumliche als auch die zeitliche Nähe von Elementen beeinflusst die Wahrscheinlichkeit einer Gestaltwahrnehmung.
- *Ähnlichkeit.* Ähneln sich Elemente in Form, Farbe, Klang und weiteren Eigenschaften, werden sie wahrscheinlicher einer gemeinsamen Gestalt zugeordnet.
- *Geschlossenheit.* Durch die Fehlertoleranz der Gestaltwahrnehmung können lückenhafte Figuren (z. B. Kreise) dennoch als Gestalt wahrgenommen werden.

Bezieht man diese visuell definierten Gesetze auf die beschriebenen auditorischen Streaming-Experimente und ersetzt die Raumkoordinaten der visuellen Wahrnehmung durch die Dimensionen Zeit und Frequenz der akustischen Stimuli, können diese mit Begriffen der Gestaltpsychologie beschrieben werden. Stimulusfolgen mit ähnlichem Frequenzspektrum beziehungsweise Zeitstruktur (Stimulationsrate) werden dabei zu

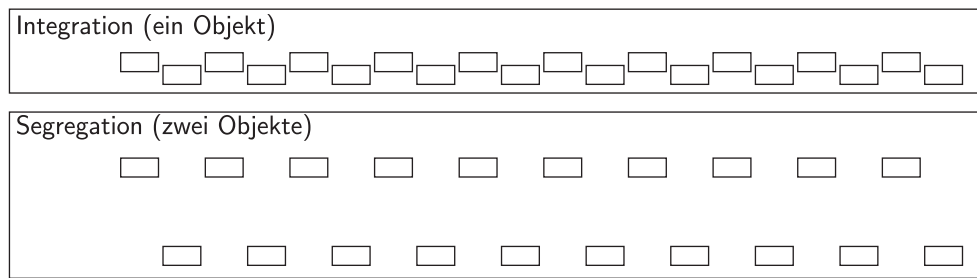


Abb. 2.2: Visualisierung der alternierten Präsentation von zwei Tönen unterschiedlicher Frequenz im *Streaming*-Phänomen. Durch die Veränderung der räumlichen Nähe der Figuren (als Analogon zum Frequenzabstand) wird das visuelle Perzept verändert und kann zur Wahrnehmung eines oder zweier visueller Objekte führen.

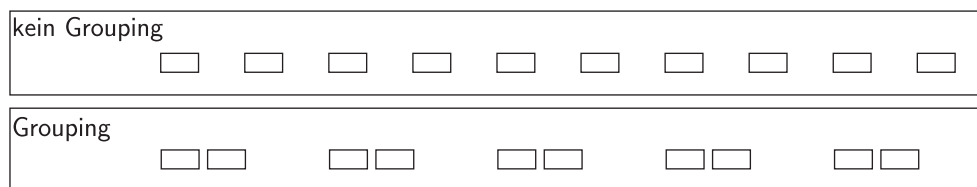


Abb. 2.3: Visualisierung des *Grouping*-Phänomens. Durch Veränderung der räumlichen Nähe der Figuren (als Analogon zum zeitlichen Abstand von Tönen) bilden sich in der visuellen Perzeption Gruppen, die hier aus jeweils zwei Figuren bestehen.

einem Objekt zusammengefasst und von anderen getrennt. Diese Überlegungen wurden von weiteren Arbeiten unterstützt (Bozzi und Vicario 1960; Warren 1968; Bregman und Campbell 1971). Zur Illustration sei auf Abb. 2.2 verwiesen, in der eine aus zwei alternierten Tönen bestehende Stimulationssequenz visualisiert wird. Dort wird allein durch Veränderung des räumlichen Abstands der Muster das visuelle Perzept verändert. Bezieht man die räumliche Nähe auf den zeitlichen Abstand der Töne, kann durch dessen Variation ebenfalls eine Objektbildung erzielt werden (Bozzi und Vicario 1960). Dieses *Grouping*, also die Zuordnung von (auditorischen) Reizfolgen zu Objekten, ist in Abb. 2.3 visualisiert dargestellt.

Einen Höhepunkt erreichten die Arbeiten zu auditorischem Streaming zunächst in den Studien von van Noorden (1975, 1977), die wiederum alterniert präsentierte Töne hoher und tiefer Frequenz untersuchten. Insbesondere die systematische Analyse der Abhängigkeit des auditorischen Perzepts von den Parametern *Stimulationsrate* und *Frequenzabstand* führten zu wesentlichen neuen Erkenntnissen. In einer Folge, wie in Abb. 2.1 dargestellt, wurden 40 ms lange Sinustöne mit einem SOA von 60 bis 150 ms präsentiert.

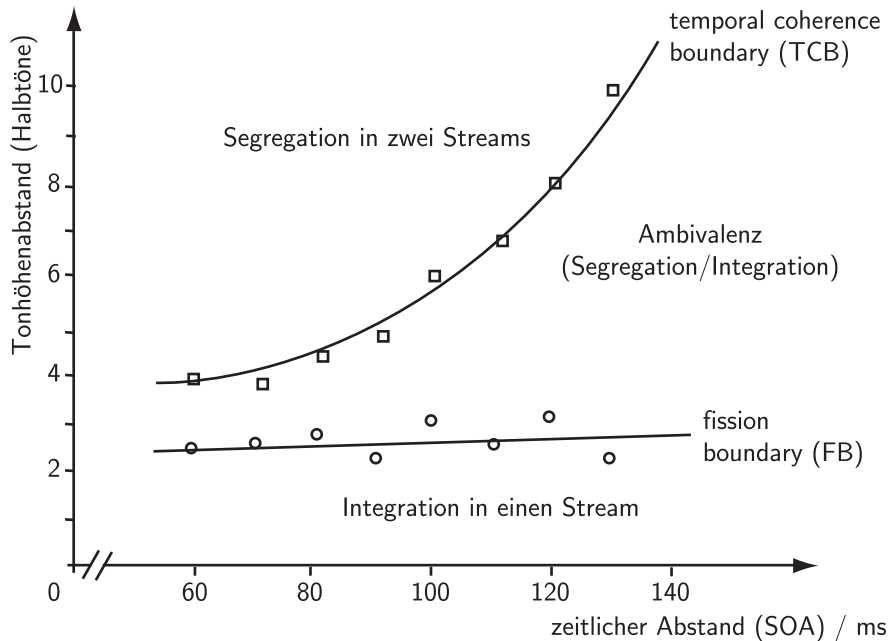


Abb. 2.4: Van-Noorden-Diagramm (nach Bregman 1999, S. 60). Der Tonhöhenabstand ist ein Maß für das Frequenzverhältnis der Töne (1 Halbton $\triangleq f_2/f_1 = \sqrt[12]{2}$). Die Bereiche der Integration und Segregation werden durch die *fission boundary* und die *temporal coherence boundary* begrenzt. Dazwischen ist das auditorische Perzept ambig.

Einer der alternierten Töne hatte eine konstante Frequenz (f_F), die des anderen Tones (f_V) variierte. Beide Parameter, SOA und Frequenzabstand der Folge $V F V - V F V - \dots$ wurden systematisch variiert. Wenn die Sequenz als ein Stream wahrgenommen wurde, waren Töne mit zwei unterschiedlichen Frequenzen in einem Galopp-Rhythmus zu hören. Bei einem großen Frequenzabstand, der zur Segregation führte, konnten zwei separate Streams wahrgenommen werden. Der Stream mit der Frequenz F hatte dabei den Rhythmus $F - - - F - - - F - - - \dots$, während der Rhythmus des anderen Streams ($V - V - V - V - \dots$) schneller war. Die unterscheidbaren Rhythmen eigneten sich somit gut zur Psychometrie des auditorischen Perzepts der Probanden, das in diesem Fall bei $f_F = 1$ kHz registriert wurde.

Durch diese Experimente wurde für verschiedene SOA der minimale Frequenzabstand ermittelt, ab dem es unmöglich wurde, einen zusammenhängenden Stream zu hören. Dessen Verlauf stellt die *temporal coherence boundary* (TCB) dar. Der Frequenzabstand, ab dem die Stimuli nicht mehr segregiert werden können, bildet in Abhängigkeit vom SOA die *fission boundary* (FB). Beide Verläufe können in einem Parameterdiagramm – dem



Abb. 2.5: Beispiel für die Verwendung von Streaming (gepunktete Linien) in barocken Kompositionen. Die wiederholten Töne gleicher Frequenz bilden einen Stream, der sich von dem durch die Töne mit veränderter Frequenz und langsameren Rhythmus gebildeten Stream abgrenzt (G. F. Telemann, Sonate Nr. 4 in C: Allegro aus *Der getreue Musikmeister*).

Van-Noorden-Diagramm – dargestellt werden (vgl. Abb. 2.4). Die Abhängigkeit beider Grenzen vom SOA ist unterschiedlich. Die FB wird vom SOA nur schwach beeinflusst, wohingegen die TCB deutlich mit wachsendem SOA zu größeren Frequenzabständen ansteigt (van Noorden 1975). Auf der Wahrnehmungsebene stellt die TCB die Grenze dar, ab der das auditorische System durch automatische, basale Prozesse das akustische Signal in zwei Streams trennt. Die FB hingegen kennzeichnet die Grenzen der aufmerksamskontrollierten Prozesse, Streams durch Selektion zu generieren (van Noorden 1977). Intrakortikale elektrophysiologische Messungen, die eine objektive Beurteilung der Neuronenaktivität zulassen, wurden kürzlich bei Primaten durchgeführt und konnten das Van-Noorden-Diagramm reproduzieren (Fishman *et al.* 2004).

Das heute unter dem Namen „Streaming“ bekannte Phänomen war, nebenbei bemerkt, neben anderen auch schon den Barock-Komponisten Bach, Vivaldi und Telemann bestens bekannt. Ein Beispiel aus den Kompositionen Telemanns illustriert die Wahrnehmung von zwei parallelen Streams, obwohl alle Töne aufeinander folgend notiert sind (vgl. Abb. 2.5).

In jüngster Vergangenheit sind zwei Mechanismen, mit denen komplexe auditorische Reize analysiert werden, verstärkt in das Interesse der Forschung gerückt. Zum einen handelt es sich um die Zuordnung simultaner Frequenzkomponenten in auditorische Objekte

(Grouping) und zum anderen um die Organisation von sequentiellen Tönen in Streams (Streaming). Über die dem Grouping zu Grunde liegenden neuronalen Mechanismen ist erst wenig bekannt. Wesentlich umfangreichere Erkenntnisse konnten bereits über das Streaming gewonnen werden. Im Fokus des Interesses stand dabei auch die Aufklärung von neuronalen Streaming-Korrelaten, wobei es derzeit als eher unwahrscheinlich angesehen wird, deren Ursprung auf eine bestimmte Neuronenpopulation eingrenzen zu können – auch, weil Streaming durch viele Faktoren wie *Selektivität*, *Adaptation* und *Aufmerksamkeit* beeinflusst werden kann. Vielmehr sollte versucht werden, die das Streaming beeinflussenden Prozesse sowie ihre Interaktionen besser zu verstehen (Carlyon 2004).

Fast alle Untersuchungen zu Streaming bei Menschen sind an erwachsenen Probanden durchgeführt worden. Aber auch bei Kindern tritt das Streaming-Phänomen nachweislich auf. Sussman *et al.* (2001) zeigten durch MMN-Studien (MMN: vgl. Kap. 3.4) mit Kindern im Alter von 7 Jahren, dass Frequenzunterschiede zwischen sequentiell präsentierten Tönen zu deren Segregation führen. Durch ein psychoakustisches Experiment wurden die Ergebnisse verifiziert und festgestellt, dass Streaming in diesem Alter präattentiv geschieht. Plausibel erscheint den Autoren, dass die Anlage für Streaming schon von Geburt an vorhanden ist und seine Reifung der Entwicklung des Hörsystems folgt. So sind, wie auch weitere MMN-Studien vermuten lassen, möglicherweise in früheren Lebensjahren größere Frequenzabstände notwendig, um Streaming hervorzurufen (Morr *et al.* 2002).

2.2 Streaming

Das auditorische Streaming ist ein wesentlicher Aspekt der auditorischen Szenenanalyse. Sämtliche Experimente der vorliegenden Arbeit basieren auf Stream-Segregation, vorwiegend im ambigen Bereich des Van-Noorden-Diagramms. Im Folgenden soll deshalb das Phänomen des Streamings aus der auditorischen Szenenanalyse herausgegriffen und näher betrachtet werden. Nach einleitenden Bemerkungen zum ambigen Perzept wird ein Schwerpunkt auf die vielfältigen Einflussfaktoren für Streaming gelegt.

2.2.1 Ambiguität

Betrachtet man das Van-Noorden-Diagramm (vgl. Abb. 2.4), stellt man fest, dass die FB und TCB nicht identisch sind. Zwischen beiden Grenzen existiert ein Parameterbereich, in dem beide Perzepte, Integration und Segregation, möglich sind. Die Breite dieses ambigen

Abb. 2.6: Rubin'sche Vase als Beispiel für Mehrdeutigkeit auf der Wahrnehmungsebene. Die Abbildung kann entweder als weiße Vase auf schwarzem Hintergrund oder als zwei schwarze Gesichter, zwischen denen sich eine weiße Fläche befindet, gesehen werden (nach Zimbardo 1995).



Bereichs nimmt mit dem SOA stark zu, wobei diese Zunahme vor allem durch dessen Einfluss auf die TCB bestimmt wird.

Die Wahrnehmung im ambigen Bereich ist empfindlich für Instruktionen, denen die Probanden während des Hörens folgen sollen. Beispielsweise kann man dazu auffordern, eine Melodielinie oder Klangfarbe zu verfolgen. Wenn dies gelingt, entspricht das vorliegende Perzept der Segregation. Dabei wurde im ambigen Bereich eine Zunahme der Segregationstendenz beobachtet, je länger die Stimulation verfolgt wird (Bregman 1978). Die Initialerwartung des auditorischen Systems ist offenbar das Perzept der Integration. Erst Eindrücke, die dieser Grundannahme in einem überkritischen Maß widersprechen, erzeugen segregierte Streams. Dieses Prinzip gilt auch für den Bereich oberhalb der TCB, wo sich nicht sofort nach Stimulationsbeginn, sondern erst nach etwa 10 – 20 s das Perzept segregierter Streams einstellt. Nach dieser *buildup*-Zeit ist das Perzept für sinusförmige Stimuli relativ stabil, während es bei anderen Stimulusformen im Laufe der Stimulationszeit stärkeren Veränderungen unterliegt (Anstis und Saida 1985). Bei der Präsentation ambiger Musikfiguren wurde beobachtet, dass die Mehrdeutigkeit bei Stimuli unterschiedlicher Klangfarbe, Lautheit oder Stimmung leichter aufgelöst werden kann als bei Verwendung von Reintönen (Smith *et al.* 1982).

Auch für den ambigen Bereich des auditorischen Streamings existiert ein visuelles Analogon. Das visuelle Perzept der in Abb. 2.6 dargestellten Rubin'schen Vase ist, wie der Betrachter leicht feststellt, ambig. Entweder wird die Abbildung als ein weißer Gegenstand auf schwarzem Hintergrund oder zwei schwarze Gegenstände, zwischen denen sich eine weiße Fläche befindet, interpretiert. Auch hier kann durch Instruktionen, z. B. konzentriertes Betrachten der Vase oder des Gesichtes, das jeweilige Perzept stabilisiert werden. Bei unkonzentrierter Betrachtung der Abbildung hingegen ändert sich das Perzept wiederholt.

Wenn beim Hören einer *ABA*-Tonsequenz durch kontinuierliche Veränderung der auditorischen Parameter (SOA und Frequenzabstand) der ambige Bereich von verschiedenen Richtungen durchlaufen wird, können Hystereseeffekte in der Wahrnehmung gefunden werden. Diese Ergebnisse sind zwar bereits in der wahrnehmungspsychologischen Definition von FB und TCB durch van Noorden (1975) implizit enthalten, wurden aber auch durch die Entwicklung von analytischen Modellen zu Streaming bestätigt (Almonte *et al.* 2005).

Eine segregierte Wahrnehmung spiegelt sich im Magnetenzephalogramm (MEG: vgl. Kap. 3.2.2) in der Zunahme bestimmter Komponenten evozierter Felder wider, die ein bei der Selektion eines der beiden Streams vergrößert wahrgenommenes SOA anzeigen (Gutschalk *et al.* 2005). Zudem konnte gezeigt werden, dass vermutlich Repräsentationen von Streaming in nichtprimären auditorischen Kortextbereichen existieren. Durch funktionelle Kernspintomographie (fMRI), ebenfalls mit Stimuli im ambigen Bereich, wurden Korrelate für Streaming im intraparietalen Sulcus gefunden (Cusack 2005).

2.2.2 Mechanismen von Streaming

Die von Bregman (1999, S. 47ff) entwickelte Streaming-Theorie postuliert die Existenz von zwei verschiedenen neuronalen Prozessen. Ein von externen Stimulusparametern abhängiger, automatisch wirkender *bottom-up*-Prozess bewirkt demnach die Partitionierung der Töne, wohingegen ein von Erfahrung, Erwartung und Aufmerksamkeit abhängiger (*schemabasierter*) Prozess die Selektierung von Streams steuert und auch *top-down*-Prozesse einschließt. Alle Prozesse laufen interaktiv ab, um die auditorische Szene zu rekonstruieren. Das auditorische System ist dadurch mit einem Mechanismus ausgestattet, der es erlaubt, unterschiedliche Repräsentationen von verschiedenen auditorischen Quellen zu erzeugen (Bregman 1999; Bey und McAdams 2002). Der aus *bottom-up*- und *top-down*-Komponenten bestehende Streaming-Prozess stellt damit ein kognitives System dar, das optimal an die Umgebung adaptiert ist. Der *top-down*-Einfluss spielt besonders dann eine Rolle, wenn die akustischen Parameter, wie in Kap. 2.2.1 beschrieben, verschiedene Organisationen zulassen (Sussman *et al.* 1998a).

Exogene Faktoren

Eine Eigenschaft der *bottom-up*-Prozesse von Streaming ist ihre Robustheit gegenüber Aufmerksamkeitsveränderungen. Im Bereich der automatischen Segregation, also oberhalb

der TCB, stellt sich das Perzept segregierter Streams schon nach wenigen Tönen ein, auch wenn bewusst versucht wird, diese zu integrieren (Bregman *et al.* 2000). Die FB ist gegenüber Veränderungen des SOA und des Frequenzabstands relativ robust. Für den Verlauf der TCB hingegen ist die durch das SOA repräsentierte Stimulationsrate der einflussreichste Faktor, mehr noch als die Dauer der Töne oder deren Frequenzabstand.

Dass der Frequenzabstand aufeinander folgender Töne einen großen Einfluss auf Streaming hat, wird oft auf die bei geringem Frequenzabstand der aufeinander folgenden Töne entstehende Überlappung der erzeugten Aktivierungsmuster auf der Cochlea zurückgeführt (Moore 1998). Somit müsste Streaming peripheren Ursprungs sein. Wäre dies der Fall, müsste man allerdings bei Personen mit einseitigem cochleären Hörverlust eine Verschiebung der FB zu größeren Frequenzabständen im Vergleich zur normal funktionierenden Seite beobachten. Das konnte aber nur in wenigen Fällen nachgewiesen werden, während die meisten eine unveränderte FB aufwiesen (Rose und Moore 1997). Gegen das cochleäre Aktivierungsmodell spricht auch das Auftreten von Streaming, wenn die Töne der verschiedenen Streams gleiche cochleäre Erregungsmuster haben und demselben Ohr präsentiert werden. Die Annahme eines allein peripheren Ursprungs für Streaming kann somit nicht aufrechterhalten werden (Moore und Gockel 2002). Systematische Experimente bestätigten dies auch dadurch, dass die FB nicht ausschließlich von der Diskriminierbarkeit aufeinander folgender Töne abhängt (Rose und Moore 2005). Stream-Segregation ist demnach immer dann möglich, wenn es einen wahrnehmbaren Unterschied zwischen den Streams gibt, der nicht auf den Frequenzunterschied beschränkt sein muss (Moore und Gockel 2002).

In komplexen auditorischen Szenen ist interessanterweise die Frage nach dem Frequenzabstand zweier Streams von den Probanden nicht immer leicht zu beantworten (Baker *et al.* 2000). Auf der Suche nach einer entsprechenden Metrik wurden drei Töne unterschiedlicher Frequenz in einem Streaming-Paradigma zyklisch angeordnet. Der Ton mit der mittleren Frequenz wurde in seiner zeitlichen Anordnung (Abstand zum vorhergehenden und folgenden Ton) sowie seiner Frequenz variiert. Die Zuordnung dieser Töne zu den beiden anderen Streams wurde psychoakustisch überprüft. Dabei erkannte man, dass der Mittelpunkt der psychometrischen Funktion nicht mit dem räumlichen Mittelpunkt auf der Cochlea, die die Frequenz logarithmisch auf die Länge abbildet, übereinstimmt. Es scheint somit eine bevorzugte Zuordnung von Stimuli zu tieffrequenten Streams zu geben, was eine leichtere Segregation von tieffrequenten Streams bedeutet.

Als eine Erklärung zur Funktionsweise von Streaming geht Nelken (2004) zunächst davon aus, dass die auditorischen Reize im Kortex mehr in Form von auditorischen Objekten

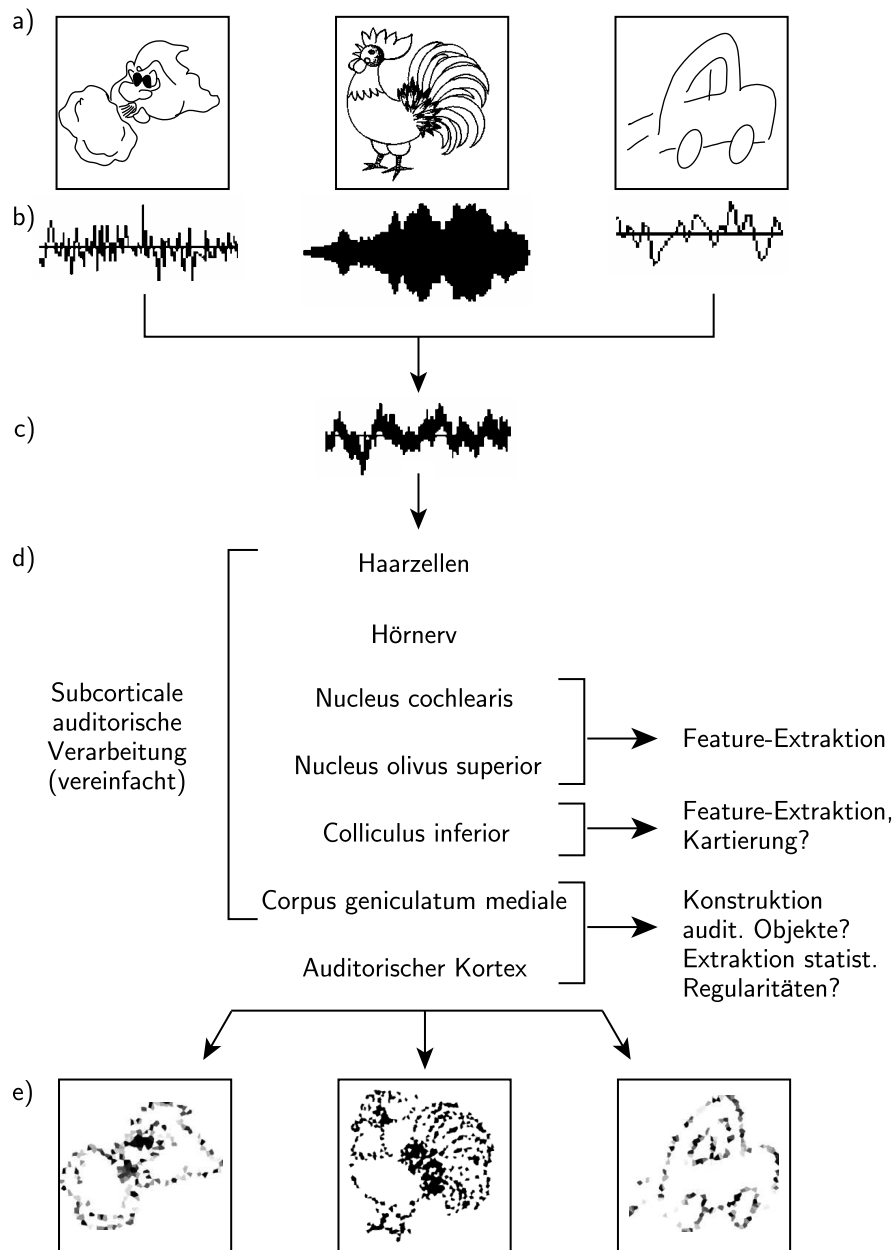


Abb. 2.7: Prozesse im afferenten auditorischen Pfad der Hörbahn (nach Nelken 2004). Multiple Schallquellen a) produzieren verschiedene akustische Signale b). Diese werden als Überlagerung c) vom Ohr aufgenommen und an den Haarzellen in neuronale Impulse gewandelt. Das auditorische System d) führt Feature-Extraktionen bereits auf den Verarbeitungsstufen des Hirnstamms aus. Im Colliculus inferior (Mittelhirn) konvergieren bereits mehrere Feature-Streams und generieren (möglicherweise überlappende) Feature-Maps. Im auditorischen Kortex e) ist der Schall nicht mehr in Form physikalischer Eigenschaften repräsentiert, sondern als auditorische Objekte.

und weniger als invariante akustische Eigenschaften abgelegt werden. Auf dem Weg der Signale zum auditorischen Kortex werden demnach die Eigenschaften extrahiert und durch Konvergenz entsprechende *Feature-Maps* generiert, die zur Objekt-Repräsentation führen (vgl. Abb. 2.7).

Einen bedeutenden Einfluss auf die Messbarkeit und damit den experimentellen Zugang zu Streaming hat die Aufgabe, die von den Probanden während einer Messung erfüllt werden muss. Generell sind Aufgaben, die eine Beurteilung der zeitlichen Anordnung von Tönen oder des Rhythmus beinhalten, schwieriger zu erfüllen, wenn die zu vergleichenden Töne Teil zweier segregierter Streams sind (Cusack und Roberts 2000; Bregman und Campbell 1971). Bei einigen Instruktionen wird immer dann eine gute Leistung der Probanden beobachtet, wenn in komplexen Streaming-Experimenten einfache, von den Probanden leicht zu segregierende Muster enthalten sind. So wird zum Beispiel, wenn zwei verschachtelte Melodien segregiert gehört werden sollen, die beste Leistung erzielt, wenn schon die in den jeweiligen Melodien enthaltenen Töne leicht segregiert werden können (Hartmann und Johnson 1991). Eine der einfachsten Aufgaben stellt die Suche nach Zielreizen (Targets) dar. In dieser Arbeit werden visuelle Targets verwendet, deren Auftreten von den Probanden durch einen Tastendruck quittiert werden sollte.

Ein weiteres Phänomen ist, dass ein Streaming-Perzept offenbar in den Ausgangszustand zurückgesetzt werden kann, wenn plötzlich Änderungen in der Stimulationssequenz auftreten. Anstis und Saida (1985) präsentierten eine *ABA*-Tonsequenz zunächst nur einem Ohr. Nach einer Zeit, die lang genug war, um die Töne zu segregieren, wechselten sie die Seite und präsentierten die Stimuli dem anderen Ohr. Die Töne wurden zunächst wieder integriert wahrgenommen wurden. Rogers und Bregman (1993) konnten dies durch weitere Experimente betätigen. Sie zeigten zudem die Möglichkeit, ein Zurücksetzen auch durch Lautheitsveränderungen hervorzurufen. Aber auch kurze Lücken in der Präsentation führen zum Zurücksetzen des Streaming-Perzepts (Cusack *et al.* 2004). Lässt man hingegen die Stimuluspräsentation konstant, führt das Fehlen von plötzlichen Änderungen von Eigenschaften einer Sequenz wiederum zu einem anderen Phänomen, der Modulation von Streaming in ambigen Situationen: Stream-Segregation wird, über einen Zeitraum von mehreren Sekunden betrachtet, stärker und schwächer (Anstis und Saida 1985). Erklärbar ist dies dadurch, dass dem Streaming zentrale Mechanismen zu Grunde liegen (Moore und Gockel 2002). Die Zeitkonstante dieser Schwankung ist interindividuell verschieden und liegt im Bereich von 1,4 s bis 7,8 s und hängt offenbar von der Erfahrung der Probanden, z. B. deren musikalischer Ausbildung, ab (Beauvois und Meddis 1997).

Zusätzliche Faktoren, die Streaming beeinflussen, sind Unterschiede von Tönen, die durch interaurale Zeit- oder Intensitätsdifferenzen hervorgerufen werden (Moore 1997; Gockel und Micheyl 1999; Bregman 1999). Deren Einfluss ist in ihrer Stärke mitunter gleich dem Einfluss der peripheren Einflüsse, meist jedoch eher schwach. Eine Studie von Vliegen *et al.* (1999) zeigt einen reduzierten Einfluss der Grundfrequenz von Tönen auf Streaming, wenn viele Harmonische enthalten sind. Die Grundfrequenz ist demnach der am stärksten segregierend wirkende Faktor.

Eine weitere wichtige Frage ist, welche Stellung Streaming im Kontext anderer basaler auditorischer Verarbeitungsprozesse hat. Hierzu sei zunächst das Grouping erwähnt, also die zeitliche Integration von auditorischen Reizen im auditorischen Kurzzeitspeicher (Zwislocki 1960). Hierbei wird davon ausgegangen, dass akustische Signale, die innerhalb eines Zeitfensters von 200 ms Dauer präsentiert werden, als eine Einheit wahrgenommen werden können (Cowan 1984, 1995). Diese Annahme erklärt allerdings nicht die Verarbeitung von Sprachreizen, denn verschiedene, innerhalb dieser 200 ms präsentierte Vokale könnten dann nicht mehr perzeptuell getrennt werden, sie würden sich gegenseitig maskieren. Ein Experiment von Yabe *et al.* (2001) ging der Frage nach, wie Streaming mit der zeitlichen Integration wechselwirkt. Mittels MEG-Messungen wurde die MMN auf Tonauslassungen in verschiedenen Streaming-Paradigmen ausgewertet. Es konnte gezeigt werden, dass Streaming der zeitlichen Integration möglicherweise vorausgeht. Wenn große spektrale Veränderungen des auditorischen Signals innerhalb kurzer Zeit auftreten – so folgern die Autoren – kann Stream-Segregation der Maskierung vorbeugen und somit zum Sprachverstehen beitragen. Hierzu sei angemerkt, dass diese Studie in Japan durchgeführt wurde, wo überwiegend tonale Sprachen gesprochen werden.

Zur Aufklärung der Beziehungen von Streaming und Grouping zueinander haben die Arbeiten von Sussman *et al.* (1998a) und Sussman (2005) wesentlich beigetragen. Sie zeigen, dass die Integration (Grouping) von auditorischen Reizen auf der Basis segregierter Streams arbeitet. Dadurch wird es möglich, Muster innerhalb von Streams zu identifizieren und, darauf aufbauend, Sprache oder Musik zu verstehen. Stream-Segregation selbst sei ein früher, basaler Prozess, der von Stimuluseigenschaften gesteuert wird und vor vielen anderen basalen Prozessen wirkt. Das verwendete Paradigma bestand aus einer alternierenden Präsentation eines hoch- und eines tieffrequenten Tons (Sussman 2005), die gelegentlich durch Frequenz-Devianten ersetzt wurden. In einer der beiden so entstehenden Melodielinien traten immer zwei unmittelbar nacheinander und in der anderen Melodielinie entweder zwei Devianten unmittelbar nacheinander oder einer einzeln auf. Wenn der Doppel-Deviant zufällig auftrat, wurde auch durch den zweiten Deviant eine

MMN evoziert. Daraus wird gefolgert, dass, wenn der Proband nicht auf die Reize achtet, die auditorische Objektbildung erst nach der Verarbeitungsstufe von Streaming geschieht.

Endogene Faktoren

Unabhängig vom Einfluss peripherer Faktoren können auch zentrale Prozesse Streaming hervorrufen oder es beeinflussen. Die Aufmerksamkeit oder Konzentration auf die Stimuli ist solch ein *top-down*-Prozess. Im Experiment von Sussman *et al.* (2002) wurden Deviantreize zum einen in einer Fünf-Ton-Sequenz mit fester Reihenfolge und zum anderen in einer um Target-Reize erweiterten zufälligen Reihenfolge präsentiert. Durch EEG- und MEG-Messungen wurde hierbei die MMN (vgl. Kap. 3.4) untersucht und zwischen den unterschiedlichen Stimulationen in Abhängigkeit von der Aufmerksamkeit der Probanden verglichen. Ein Auftreten der MMN bedeutet, dass die Deviantreize detektiert wurden und kein Grouping vorliegt. Zunächst wurde bei Konzentration auf die Targetreize eine in der Oddball-Bedingung stärker als in der Passivbedingung ausgeprägte MMN gezeigt. Dies wird auch bei der geordneten Stimuluspräsentation gefunden. Die Konzentration auf ein in der Stimmulationssequenz vorhandenes Tonmuster hingegen führte dazu, dass die MMN verschwindet. Hierdurch wurde gezeigt, dass sich zumindest sequentielles Grouping durch *top-down*-Prozesse steuern lässt, wovon auch eine frühere Arbeit von Sussman *et al.* (1998a) ausgeht.

Auch Carlyon *et al.* (2001) studierten den direkten Einfluss der Aufmerksamkeit auf Streaming und verwendeten dazu ein *ABA*-Schema, das dem linken Ohr präsentiert wurde. Auf der rechten Seite wurden in der Nullbedingung keine Stimuli und in den ersten 10 s der Kontroll- und Aufmerksamkeitsbedingung Schmalbandrauschen präsentiert. Dieses sollte von den Probanden in der Kontrollbedingung ignoriert und in der Aufmerksamkeitsbedingung bewertet werden. Die Ergebnisse der Null- und Kontrollbedingung sind vergleichbar, weisen aber, nachdem die Aufmerksamkeit wieder auf die *ABA*-Stimuli gerichtet war, einen reduzierten Segregationseffekt in der Aufmerksamkeitsbedingung auf. Zwar wird hier ein Aufmerksamkeitsseffekt bei Streaming gezeigt, jedoch ist nicht klar, inwieweit nicht auch die räumliche Lokalisation oder die Verschiedenartigkeit der Stimuli einen Einfluss haben (Moore und Gockel 2002).

Ein weiteres bedeutendes Experiment, das die Aufmerksamkeitsinflüsse auf Streaming untersucht, ist von Sussman *et al.* (2005) durchgeführt worden. In drei parallel präsentierten Melodielinien waren Deviantreize enthalten. In einer der Bedingungen sollten die Probanden die hochfrequente Stimulusfolgen aufmerksam verfolgen, in einer anderen zur

Ablenkung einen Videofilm sehen. Die Aufmerksamkeit auf die jeweilige Melodie führte dazu, dass nur durch die in ihr enthaltenen Devianten eine MMN evoziert wurde. Die Bedingung ohne Aufmerksamkeit hingegen führte zur Evozierung einer MMN durch die Devianten aller drei Melodielinien. Es wird somit deutlich, dass die Verarbeitung auditorischer Stimuli durch Aufmerksamkeit verändert werden kann, insbesondere, wenn einzelne Streams durch diese selektiert werden. Der Aufmerksamkeitseinfluss ist so stark, dass damit sogar Pegelunterschiede von bis zu 15 dB zwischen den Stimuli verschiedener Streams ausgeglichen werden können (Botte *et al.* 1997). Erst danach rückt der lautere Stream unfreiwillig in den Fokus der Aufmerksamkeit im Sinne von Streaming.

Auch unterschiedliche Klangfarben (Timbre) können einen segregierenden Effekt haben. Obwohl die Verarbeitung von Timbre gewöhnlich in der rechten Hemisphäre erwartet wird (Zatorre *et al.* 2002), zeigen die Ergebnisse von Deike *et al.* (2004) eine verstärkte Aktivierung des linken auditorischen Kortex. Auch dadurch wird ein Hinweis auf einen *top-down* gesteuerten selektiven Prozess gegeben, der bei Streaming wirkt.

2.2.3 Modelle für Streaming

In den letzten Jahren sind vielfach robuste und reproduzierbare Studien zu Stream-Segregation durchgeführt worden, die auch in Beziehung zu anderen Integrations- und Segregationsphänomenen der Wahrnehmung gesetzt werden können (Review: Moore und Gockel 2002). Vielfach wurde versucht, Streaming durch die Prinzipien der Gestalt-Theorie zu erklären (Bodenmann und Schaer 2006; Bregman 1999). Ein weiteres Modell für Streaming ist auch das der *Kanalisierung* (Broadbent 1958). Es beruht auf der Annahme, dass selektive Aufmerksamkeit auf eine singuläre Wahrnehmungsdimension (z. B. Tonhöhe) zu Streaming führt. Allerdings konnten hiermit komplexere Streaming-Phänomene, die z. B. auf Klangfarbe beruhen, nicht erklärt werden (Almonte *et al.* 2005).

Neben diesen und weiteren psychologischen Erklärungen werden zur Zeit zwei neurophysiologische Theorien über die Integrationsweise des neuronalen Systems diskutiert: Die *Binding*-Theorie nimmt an, dass ein integriertes Perzept immer dann entsteht, wenn kortikale Aktivität synchronisiert ist (Engel *et al.* 2001). Die andere Theorie ist auf dem Gebiet der multimodalen Integration entstanden und nimmt die Existenz von informativen Konvergenzzonen an (Calvert 2001). Die Konvergenzzonen sind dabei kortikale und subkortikale Netzwerke, die dort ankommende Informationen integrieren und damit ein Segregations- bzw. Integrationsperzept erzeugen.

Ein Modell, das mit basalen physiologischen Prinzipien arbeitet, wurde von Beauvois und Meddis (1991, 1996) entwickelt. Es erklärt die Entstehung von Stream-Segregation, die im Van-Noorden-Diagramm (vgl. Abb. 2.4) gemessenen Grenzen *fission boundary* und *temporal coherence boundary* und die Schwelle des minimalen Frequenzunterschiedes, um Töne getrennten Streams zuordnen zu können. Es beinhaltet auch, dass die auditorische Szenenanalyse auf der Basis von Prinzipien der Gestalt-Psychologie arbeitet und somit ein Produkt basaler Verarbeitungsprozesse ist. Ein wesentliches Prinzip dieses Modells ist der Überlapp von Erregungsmustern, die durch eine Serie von Stimuli auf der Cochlea entstehen. Je höher der Grad der Überlappung ist, umso ausgeprägter ist das Grouping. Neuere Ergebnisse zeigen jedoch, dass Streaming auch dann auftreten kann, wenn das Leistungs- und Frequenzspektrum beider Streams ähnlich oder gleich ist, was zu gleichen Erregungsmustern auf der Cochlea führt und somit *per se* von diesem Modell nicht erklärt werden kann (Moore und Gockel 2002).

Von Almonte *et al.* (2005) sind zwei weitere Modelle vorgeschlagen worden, die ebenfalls verschiedene Streaming-Szenarien erklären und hier näher beschrieben werden sollen. Das erste ist das *stream classification*-Modell. Es geht von zwei verbundenen Netzwerken aus, von denen das erste tonotop (eindeutige Ortsabbildung der Frequenz) organisiert ist. Dort ist die neuronale Aktivität $\Psi(x,t)$ orts- und zeitabhängig definiert. Das zweite System klassifiziert die resultierende raumzeitliche Dynamik dieses neuronalen Feldes. Man geht von einem Input-Signal $p(x,t)$ aus, das die spektralen Komponenten zur Zeit t am Ort x kodiert. Zudem existiert ein neuronales Netzwerk mit n Unternetzwerken mit deren jeweiligen Aktivitäten Ψ_i

$$\Psi(x,t) = (\Psi_1(x,t), \Psi_2(x,t), \dots, \Psi_n(x,t)) \quad (2.1)$$

Das ankommende auditorische Signal wird in einem tonotop organisierten neuronalen Feld zwischenverarbeitet, dessen Dynamik sich wie folgt beschreiben lässt:

$$\dot{\Psi}(x,t) = G(\Psi(x,t)) + p(x,t) \quad (2.2)$$

wobei G eine nichtlineare Funktion ist, die die lokale Aktivität an jedem Netzwerkknoten beschreibt. $p(x,t)$ ist die normierte Input-Funktion für jeden Knoten. Ein auf $\Psi(x,t)$ wirkendes Funktional $I_i(t) = I_i(\Psi(x,t))$ ordnet der i -ten Komponente von Ψ das aktivierte Perzept $y_i(t)$ zu:

$$\dot{y}_i(t) = g_i(y_i) - \sum_{j \neq i} c_{ij}(y_i - y_j) + I_i(t) \quad (2.3)$$

g_i ist hierbei eine nichtlineare Funktion, die die intrinsische Dynamik von $y_i(t)$ beschreibt, $c_{ij}(y_i - y_j)$ die Kopplung durch die typischerweise sigmoidale Funktion c_{ij} .

In diesem Modell findet ein Vergleich zwischen den Perzepten und nicht zwischen den Streams statt, was die Koexistenz von zwei Streams mit identischer Frequenz erlaubt. Das ist ein Vorteil des beschriebenen Modells, denn dieses Prinzip ist im folgenden Modell nicht enthalten.

Das zweite Modell geht von der Aufteilung des ankommenden Signals in Streams mit unterschiedlichen Energien aus, die im Wettbewerb miteinander stehen. Beispiele für eine solche *winner-takes-all*-Architektur, wie sie diesem *stream competition*-Modell zu Grunde liegt, sind z. B. die Cochlea-Modelle und Filterbänke. Man geht wieder vom Input-Signal $p(x,t)$ und den neuronalen Aktivitäten $\Psi(x,t)$ aus. Jedes Unternetzwerk repräsentiert einen Stream, der in Wettbewerb mit den anderen steht. Die Kopplung bezieht sich entweder auf die Frequenz (als Ort im Kortex abgebildet) oder die Zeit und ist entweder eine globale Erregung oder Inhibition. In der Konsequenz wird nur ein Stream als dominant hervorgehen, obschon es Übergänge zwischen Streams geben kann, die dieselbe Ortskomponente haben. Die Gleichung, die dieses System beschreibt, ist die folgende:

$$\dot{\Psi}_i(x,t) = G_i(\Psi_i(x,t)) - \sum_j C_{ij} \Psi_j(x,t) + Mp(x,t) \quad (2.4)$$

wobei G_i , wie zuvor, eine nichtlineare Funktion ist, die die lokale Aktivität an jedem Netzwerkknoten beschreibt. $C_{ij} < 0$ ist die Kopplungskonstante und M die Wichtungskonstante, die den Eingang jedes Knotens normiert. Als Konsequenz entsteht eine Aktivitätsverteilung $v_i(x)$ für den Stream i mit der Amplitude $s_i(t)$. Die gesamte Information über das Wahrnehmungsmuster ist dann in den aktivierten Streams $\{s_i(t)v_i(x)\}$ enthalten.

Das *stream competition*-Modell beruht somit auf der gegenseitigen Beeinflussung neuronaler Datenströme, von denen jeweils der dominantere das auditorische Perzept formt. Es kann, so konstatieren die Autoren, viele reale Streaming-Phänomene reproduzieren. Auch das Van-Noorden-Diagramm (vgl. Abb. 2.4) kann damit erklärt werden. Das *stream classification*-Modell erlaubt hingegen Streams gleicher Frequenz. Demzufolge müssten Streaming-Prozesse nur kortikal ablaufen. Es scheint also eine Kombination beider Modelle zu sein, die reale biologische Systeme beschreibt. Die peripheren Prozesse scheinen mehr wie das *stream competition*-Modell, die kortikaleren Prozesse mehr wie das *stream classification*-Modell abzulaufen (Almonte *et al.* 2005).

2.3 Multimodale Reizverarbeitung

Die Sinneseindrücke, die im Kortex entstehen, sind fast immer komplex und bestehen aus Reizen mehrerer Modalitäten. Eine bedeutende Rolle spielt die Kombination von auditorischen und visuellen Reizen, wie sie u. a. bei der zwischenmenschlichen Kommunikation, beim Lesen, Fernsehen, aber auch in vielen anderen Bereichen auftritt. Im menschlichen Kortex sind etwa etwa 10^{10} Neurone enthalten, welche schon bis zur Hälfte in visuelle Funktionen involviert sein können (Palmer 1999). Die Anzahl der Neurone im auditorischen System ist hingegen viel kleiner. Schon an diesem Punkt drängt sich die Vermutung auf, dass das visuelle System auch andere Aufgaben als die rein visuelle Verarbeitung wahrnimmt. Besonders das *Verstehen* von externen Informationen wird vom visuellen System dadurch unterstützt, dass sich viele reale Szenen in Raum und Zeit separieren lassen. Insbesondere die räumliche Komponente ist dem visuellen System zugänglich und kann dort verarbeitet werden (Kubovy und van Valkenburg 2001).

Zunächst ist hier ein kurzer Ausflug in die Psychologie nützlich, der in das Zusammenwirken der auditorischen und visuellen Modalität einführt. Die Wahrnehmung von Sinnesreizen dient, wie bereits in Kap. 2.1 ausgeführt, dem Erkennen und Abstrahieren von Objekten. Ihre Betrachtungen zu auditorischen und visuellen Objekten schließen Kubovy und van Valkenburg (2001) mit einer schematischen Übersicht, die das Zusammenwirken beider Modalitäten zusammenfasst (vgl. Abb. 2.8). Jede Modalität wird dort durch zwei Pfade, dem *Wo*- und dem *Was*-Pfad repräsentiert. In der Bildmitte wird die enge Verbindung des auditorischen mit dem visuellen System symbolisiert. Besonders deutlich wird diese durch den bemerkenswerten Umstand, dass das auditorische *Wo*-System durch die raum-zeitliche Projektion im Dienst des visuellen *Wo*-Systems steht. Hierin liegt auch die Asymmetrie dieser Verbindung begründet. Prinzipiell können jedoch beide Systeme Teile ihrer Informationen in Form von räumlichen oder zeitlichen Kodierungen auf das andere System projizieren. An den Außenseiten der Abbildung ist das *Was*-System, welches beide Modalitäten verwenden, dargestellt. Die Schlüsselaktivitäten der visuellen Informationsverarbeitung sind zum einen die Segregation von Bild und Hintergrund und zum anderen die Kantendetektion. Das visuelle System arbeitet dabei mit Raum- und Zeitinformationen. Im auditorischen System stehen Frequenz- und Zeitinformationen zur Verfügung. Die Kantendetektion entspricht dabei der Detektion von Tonanfang und -ende. Der Segregation von Bild und Hintergrund entspricht im auditorischen System der Stream-Segregation. Die Verbindung zwischen dem visuellen *Wo*- und dem auditorischen

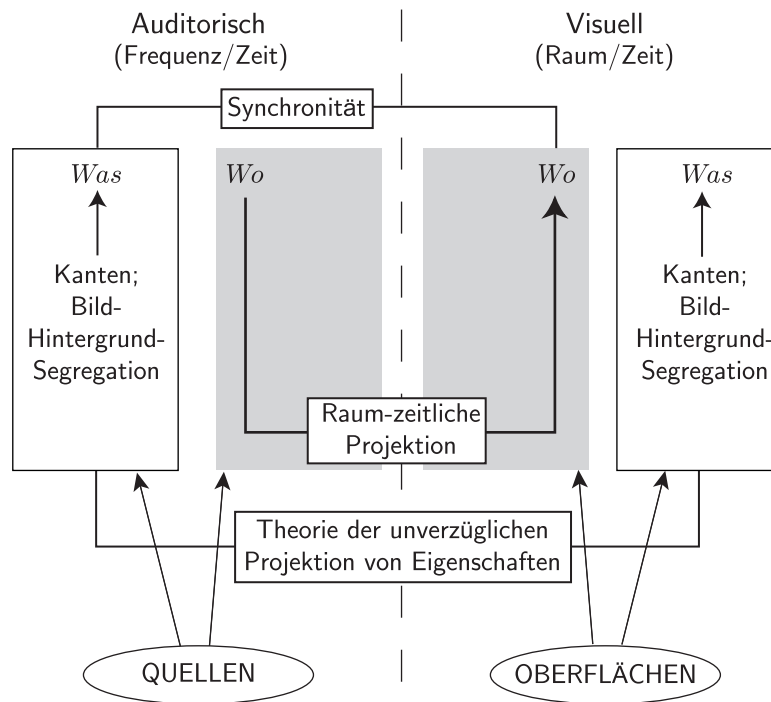


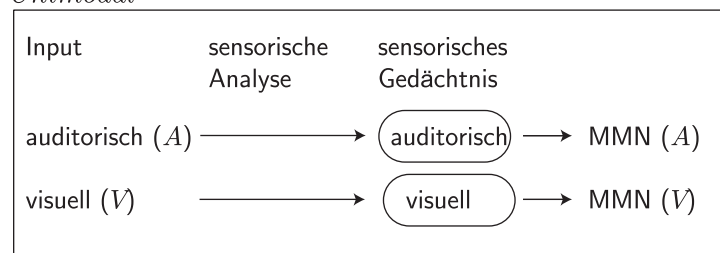
Abb. 2.8: Charakteristiken des auditorischen (links) und visuellen Systems (rechts). Beide Systeme sind stark vernetzt (nach Kubovy und van Valkenburg 2001).

Was-System kann als Erklärung für audiovisuelle Effekte wie den Bauchredner-Effekt (s. u.) dienen.

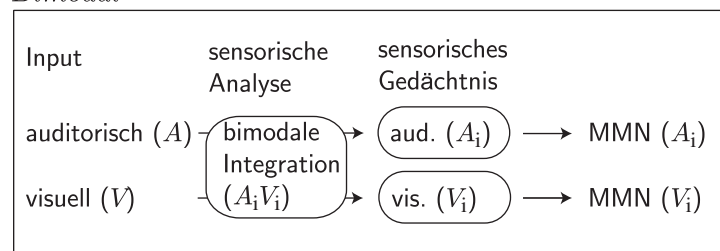
Da in dieser Arbeit die Wechselwirkung zwischen dem auditorischen und dem visuellen System untersucht wird, soll im Weiteren auf Erkenntnisse der multimodalen Verarbeitung und damit verbundene spezielle Effekte eingegangen werden. Betrachten wir zunächst die Verarbeitung von Sprache. Das Lesen von geschriebener Sprache führt sowohl zu neuronaler Aktivierung des visuellen als auch des auditorischen Kortex. Hörgeschädigte verständigen sich zusätzlich zur Sprache auch durch Gestik. Hierdurch wird ihr Verstehen von Sprache verbessert – ein Resultat audiovisueller Integration. Das Sprachverstehen kann zudem signifikant durch Beobachten der simultanen Lippenbewegung des Sprechers verbessert werden (Dodd 1977). Dass es dabei zu einer tatsächlichen Integration der audiovisuellen Informationen kommt, zeigt eindrucksvoll der McGurk-McDonald-Effekt (McGurk und McDonald 1976; Nicholls *et al.* 2004). Er weist die Veränderung auditorischer Wahrnehmung mittels synchroner visueller Reize nach. Zum Beispiel wird das auditorische Perzept der gesprochenen Silbe /ba/ plötzlich zu /ga/ oder /da/, wenn die Lippenbewegung des

Abb. 2.9: Schematisches Modell der MMN-Prozesse für unimodale und bimodale Sinnesreize (nach Besle *et al.* 2005). Wenn die auditorischen (A) und visuellen (V) Reize synchron präsentiert werden, modifizieren Interaktionen im afferenten sensorischen System die Signale, bevor sie im jeweiligen sensorischen Gedächtnis gespeichert werden. Die Prozesse der auditorischen und visuellen MMN arbeiten dann auf der Basis der veränderten Reize A_i und V_i .

Unimodal



Bimodal



Sprechers von /ba/ auf /ga/ wechselt. Dieses Phänomen ist sehr stabil, denn die meisten Personen berichten über diese Illusion, selbst wenn sie darüber aufgeklärt wurden.

Zu den zur Aufklärung audiovisueller Verarbeitungsprozesse geeigneten Werkzeugen zählt auch die MMN (vgl. Kap. 3.4). Dieses ereigniskorrelierte Potential (EKP) arbeitet nach Besle *et al.* (2005) auf der Basis der neuronalen Signale, die bereits durch bimodale Integration verändert wurden. Diese modalitätsübergreifende Interaktion (vgl. Abb. 2.9) wird auf einer Verarbeitungsstufe vor der MMN vermutet. Eine audiovisuelle Interaktion konnte auch in MEG-Messungen nachgewiesen werden, wo in Oddball-Paradigmen mit audiovisuellen phonologischen Devianten eine MMN evoziert wurde (Yumoto *et al.* 2005).

Beim McGurk-McDonald-Effekt erweist sich die MMN ebenfalls als objektive und robuste Untersuchungsmethode. Sie kann zeigen, dass dieser Effekt eine Folge von audiovisueller Integration ist (Colin *et al.* 2004). Probanden wurden Tonsilben sowohl auditorisch und visuell kombiniert als auch jeweils separat präsentiert. Zufällig eingestreut waren visuelle Deviantstimuli enthalten. Wenn die audiovisuelle Integration erst auf höheren Verarbeitungsebenen geschieht, sollte durch die von den Tonsilben abweichenden Lippenbewegungen keine auditorische MMN hervorgerufen werden. Allerdings konnte in diesem Experiment eine deutliche MMN evoziert werden, was somit auf eine frühe, basale audiovisuelle Integration hinweist. Ein neueres Experiment zeigt, dass möglicherweise die Lippenbewegungen eines Sprechers vom visuellen System extrahiert werden und als Filter auf das auditorische System projiziert werden (Gentilucci und Cattaneo 2005). Durch

MMN-Messungen ist auch aufgeklärt worden, dass bei multisensorischen Illusionen eine linkshemisphärische Dominanz entsteht (Saint-Amour *et al.* 2007).

Ein weiterer Effekt, der auf audiovisueller Integration beruht, ist die Ventriloquist-Illusion (Bauchredner-Effekt): Wenn synchron zur Sprache assoziative visuelle Informationen wie z. B. Artikulationsbewegungen aus anderer räumlicher Position wahrgenommen werden, erscheint die akustische Quelle an diesen Ort verschoben zu sein. Auch dieser Effekt beruht auf automatischen, präattentiven Mechanismen der perzeptuellen Verarbeitung (Bertelson *et al.* 2000; Colin *et al.* 2002a).

Die beschriebenen Phänomene zeigen zunächst einen Einfluss visueller Reize auf die auditorische Wahrnehmung. Aber auch der inverse Weg bimodaler Interaktion ist möglich: die Veränderung der visuellen Wahrnehmung durch auditorische Stimuli (für eine Übersicht vgl. Mitchell und Maslin 2007). Shams *et al.* (2000) präsentierten einen einzelnen visuellen Blitz in Kombination mit mehreren kurzen Tönen. Die Wahrnehmung der Anzahl der Blitze veränderte sich durch die Töne hin zu mehreren Blitzen. Diese Interaktion ist auf eine Länge von 100 ms beschränkt, was offenbar als Grenze für derartige Interaktionen betrachtet werden kann. Weiterhin kann die visuelle Leistungsfähigkeit durch störende auditorische Reize verändert werden (Guttman *et al.* 2005). Auditorische Reize wurden hier mit einem abweichenden Rhythmus präsentiert, was zur veränderten Beurteilung der visuellen Rhythmen führte.

Audiovisuelle Signaländerungen können die Reaktionszeit der Probanden gegenüber unimodalen Signaländerungen verkürzen (Schröger und Widmann 1998). Auf der Suche nach einer Erklärung, ob dies tatsächlich durch die audiovisuelle Integration geschieht oder durch Summation der auditorischen und visuellen Aktivitäten erklärt werden kann, wurden mit audiovisueller Integration verknüpfte ereigniskorrelierte Potentiale gefunden. Diese lieferten zudem die Erkenntnis, dass die Integration zwar nach der basalen (low-level) sensorischen Verarbeitung, aber weit vor der Ausführung einer motorischen Reaktionsaktivität stattfindet. Ähnliche Ergebnisse wurden auch im Kontext des McGurk-McDonald-Effekts erzielt (Colin *et al.* 2002b). Als Konsequenz wird die audiovisuelle Integration in Verarbeitungsstufen vor der aufmerksamkeitsbasierten Selektion von Silben eingeordnet (Soto-Faraco *et al.* 2004).

Eine modalitätsübergreifende Integration setzt in vielen Fällen voraus, dass die zur Verfügung stehende Aufmerksamkeit auf beide Modalitäten aufgeteilt werden kann. Dass dies möglich ist, zeigt eine Studie, bei der zwei auditorische und zwei visuelle Stimulusfolgen gleichzeitig präsentiert wurden (Duncan *et al.* 1997). Interessanterweise konnte dort die

Aufmerksamkeit weit besser zwischen zwei Modalitäten aufgeteilt werden, als auf zwei Streams innerhalb einer Modalität.

Welche kortikalen Strukturen und Neuronenpopulationen die Integration von auditorischen und visuellen Reizen durchführen, blieb lange ungeklärt. Meredith *et al.* (1987) vermuten polysensorische Neurone im Zwischenhirn (Colliculus superior), die für die Synthese von multimodalen Reizen verantwortlich sind. Diese Synthese von Informationen verschiedenen Modalitäten wird als wichtige Basis für die Steuerung von Aufmerksamkeit und Orientierung betrachtet (Wallace und Stein 1994). Durch Läsionsexperimente konnte gezeigt werden, dass der assoziative Kortex, speziell die anteriore ectosylvanische Furche, wesentliche Beiträge liefert. Auch durch fMRI-Experimente konnten neuronale Strukturen in der linken temporalen Furche nachgewiesen werden, die für bimodale Interaktionen verantwortlich sind (King und Calvert 2001).

Sams *et al.* (1991) konnten mit MEG-Messungen eine spezifische Wellenform im supratemporalen auditorischen Kortex nachweisen, die nur bei Integration auditorischer und visueller Reize auftritt. Somit könnten visuelle Informationen über artikulierende Bewegungen direkt Eingang in den auditorischen Kortex finden. Selbst Hirnstammpotentiale konnten im afferenten Pfad der neuronalen Reizleitung als Korrelate audiovisueller Integration identifiziert werden. So findet sich z. B. im Bereich von 10–20 ms nach Einsatz des Stimulus eine deutliche Negativierung bei übereinstimmenden auditorischen und visuellen Reizen (Musacchia *et al.* 2005).

Genauere Versuche, die neuronalen Quellen der audiovisuellen Integration zu lokalisieren, wurden von verschiedenen Gruppen durchgeführt. So wurde durch EEG-Messungen von Giard und Peronnet (1999) die Integration mit den daraus folgenden EKP aus dem auditorischen Kortex erklärt. Interessanterweise fanden sie aber auch Potentiale im visuellen Kortex und im rechten frontozentralen Bereich. Genauere Quellenlokalisationen konnten für diesen Bereich auf Grund der zu geringen räumlichen Auflösung der EEG-Methode nicht durchgeführt werden, eine PET-Arbeit kommt allerdings zu ähnlichen Ergebnissen (Paulesu *et al.* 1995). Die Areale im visuellen Kortex und die Aktivität auf frontozentralen sowie zentralen Bereichen der Kopfoberfläche, die zur audiovisuellen Integration beitragen, wurden auch von Molholm *et al.* (2002) durch EEG-Messungen identifiziert.

Beim Lösen von audiovisuellen Kategorisierungsaufgaben scheint der linke primäre auditorische Kortex besonders aktiv zu sein (Suchan *et al.* 2006). Probanden wurden in loser Folge Bilder gezeigt und deren Bezeichnung genannt. Stimmt ein Bild oder Begriff mit dem vorletzten Objekt in seiner Bedeutung überein, sollte dies angezeigt werden. Im fMRI

wurde immer dann im linken primären auditorischen eine erhöhte BOLD-Antwort gemessen, wenn die visuellen Reize in auditorische Kategorien transformiert werden mussten.

Durch Ableitungen an Einzelzellen wurden im superioren temporalen Sulcus Neurone gefunden, die audiovisuelle Integration reflektieren (Barraclough *et al.* 2005). Diese direkten Ableitungen liefern eine gute Möglichkeit, die Integration quantitativ auszuwerten. Es lässt sich ein Index für die Linearität der Integration I einführen, der die Antworten der Zellen (Spikes/Sekunde) bei alleiniger visueller Stimulation V oder auditorischer Stimulation A mit der audiovisuellen Stimulation AV verrechnet:

$$I = \frac{AV - (A + V)}{A + V} \quad (2.5)$$

Die Superposition von auditorischer und visueller Stimulation ergab eine Nichtlinearität ($I > 0$). Somit ist die audiovisuelle Antwort größer als die Summe der einzelnen Antworten. Diese treten als intersensorische Komponenten evozierter Potentiale in allen Bändern des EEG auf, so im ϑ -Band frontozentral parietal bei Latenzen < 100 ms. α -Band-Komponenten treten später als 200 ms nach Stimulation temporal und occipital auf (Sakowitz *et al.* 2005). Beide Komponenten reflektieren unterschiedliche Wahrnehmungsstufen (vgl. Kap. 3.3).

Abschließend sei bemerkt, dass sich durch die modalitätsübergreifende Integrationsfähigkeit die Möglichkeit ergibt, Defizite einer Modalität durch eine andere auszugleichen (Stein und Meredith 1993). Bei Patienten mit Sehstörungen kann z. B. die Erkennung der visuellen Stimuli durch die synchrone Präsentation von auditorischen Stimuli verbessert werden (Frassinetti *et al.* 2005; Rausch *et al.* 2000). Bei Hörstörungen unterstützt der Sehsinn die auditorische Wahrnehmung (Calvert *et al.* 1999; Stein 1998).

Kapitel 3

Methodik zur Messung von Streaming

Im Hinblick auf die Verbesserung der Lebensqualität von Hörgeschädigten wurden bisher verschiedene Ansätze verfolgt. In den Bereich der technischen Lösungen fallen dabei die Anpassung von externen Hörgeräten, aber auch implantierbaren Systemen wie Mittelohrprothese, Cochlea- und Hirnstammimplantat. Deren Leistungsfähigkeit und Vergleichbarkeit mit dem Hören hörgesunder Menschen könnte jedoch durch ein automatisches Erkennen der aktuellen akustischen Situation (z. B. Sprache, Sprache im Störgeräusch, Störgeräusch, Musik) noch weiter verbessert werden (Dillier 2005).

Das gesunde auditorische System verwendet verschiedene Strategien, um die Merkmale der akustischen Umwelt (u. a. spektrales Profil, Harmonizität, Onsets und Offsets, Amplituden- und Frequenzmodulationen) zu analysieren. Ein solcher Mechanismus wird im auditorischen Streaming abgebildet, das in dieser Arbeit untersucht wird. Hierbei stellt sich die Frage nach einer Messmethodik, die das Phänomen möglichst unverfälscht repräsentiert. Da beim Streaming auch kognitive Prozesse eine Rolle spielen, sind mehrere überlagerte Effekte wirksam. Um diese zu separieren, werden meist mehrere Methoden parallel verwendet.

Vorwiegend haben sich psychometrische Methoden etabliert. Diese erfordern eine aktive Beteiligung der Probanden. Zum anderen stehen aber auch objektivere elektrophysiologische Methoden zur Verfügung, bei denen evozierte elektrische Potentiale oder Magnetfelder ausgewertet werden.

3.1 Psychometrie

Psychometrische Experimente mit Streaming beruhen auf der Auswertung eines Perzepts, das entweder das Streaming-Perzept direkt, oder mit diesem korreliert ist. Dieses zeigen die Probanden entweder durch direkte Befragung oder auch indirekt an. Es haben sich einige dieser psychologischen Herangehensweisen etabliert, die von Bregman (1999, S. 55–57) zusammengefasst wurden und hier als Auswahl skizziert werden sollen.

- *Ausjustierung.* Von den Probanden wird gefordert, einen physikalischen Parameter während des Experiments kontinuierlich in Abhängigkeit von ihrem auditorischen Perzept zu verändern. Daraus werden die Umkehrpunkte ermittelt, ab dem sie zwei Streams hören oder – in der anderen Richtung – sich die Töne zu einem Stream vereinigen. Aus dem Mittelwert beider Punkte kann eine Schwelle ermittelt werden. Problematisch hierbei ist jedoch die Hysterese von perzeptueller Organisation, wenn die Schwellen für die jeweilige Perzeptänderung asymmetrisch sind.
- *Schwellenbestimmung.* Bei dieser Methode werden die Parameter vom Experimentator systematisch verändert und, ähnlich wie bei der vorigen Methode, Schwellen ermittelt. Der Proband hat hier nur die Aufgabe, die Perzeptänderung anzuzeigen. Auch bei dabei ist das Problem der Hysterese zu berücksichtigen.
- *Verhältnis der Integrations- und Segregationzeiten.* Hierbei sollen Probanden eine Taste gedrückt halten, solange sie einen integrierten Stream wahrnehmen und eine andere Taste drücken, wenn mehrere Streams zu hören sind. Das Verhältnis der Zeiten für das jeweilige Perzept liefert ein Maß für die Anzahl der wahrgenommenen Streams. Nachteile bei dieser Methode ergeben sich bei Stimulationsbedingungen, die je nach Aufmerksamkeit des Probanden multiple Perzepte zulassen.
- *Bewertung.* Die Probanden sollen verschiedene Stimulationsbedingungen auf einer Skala mit den Extrempercepten *Integration* und *Segregation* bewerten. Da Probanden bei verschiedenen auditorischen Szenen zur Variation der Skala neigen, ist diese Methode zum Vergleich verschiedener Experimente wenig geeignet.
- *Mustererkennung.* In Stimulationssequenzen werden Target-Muster eingefügt, die von den Probanden zu detektieren sind. Die Sensitivität d' als ein Maß für die Detektionsleistung wird ausgewertet. Hierbei muss sichergestellt werden, dass die gemessene Sensitivität nicht spezifisch für das Streaming ist, sondern auch die Mustererkennung beinhaltet.

Bei der psychoakustischen Beurteilung von audiovisuellen Szenen hängt das Ergebnis vielfach von der Aufmerksamkeit der Probanden auf die Stimuli ab. Bei multisensorischen Interaktionen, die auf frühen Verarbeitungsstufen ablaufen, ist diese Aufmerksamkeitsabhängigkeit jedoch weniger ausgeprägt: Sowohl der Bauchredner-Effekt (Vroomen *et al.* 2001) als auch die Wahrnehmung eines kurzen visuellen Stimulus als Blitzfolge durch entsprechende synchrone auditorische Stimulation (Watkins *et al.* 2006) sind von der Fokussierung der Aufmerksamkeit unabhängig.

In dem psychoakustischen Experiment dieser Arbeit wird die Methode *Verhältnis der Integrations- und Segregationszeiten* verwendet. Zur Kontrolle der Aufmerksamkeit eignet sich die Methode der *Mustererkennung*. Sie ist Bestandteil der elektrophysiologischen Experimente.

3.2 Messung von Hirnaktivität

3.2.1 Elektroenzephalographie

Elektrophysiologische Messungen sind sowohl in der Forschung als auch in der Klinik etablierte Verfahren zur Untersuchung des Hörsystems. Sie beruhen auf der Erfassung elektrischer Aktivität von Nervenzellen (Neuronen). Diese bestehen aus einem Zellkörper (Soma), Verzweigungen des Somas (Dendriten) und dem Axon mit Synapsen als Verbindung zu anderen Nervenzellen. Seinen Input erhält das Neuron chemisch über Transmittermoleküle an den Dendriten, die die Informationen von verschiedenen anderen Neuronen sammeln. Es integriert diese zu einem Potentialsignal und leitet es zum Axon, wo die Informationen über Synapsen an Dendriten anderer Neurone weitergegeben werden (Rubinstein 2004). Wird der erregende Neurotransmitter (z. B. Glutaminsäure) in den synaptischen Spalt ausgeschüttet, führt dies zu einem exzitatorischen postsynaptischen Potential (EPSP) beziehungsweise inhibitorischen postsynaptischen Potential (IPSP) an der postsynaptischen Membran. Im exzitatorischen Fall bewirken die in den Dendriten strömenden positiven Ladungsträger extrazellulär ein negatives Potential. Die eingeströmten positiven Ladungsträger fließen zum Soma und treten dort, und auch schon auf dem Weg dorthin, wieder in den Extrazellulärraum über. Somit entsteht außerhalb des Dendriten am Soma und in dessen Nähe eine Stromquelle. Durch die extrazelluläre Negativität (Senke) und Positivität (Quelle) kommt es zu einem extrazellulären Stromfluss, der den Stromkreis schließt (Breckow und Greinert 1994; Hoppe *et al.* 1982).

Durch Potentialdifferenzen an der Membran, die durch Konzentrationsunterschiede der K^+ - und Na^+ -Ionen innerhalb und außerhalb der Zelle (des Neurons) gebildet werden, entsteht ein elektrisches Feld. Typischerweise ist die Kaliumkonzentration intrazellulär etwa 30 mal so groß wie extrazellulär. Dagegen ist die Natriumkonzentration extrazellulär etwa 10 mal so hoch wie intrazellulär. Die Verteilung der Cl^- -Konzentration ist der der K^+ -Konzentration reziprok.

Unter Einbeziehung der Permeabilitäten g der Zellmembran für die unterschiedlichen Ionenarten und die Ionenbeweglichkeiten erhält man die Goldman-Gleichung für das Ruhemembranpotential:

$$E_R = \frac{RT}{F} \ln \frac{g_K[K^+]_e + g_{Na}[Na^+]_e + g_{Cl}[Cl^-]_i}{g_K[K^+]_i + g_{Na}[Na^+]_i + g_{Cl}[Cl^-]_e} \quad (3.1)$$

Es beträgt für Nerven- und Muskelzellen bei Verwendung des Extrazellulärtraumes als Bezugspotential:

$$E_R = -80 \dots -90 \text{ mV} \quad (3.2)$$

Der durch die Permeabilität der Membran für Ionen verursachte Potentialabbau wird durch einen aktiven Ionentransport mittels in der Zellmembran befindlicher Ionenpumpen ausgeglichen.

Signale im Sinne einer Information werden erst durch zeitlich veränderte Membranpotentiale erzeugt. Durch ein Ersatzschaltbild lassen sich diese modellieren und mit der Kabelgleichung beschreiben (Rall 1977):

$$\frac{1}{R_e + R_i} \cdot \frac{d^2 U_m}{dx^2} = C_m \frac{dU_m}{dt} + \frac{U_m - E_R}{R_m} \quad (3.3)$$

R_e und R_i sind die Widerstände entlang der Membranoberflächen, C_m die Kapazität über die Membran und U_m die Membranspannung (Signal). Die Leitfähigkeiten g werden hier durch Transmembranwiderstände R_m dargestellt und als konstant angenommen.

Durch numerisches Lösen von Gl. (3.3) findet man, dass sich die Potentialänderung $U_m - E_R$ räumlich und zeitlich exponentiell abklingend ausbreitet (vgl. Crank und Nicolson 1996). Diese Eigenschaften sind für den Transport von Signalen entlang von Neuronen essentiell. Um ein unabgeschwächtes Signal weiterzuleiten, bedienen sich die Neurone zusätzlich eines anderen Prinzips: Sie sind entlang der Axone erregbar. Erregbarkeit bedeutet hierbei, dass die Leitfähigkeiten entlang der Axonmembran spannungs- und zeitabhängig sind und das System als aktiver Verstärker wirkt.

In Erweiterung zu Gl. (3.3) kann der Spannungsverlauf an einer Membran durch die Hodgkin-Huxley-Gleichung beschrieben werden:

$$\frac{1}{R_e + R_i} \cdot \frac{d^2U}{dx^2} = C_m \frac{dU}{dt} + U \cdot g_K(U,t) + U \cdot g_{Na}(U,t) \quad (3.4)$$

Sie beschreibt den zeitlichen und räumlichen Verlauf der Spannung unter Einbeziehung spannungs- und zeit- sowie von der Art der Ionen abhängiger Membranpermeabilitäten.

Will man kortikale Aktivität auf der Kopfhaut erfassen, muss das Signal stark sein, da es nur als Fernfeld an der Kopfhaut gemessen werden kann. Grundlage dafür ist die palisadenförmige Anordnung der im gesamten Kortex vorhandenen Pyramidenzellen, deren Dendritenbäume senkrecht zur Oberfläche stehen. Wenn viele Pyramidenzellen (als Neuronenverbund) in gleicher Weise erregt werden, sind die entstehenden Dipole gleichgerichtet. Dann entsteht ein großes Signal, das an der Kopfhaut als Spannung gemessen werden kann.

Die Registrierung erfolgt über ein Elektroenzephalogramm (EEG), indem mindestens drei Elektroden (Masse-, Referenz- und Messelektrode) an der Kopfhaut angebracht werden. Mit einem Differenzverstärker wird die Spannung zwischen Referenz- und Messelektrode gemessen, die im Bereich einiger Mikrovolt liegt. Üblicherweise wird das EEG durch Verwendung vieler (z. B. 32), auf dem Kopf verteilter Elektroden registriert, was bei der späteren Analyse Rückschlüsse auf die Lage der Neuronenverbände zulässt. Die Referenzelektrode wird oft an der Nase, aber auch an den Mastoiden angelegt.

3.2.2 Magnetenzephalographie

Die in den Neuronen fließenden Ionenströme verursachen neben dem elektrischen Feld auch ein schwaches Magnetfeld. Durch die säulenförmige Anordnung der Neurone in vielen kortikalen Strukturen (Pyramidenzellen) sind die Ionenströme synchron und gleichgerichtet. Ihr Magnetfeld überlagert sich derart, dass es auch außerhalb des Schädels gemessen werden kann. Bei solchen biomagnetischen Messungen (Magnetenzephalographie, MEG) findet man, dass etwa 10^5 – 10^6 Neurone ein im MEG messbares magnetisches Feld von etwa 50–500 fT erzielen. Damit ist dieses Feld um neun Größenordnungen kleiner als das Erdmagnetfeld (20–30 μ T). In der Elektrophysiologie wird das an der Kopfhaut gemessene Potential durch die elektrische Leitfähigkeit kortikaler Strukturen kaum abgeschwächt. Im Gegensatz dazu nimmt das Magnetfeld eines Dipols mit dem Quadrat des Abstands ab ($B_C \sim 1/r^2$). Das messbare Magnetfeld von im Inneren des Kortex liegenden Strukturen ist daher schwach (Hämäläinen *et al.* 1993).

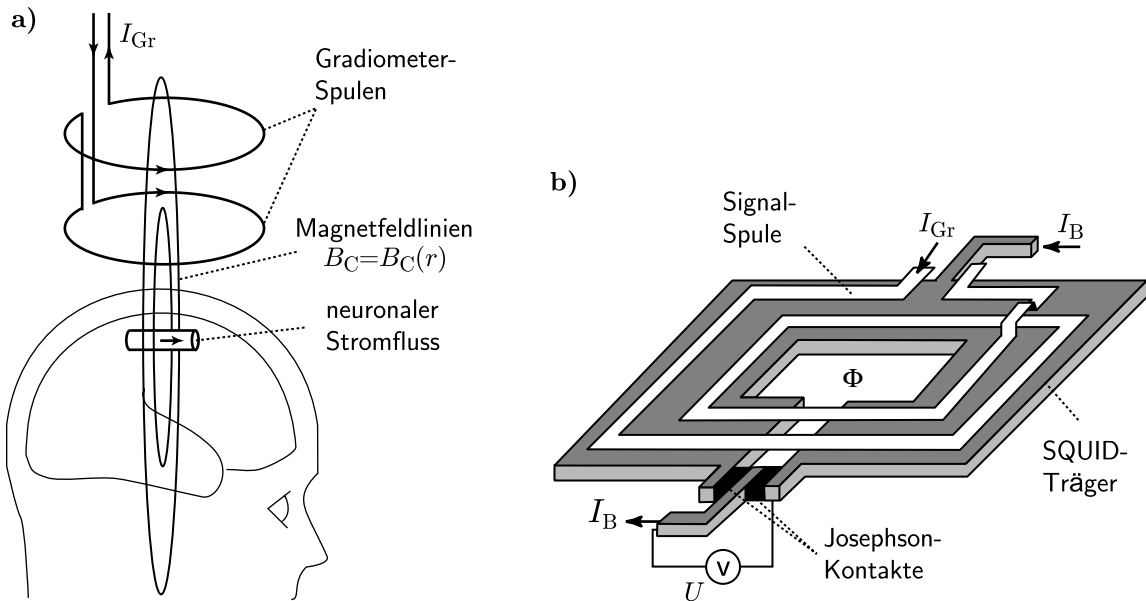


Abb. 3.1: Schematische Funktionsweise einer Gradiometerspule (nicht maßstabsgetreu). **a)** Durch einen neuronalen Stromfluss entsteht auch außerhalb des Schädels ein Magnetfeld B_C , das als Differenzfeld in den Gradiometerspulen einen Stromfluss I_{Gr} induziert. **b)** Dieser wird zur isoliert auf dem SQUID befindlichen Signalspule geleitet und erzeugt dort einen magnetischen Fluss Φ , der im SQUID-Träger wiederum einen Stromfluss induziert. Der SQUID-Träger wird durch einen Bias-Strom I_B auf den Arbeitspunkt gebracht, um über den Josephson-Kontakten eine Gleichspannung U messen zu können. Sowohl die Gradiometerspulen als auch das SQUID werden gekühlt und befinden sich im supraleitenden Zustand.

MEG-Messungen werden in abgeschirmten Kabinen durchgeführt, um den Einfluss äußerer Magnetfelder (z. B. vom Stromnetz, Erdmagnetfeld) gering zu halten. In der Nähe des Kopfes befinden sich in einer helmartigen Anordnung Spulen, in denen durch das kortexnahe Magnetfeld Ströme induziert werden (vgl. Abb. 3.1). Dabei handelt es sich meist um Gradiometer, die aus zwei gegenläufigen, in unterschiedlichem Abstand vom Kopf angeordneten Spulen bestehen. Hierdurch wird der Einfluss von Magnetfeldern, deren Ursprung sich in einer relativ zum Spulenabstand großen Distanz befindet, reduziert und im Wesentlichen der Gradient der im Kortex generierten Felder B_C als Differenzstrom I_{Gr} erfasst. Zusätzlich werden Referenzdetektoren, die weiter vom Kopf entfernt sind, eingesetzt, um den tieffrequenten Anteil der Störfelder (von bewegten Störquellen wie z. B. Fahrstühlen) von den gemessenen Feldern zu subtrahieren. Die Gradiometerspulen sind mit den SQUIDS (Superconducting QUantum Interference Device, s. u.) verbunden, um die mit dem Magnetfeld korrelierte Spannung U im Zeitverlauf registrieren zu können (vgl. König *et al.* 2007).

Zunächst sei angenommen, dass die magnetische Permeabilität des Kopfes μ kaum variiert. Betrachtet man das gemessene Feld als lineare Überlagerung der Dipolfelder verschiedener Neurone, kann das an einem Punkt r gemessene Feld $B(r)$ außerhalb eines Volumens V als Integral über dieses Volumen beschrieben werden (Biot-Savart'sches Gesetz):

$$B_C(r) = \frac{\mu}{4\pi} \int_V \frac{j(r') \times (r - r')}{|r - r'|^3} dV \quad (3.5)$$

r' bezeichnet den Ort der Quelle, j die Gesamtstromdichte.

$$j = j_p + j_v \quad (3.6)$$

Hier ist j_p die intrazelluläre (primäre) Stromdichte, die in direkter Beziehung zum physiologischen Prozess steht und die biologische Aktivität der Quelle beschreibt. Die makroskopische (passive) Stromdichte j_v kennzeichnet die durch das umgebende Gewebe hervorgerufene biologische Aktivität.

Bei der Verwendung von Gradiometerspulen sind die durch elektrische Dipole generierten Magnetfelder nicht vollständig messbar. Es können, unter der Annahme einer homogenen Kopfgeometrie, nur die Normalkomponenten von $B(r)$ erfasst werden, was besonders bei radial orientierten Dipolen problematisch ist.

Durch die geringe Feldstärke des außerhalb des Kopfes registrierbaren Magnetfeldes ergibt sich die Notwendigkeit für besonders empfindliche Detektionssysteme. Ein solches

Messsystem für den magnetischen Fluss ist das SQUID (Superconducting QUantum Interference Device), das den Josephson-Effekt nutzt. Dieser quantenmechanische Effekt beschreibt das Tunneln von Cooper-Paaren durch eine dünne Isolationsschicht zwischen zwei supraleitenden Kontakten. Bis zu 300 von ihnen erfassen, durch flüssiges Helium gekühlt, in modernen Systemen die Magnetfeldverteilung über der Kopfhaut von Probanden. Dazu befinden sie sich in einer helmartigen Anordnung in möglichst geringem Abstand von der Kopfhaut der Probanden (Wikswow Jr. 1995).

Die in den Gradiometerspulen induzierten Ströme werden zu den SQUIDs geleitet (vgl. Abb. 3.1). Deren Träger besteht oft aus Niob- (Nb) oder Niobnitrid (NbN) und ist meist an zwei Stellen durch eine wenige Nanometer dicke Isolationsschicht aus Magnesium- oder Aluminiumoxid (MgO , Al_2O_3) unterbrochen (Josephson-Kontakte). Auf den SQUID-Träger ist die Signalspule aufgebracht, die von der Gradiometerspule gespeist wird. Der SQUID-Träger wird mit einem Bias-Strom auf seinen Arbeitspunkt gebracht, um an den Josephson-Kontakten einen Gleichstrom zu erzeugen, der über diesen als Spannungsabfall messbar ist. Wird nun durch die Signalspule ein magnetischer Fluss erzeugt, ändert sich der Strom in beiden Teilen des SQUID-Trägers, was sich in einer Spannungsänderung über den Josephson-Kontakten abbildet. Dieser Spannungsverlauf wird verstärkt und als MEG-Signal gespeichert.

Mittels verschiedener Annahmen können durch das MEG-Signal die generierenden Quellen lokalisiert werden. Das einfachste physiologisch bedeutende Modell ist das Modell der äquivalenten Dipole. Hierbei wird angenommen, dass die räumliche Ausdehnung des Dipols wesentlich kleiner als sein Abstand zum Detektor ist. Damit kann dieser als punktförmig betrachtete Dipol Q durch die Dirac'sche Deltafunktion an dem Ort r_Q beschrieben werden:

$$j_p = Q\delta(r - r_Q) \quad (3.7)$$

Eine postulierte Anzahl derartiger Dipole dient Algorithmen zur Bestimmung von deren Lage und Orientierung als Basis. Die Parameter werden dabei iterativ verändert, bis die Dipole den Zeitverlauf des MEG-Signals bestmöglich reproduzieren (Sarvas 1987; Scherg 1990).

Die MEG eignet sich durch ihre hohe räumliche und zeitliche Auflösung für die Analyse evozierter Felder und zur Lokalisation deren schädelnaher kortikaler Quellen. Die zeitliche Auflösung von MEG und EEG ist gleich, die räumliche Auflösung des MEG ist dem EEG jedoch überlegen, da beim EEG die Volumenleitung von Strömen nicht vernachlässigt

werden kann (Pantev *et al.* 1988; Pantev und Lütkenhöner 2000; Vrba und Robinson 2001; König *et al.* 2007).

3.3 Evozierte Potentiale und Felder

Der humane Kortex ist anatomisch so organisiert, dass bestimmte Areale vor allem nur für eine Sinnesmodalität verantwortlich sind (funktionelle Organisation). Im auditorischen Kortex nahe der Sylvischen Furche ist die Verarbeitung von Schallreizen, und im occipitalen Bereich des Kortex (visueller Kortex) die des visuellen Inputs lokalisiert (Shepherd 1993). Weil jeweils eine Gruppe gleich ausgerichteter Pyramidenzellen in gleicher Weise erregt wird, kann deren Funktion durch gezielte Stimulation als evoziertes Potential im EEG und als evoziertes Feld im MEG gemessen werden. Die folgenden Ausführungen beziehen sich auf evozierte Potentiale, gelten aber in ähnlicher Weise auch für evozierte Felder (Hari *et al.* 1992).

In dieser Arbeit sollen nur die transienten evozierten Potentiale betrachtet werden. Sie lassen sich verstehen als zeitlich begrenzte Reaktion auf einen Reiz. Im Gegensatz zum unvermeidlich zusammen mit den evozierten Potentialen gemessenen Ruhe-EEG, haben sie eine zeitliche Beziehung zum akustischen Reiz. Ihre Amplitude ist meist kleiner als die Amplitude des EEG. Um sie dennoch darstellen zu können, wird das EEG in viele Segmente unterteilt, die jeweils den gleichen zeitlichen Bezug zum dargebotenen Reiz haben. Die so entstandenen Potentiale werden gemittelt. Durch diese reizkorrelierte Mittlung des EEG werden die nicht mit dem Reiz korrelierten Bestandteile (Rauschen) unterdrückt, und die evozierten Potentiale können isoliert werden.

Auditorisch evozierte Potentiale (AEP) entstehen auf allen Stufen der Hörbahn, von der Cochlea über den Hirnstamm bis hin zum auditorischen Kortex (Moore 1983). Potentiale mit kleiner Latenz werden ersten Verarbeitungsstufen auf der Hörbahn zugeschrieben, bei Potentialen großer Latenz geht man von Entstehungsorten der hohen Verarbeitungsstufen, wie z. B. dem auditorischen Kortex, aus. Demzufolge wird auch die Klassifizierung der AEP nach ihrer Latenz vorgenommen (vgl. Abb. 3.2). Frühe akustisch evozierte Potentiale (FAEP, kleine Latenz) werden im Hirnstamm generiert. Die mittellatenten Potentiale (MAEP) entstehen im Mittelhirn, Thalamus und zum Teil auch im primären auditorischen Kortex. Späte Potentiale (SAEP, große Latenz), entstammen Generatoren im auditorischen Kortex. Eine Übersicht über evozierte Potentiale findet sich bei Gelfand (1998). In dieser Arbeit werden nur die kortikalen SAEP betrachtet.

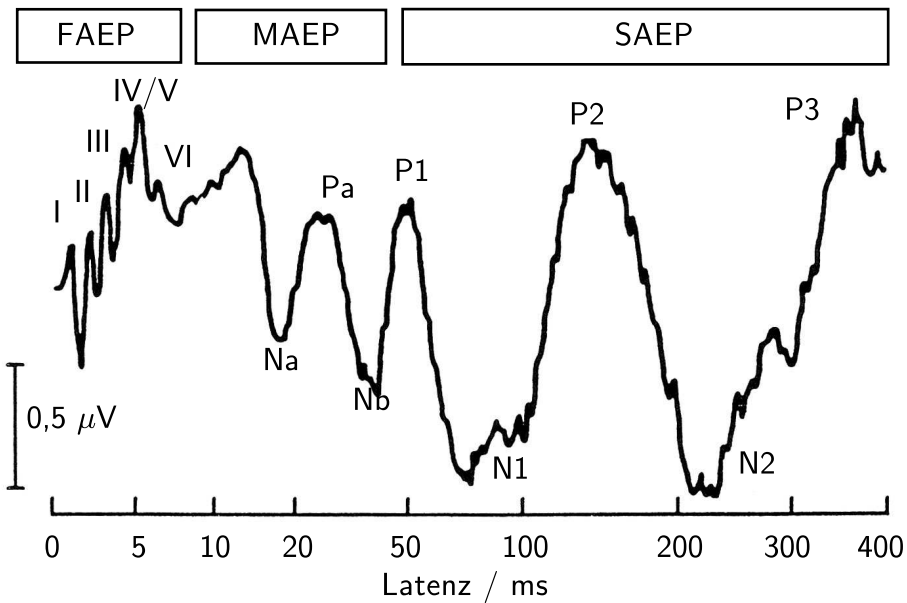


Abb. 3.2: Frühe, mittellatente und späte auditorisch evozierte Potentiale (FAEP, MAEP, SAEP). Die Abszisse ist logarithmisch geteilt (nach Gelfand 1998).

SAEP zeigen oft eine starke Abhängigkeit vom Aufmerksamkeitszustand des Probanden. Ihre Extrema werden, ihrer Polarität entsprechend, mit P1 bei einer Peak-Latenz von etwa 60 ms, N1 bei 100 ms, P2 bei 160 ms, N2 bei 250 ms und P3 bei 250–500 ms bezeichnet. Zur Erzeugung von SAEP können z. B. Tonpulse einer bestimmten Frequenz verwendet werden.

Potentiale, die überwiegend von den physikalischen Reizparametern abhängen, werden als *exogene* Potentiale bezeichnet. Im Gegensatz dazu stehen die *endogenen* Potentiale, die vor allem von deren psychologisch-subjektiver Einordnung abhängen. Im Latenzbereich der SAEP auftretende Potentiale sind, mit steigender Latenz, zunehmend kognitiver Natur (endogene Potentiale). Deren Verlauf hängt davon ab, inwieweit der Proband den Reizen bestimmte Bedeutungen zuordnet (Katayama und Polich 1999; Altenmüller 1996).

Der Potentialkomplex N1-P2 (SAEP) stellt die automatische Antwort auf die Wahrnehmung eines akustischen Reizes dar. Einem negativen N1-Peak mit einer Latenz von 100 ms folgt meist ein positiver P2-Peak bei ungefähr 175 ms. Ausgelöst werden kann der N1-P2-Komplex von allen üblichen Stimuli: Klicks, Bursts, Tönen und Sprachelementen. Außerdem entsteht er bei Änderungen der Lautstärke oder Frequenz (Näätänen und Picton 1987). Der N1-P2-Komplex wird Generatoren im sekundären auditorischen Kortex zugeschrieben, die für N1 und P2 verschieden sind (Eggermont und Ponton 2002).

Generell nimmt die Amplitude des N1-P2-Komplexes mit dem Reizpegel zu. Zudem hängt sie von der Reizrate, also dem Abstand zweier aufeinanderfolgender Stimuli (SOA) ab. Mit kleiner werdendem SOA nimmt auch die N1-P2-Amplitude ab und kann mit kurzen

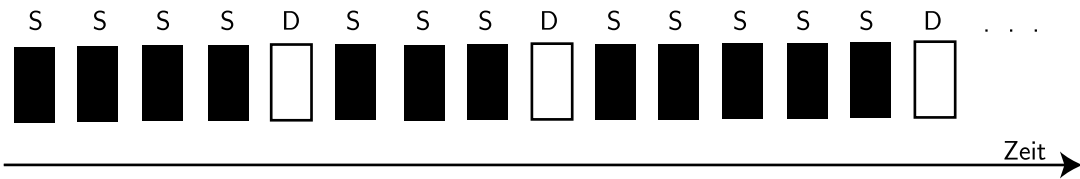


Abb. 3.3: Schematische Darstellung eines Oddball-Paradigmas mit häufigen Standard- (S) und seltenen Deviantreizen (D).

SOA fast vollständig unterdrückt werden, wobei sich die Latenzen nur geringfügig ändern. Diese Abhängigkeit kann durch die bei kurzen SOA stärkere Habituation an die Reizfolge erklärt werden: Neuronen adaptieren auf einen wiederholt dargebotenen Reiz und die ursprüngliche Reaktion auf diesen Reiz nimmt ab. Der N1-P2-Komplex ist vigilanzabhängig, wird aber dennoch automatisch generiert. Eine Bedeutungszumessung durch den Probanden ist hierbei nicht erforderlich (Hyde 1997; Näätänen und Picton 1987).

Ein weiteres, ereigniskorreliertes SAEP ist die P3 mit einer Latenz von 250–500 ms. Es hat relativ große Amplituden, so dass es schon nach wenigen Mittelungen nachgewiesen werden kann. Seine Entstehung ist mit kognitiven Prozessen, die mit den Stimuli zusammenhängen, verbunden. So können gleiche Stimuli zu einer P3 führen oder auch nicht, je nachdem, ob der Stimulus eine Relevanz für den Probanden besitzt. Diese Abhängigkeit von der Bedeutungszumessung des Probanden, aber auch die Vigilanzabhängigkeit unterscheidet die P3 von der automatischen N1 und ist somit ein endogenes EKP.

Um ein P3-Potential zu messen, werden meist so genannte *Oddball*-Paradigmen verwendet (vgl. Abb. 3.3). Ein solches Paradigma enthält zwei oder mehr verschiedene Stimuli, wovon der eine häufig dargeboten wird (Standard) und der andere mit einer geringeren Wahrscheinlichkeit zufällig eingestreut wird (Deviant). Standard- und Deviantstimuli können sich beispielsweise in der Frequenz oder Lautstärke unterscheiden. Der Proband wird aufgefordert, die in dieser Serie gleicher Stimuli auftretenden Deviantreize zu registrieren. Dabei soll er sich auf die Stimuli konzentrieren und ihnen eine Bedeutung beimessen. Dies kann gewährleistet sein, indem er die seltenen Stimuli durch Tastendruck anzeigt, oder diese zählt (Jacobson 1994).

Die durch die Standardreize ausgelösten Potentiale werden getrennt von den durch die Deviantreize ausgelösten gemittelt. Dabei zeigt sich, dass die P3-Komponente nur in letzteren enthalten ist. Wenn Standard- und Deviantreiz schwer zu unterscheiden sind, ist es auch möglich, dass der Proband einem Standardreiz irrtümlich die Bedeutung eines Devi-

antreizes zumisst. In diesem Fall wird die P3-Komponente zusätzlich durch die fälschlich als Deviantreiz registrierten Standardreize evoziert (Salisbury *et al.* 2001).

Aus der P3 lassen sich verschiedene Unterkomponenten separieren. So zeigt die P3a vermutlich die automatische Orientierungsreaktion gegenüber den Stimuli an, die spätere P3b die eigentliche Reizklassifikation und die nachfolgende *slow wave* die weitergehende Reizverarbeitung (Jacobson 1994; Salisbury *et al.* 2001).

3.4 Mismatch-Negativity

3.4.1 Grundlegendes

In alltäglichen Hörsituationen ändert sich der auditorische Input ständig. Diese Veränderungen zu detektieren, ohne permanent höchste Aufmerksamkeit auf die auditorische Szene zu richten, kann für den Hörer wichtig sein. Die Veränderungen enthalten z. B. Informationen darüber, ob eine neue Schallquelle aktiv ist oder verstummt. Mit der neuronalen Detektion solcher Veränderungen ist die Mismatch-Negativity (MMN) korreliert, die als ein ereigniskorreliertes Potential 1978 von Näätänen, Gaillard und Mäntysalo beschrieben wurde (Näätänen *et al.* 1978). Sie wird durch jegliche unterscheidbare Veränderungen in einer regelmäßigen Folge auditorischer Stimuli evoziert und kann auch ohne Aufmerksamkeit der Probanden auf die Stimuli abgeleitet werden (Giard *et al.* 1995). Als negative Komponente der AEP tritt ihr Maximum im Latenzbereich von 100–200 ms nach der Veränderung auf (Sams *et al.* 1985). Durch ein Experiment mit Kontext-Änderungen von auditorischen Stimuli wurde ebenfalls gezeigt, dass die MMN an der Detektion solcher Änderungen beteiligt ist (Sussman *et al.* 2001). Auch die P3a ist ein Korrelat für die Aufmerksamkeitsänderung, die durch plötzliche Veränderungen in der akustischen Umwelt generiert wird.

Eine MMN kann, wie die P3, durch ein Oddball-Paradigma (vgl. Abb. 3.3) evoziert werden. Die durch die häufigen Standard- und seltenen Deviantreize evozierten Potentiale werden separat gemittelt und die MMN in der Differenzkurve zwischen diesen beiden Potentialen ausgewertet. Das biomagnetische Pendant zur MMN ist das Mismatch-Field (MMF), das in MEG-Messungen analysiert werden kann.

Die Abweichungen in einer Stimulussequenz, die eine MMN auslösen können, sind vielfältig. So sind MMN als Folge von Frequenz-, Dauer- und Intensitätsänderungen bekannt

(Tervaniemi *et al.* 1999). Auch eine veränderte räumliche Lage oder phonetische Unterschiede können eine MMN hervorrufen (Näätänen und Winkler 1999). Zudem sind komplexere Paradigmen bekannt, die unter Variation mehrerer Stimulusparameter ebenfalls zur Evozierung einer MMN führen (Schröger 2005; Picton *et al.* 2000). Schröger *et al.* (1994) konnten eine MMN evozieren, indem sie die Position eines Tones in einer schnellen Sequenz verändert haben. Sie kann auch durch zufälliges Auslassen eines Stimulus, Umkehr der Frequenzmodulationsrichtung (Sams und Näätänen 1991; Pardo und Sams 1993; Paavilainen *et al.* 1995; Rahne *et al.* 2005) oder Veränderungen in der Reihenfolge von Stimuli generiert werden (Sussman *et al.* 1999). Zudem lässt sich eine MMN in einer alternierenden Tonsequenz durch Tonwiederholungen evozieren (Alain und Woods 1997).

Abb. 3.4 zeigt eine Modellvorstellung der Auslösung der MMN. Sie basiert auf so genannten Spuren im auditiven sensorischen Gedächtnis, die die regelmäßig wiederkehrenden Eigenschaften des auditorischen Inputs repräsentieren. Dieser Speicher arbeitet noch weitgehend vor der Kategorisierung, also der bewussten Bedeutungszumessung der Stimuli, und gliedert sich in zwei Typen: Der kürzere mit einer Dauer von etwa 200 ms ist für die unmittelbare Wahrnehmung des Schalls relevant. Der längere dieser Speicher, der eine Dauer von bis zu 20 s hat, stellt eine statische Form der Speicherung dar. Er ist auch unter dem Begriff *Echogedächtnis* bekannt (Cowan 1984; Giard *et al.* 1995). Er speichert die Merkmale eines Geräuschmusters als neuronale Repräsentation R' . Auf dieser Basis können die sich wiederholenden Aspekte im auditorischen Kortex modelliert werden (Näätänen und Winkler 1999). Neue Stimulusrepräsentationen R werden mit dieser Spur verglichen (Komparator) und bei Abweichung (Devianten, vgl. Abb. 3.3) wird eine MMN generiert (vgl. Abb. 3.4). Dieser Mechanismus wird automatisch – also ohne bewusste Konzentration auf die Stimuli – ausgelöst und setzt die Funktion basaler Gedächtnisfunktionen voraus (Jacobsen und Schröger 2001, 2003; Schröger 2004; Picton *et al.* 2000).

Obwohl die MMN als ein automatischer Detektionsmechanismus betrachtet wird, kann sie durch Aufmerksamkeitseinflüsse verändert werden (Näätänen *et al.* 1993). Eine signifikante Verstärkung ihrer Amplitude gegenüber dem passiven Hören tritt auf, wenn die Aufmerksamkeit der Probanden auf die Detektion der Deviantreize gerichtet ist (Woldorff *et al.* 1991). Trotzdem spiegelt die MMN automatische, präattentive Prozesse wider, denn auch Prozessen, die für veränderte Aufmerksamkeitszustände empfindlich sind, kann nicht *per se* ihre präattentive Entstehung ausgeschlossen werden (Schneider und Shiffrin 1977).

Um auszuschließen, dass es sich bei der MMN um ein durch selten eingestreute Deviantstimuli auftretendes N1-Potential mit größerer Amplitude handelt, wurden verschiedene

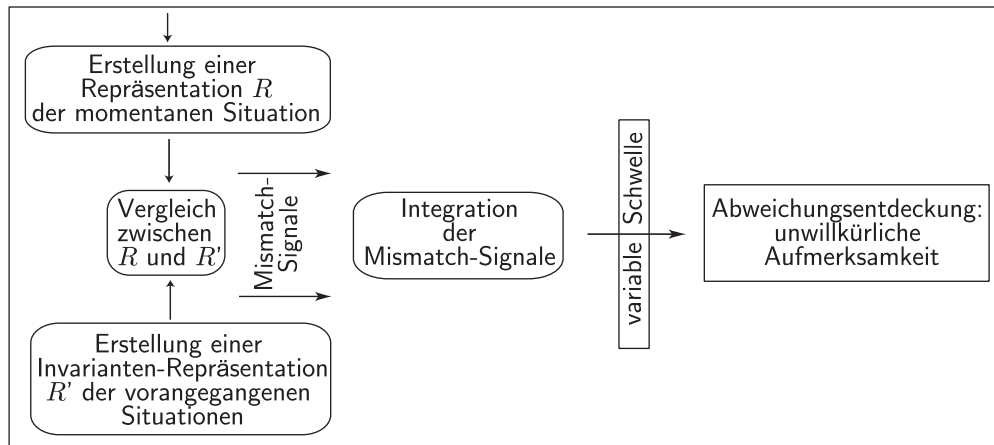


Abb. 3.4: Modell zum MMN-Auslösemechanismus. Das zentrale Hörsystem erzeugt von jedem eintreffenden Schall eine neuronale Repräsentation R und bildet Regularitäten vorangegangener Stimulationen in einer Invariantenrepräsentation R' ab. Diese Repräsentationen werden verglichen und bei Nichtübereinstimmung Mismatch-Signale generiert. Deren Integration muss eine Schwelle übersteigen, um eine MMN auszulösen (nach Schröger 1997).

Habituations-Situationen untersucht (Jacobsen und Schröger 2003). Dabei wurde durch die seltenen Deviantstimuli keine MMN evoziert, wenn sich keine neuronale Repräsentation der Standardstimuli bilden konnte. Somit handelt es sich bei der MMN in der Tat um ein Korrelat der Abweichungsdetektion als Folge eines Vergleichsprozesses im sensorischen Gedächtnis.

3.4.2 Neuronale Generatoren der MMN

Die Strukturen im Kortex, die eine MMN generieren, sind derzeit noch nicht vollständig aufgeklärt (Saint-Amour *et al.* 2007). EEG-Experimente mit Vielkanal-Ableitungen ermöglichen eine Dipolanalyse auf Grundlage der ereigniskorrelierten Potentiale (Scherg *et al.* 1989). Hierbei werden die Dipolquellen im Kortex ermittelt, die das gemessene Potential erklären. In Kombination mit der Visualisierung der Quellenstromdichte (Source Current Density, SCD) können Rückschlüsse auf die Generatoren der MMN gezogen werden (Giard *et al.* 1990). Entsprechende Regionen wurden z. B. bilateral im supratemporalen Planum (im Bereich des auditorischen Kortex) lokalisiert (Alho 1995). Dipolanalysen konnten den verschiedenen Devianttypen, die eine MMN auslösen (Dauer, Frequenz und Intensität) ebenfalls unterschiedliche aktive Regionen im auditorischen Kortex zuord-

nen (Rosburg 2003). Auch bei Verwendung eines Paradigmas mit Sprachreizen (Silben) wurden bilaterale MMN-Quellen im auditorischen Kortex gefunden (Saint-Amour *et al.* 2007).

Vielfach wird zudem über eine Polarisationsinversion als Charakteristikum für die MMN berichtet: An den Mastoiden wird eine positive Spannung gemessen, wenn als Referenzelektrode die Nase verwendet wurde. Dieses Phänomen konnte durch Dipolanalysen vielfach bestätigt werden (Giard *et al.* 1995) und ist auf Generatoren im auditorischen Kortex zurückzuführen. Elektroden, die darunter liegen (z. B. am Mastoid) weisen dann eine entgegengesetzte Polarität als die darüber liegenden Elektroden (z. B. Cz, Fz, vgl. Kap. 4.2.5) auf.

In MEG-Experimenten konnten ebenfalls mehrere Generatoren im auditorischen Kortex, insbesondere in der rechten Hemisphäre gefunden werden (Levänen *et al.* 1996). Auch bei Katzen wurden, hier jedoch durch intracranielle Ableitungen, Generatoren im auditorischen Kortex (A1) lokalisiert (Ulanovsky *et al.* 2003). Zur MMN tragen aber auch Bereiche außerhalb des auditorischen Kortex bei. So sind separate Quellen im Bereich des Frontallappens (Opitz *et al.* 2002) vermutlich mit der durch den Deviantstimulus hervorgerufenen unwillkürlichen Aufmerksamkeitsänderung verbunden (Paavilainen *et al.* 2003). Neuere Untersuchungen hierzu ergaben, dass auch die P3a, die allgemein als Index für diese Aufmerksamkeitsänderung verwendet wird, im Kontext von MMN-Messungen auftritt und dem selben Mechanismus zuzuordnen ist (Rinne *et al.* 2006). Untersuchungen an Patienten, bei denen verschiedene Bereiche des Kortex geschädigt waren, bestätigen die genannten Erkenntnisse, in dem Generatoren in den läsierten Arealen ausgeschlossen werden konnten (Alho 1995).

Bei bestimmten Deviantarten kann die MMN aus zwei Komponenten bestehen (Scherg *et al.* 1989). Die erste Komponente hat Latenz und Quellen ähnlich der N1 und tritt bei geringen Differenzen zwischen Standard- und Deviantreiz nicht auf. Demnach ist die zweite Komponente, die im supratemporalen Planum etwas anterior zum N1-Generator lokalisiert ist und eine frontolaterale Orientierung des Dipols aufweist, die eigentliche MMN, während die erste Komponente die Aktivierung des N1-Generators darstellen könnte.

Als weitere Methoden zur MMN-Quellendetektion haben sich intrakranielle Ableitungen (Kropotov *et al.* 1995; Opitz *et al.* 2002), die Positronemissionstomographie (PET) (Tervaniemi *et al.* 2000) und die funktionelle Magnetresonanztomographie (fMRI) (Opitz *et al.* 1999) etabliert. Besonders in den fMRI-Experimenten konnte eine rechtshemisphärisch

erhöhte hämodynamische Aktivität nachgewiesen werden, wenn Oddball-Stimuli präsentiert wurden (Stevens *et al.* 2005).

Generell wird davon ausgegangen, dass die exakte Lokalisation von MMN-Generatoren von den Stimulusparametern und der Komplexität der Stimuli abhängt (Giard *et al.* 1995; Alho *et al.* 1996). Die Hauptbeiträge zur MMN entstammen demnach dem auditorischen Kortex, wobei rechtshemisphärische frontale Generatoren weitere Beiträge liefern. Diese sind dann meist durch Deviantstimuli getriggert, also mit der dadurch verursachten unwillkürlichen Aufmerksamkeitsänderung (vgl. Abb. 3.4) korreliert (Escera *et al.* 2000; Näätänen *et al.* 2001).

3.4.3 Anwendungsmöglichkeiten – Messparadigmen

MMN-Messungen sind auf Grund ihres relativ geringen technischen Aufwandes und der weitgehenden Unabhängigkeit der MMN von der Art des Devianten eine in Grundlagen- und klinischer Forschung vielfältig angewandte Methode. Die MMN stellt, da sie passiv messbar ist, ein objektives Hilfsmittel zur Diagnose von zentralen auditorischen Prozessen, die durch Alterung und vielfältige kortikale Störungen beeinflusst werden, dar (Näätänen und Escera 2000). Die Amplitude und Latenz der MMN wird zudem durch Trainingseffekte modifiziert, weshalb sie auch zur Bestimmung von Rehabilitationserfolgen einsetzbar ist (Dalebout und Fox 2001).

Mögliche Anwendungsgebiete der MMN reichen von der Pädiatrie und Neuropädiatrie über Sprach- und Sprecherziehung (Dyslexie) bis hinein in die Psychiatrie (Alkoholismus, Schizophrenie) und Neurologie. Auch das Leistungsvermögen technischer Hörhilfen (Cochlear Implant, Hörgeräte) kann mit ihrer Hilfe untersucht werden. Die MMN kann als Prediktor für das Aufwachen aus dem Koma verwendet werden, bei Patienten mit M. Parkinson und M. Alzheimer tritt sie mit geringerer Amplitude oder verzögert auf (Tervaniemi *et al.* 1999). Eine umfangreiche Aufstellung findet sich in (Näätänen und Escera 2000) und (Näätänen *et al.* 2001).

Im Kontext der vielfältigen Einsatzmöglichkeiten soll aber auch auf Nachteile hingewiesen werden. So ist die MMN interindividuell in Amplitude und Latenz sehr verschieden. Gesicherte Messwerte erfordern demzufolge lange Messzeiten. Auch ist ihre eindeutige Identifikation nach Dalebout und Fox (2001) ein Problem, das gegebenenfalls den Einsatz der Signalentdeckungstheorie erfordert.

Ein weiteres Problem bei Potentialen wie der MMN ist ihre recht geringe Reproduzierbarkeit. Bei Tests mit Abweichungen in Tonsilben (/ga/ und /da/) ließ sich bei Probanden ohne Hörstörung z. B. nur eine Reproduzierbarkeit der MMN von 29 % erreichen. Bei Variation der Stimulusdauer, Frequenz oder Intensität wird eine Reproduzierbarkeit der Messungen von 30–60 % erreicht (Dalebout und Fox 2001). Wichtig ist dabei auch die geeignete Justierung der Abweichung. Der Unterschied zwischen Standard- und Deviantreiz (z. B. Pegel) soll zwar deutlich hörbar sein, zu große Unterschiede provozieren jedoch ein bewusstes Hinhören, was zusätzliche Aufmerksamkeitseffekte auf die MMN erzeugt (Tervaniemi *et al.* 1999). Dann würden auch andere, mit der MMN verwandte Potentiale (z. B. N2b) diese überlagern. Die MMN ist demzufolge – trotz ihrer Vorteil – für die klinische Diagnostik nur sehr begrenzt geeignet.

Mit einer Latenz von etwa 150 ms überlagert die MMN den N1-P2-Komplex. Da sich dessen Amplitude mit dem SOA verringert (vgl. Kap. 3.3), kann durch ein entsprechend kurzes SOA diese Überlagerung begrenzt werden. Zu kurze SOA führen jedoch dazu, dass bereits im Latenzbereich der interessierenden AEP eines Stimulus bereits der folgende präsentiert wird. Für die Messung einer MMN geeignete SOA liegen im Bereich von 300–700 ms (Näätänen 1992; Scherg *et al.* 1989).

Die MMN wird des Weiteren von der Deutlichkeit des Unterschieds zwischen Standard- und Deviantreiz und der Diskriminationsfähigkeit der Probanden beeinflusst. Untersuchungen der MMN-Amplitude und -Latenz zeigten, dass die Amplitude der MMN auf Intensitäts-Devianten z. B. von der Stärke des Intensitätsunterschieds der Stimuli abhängt. Deutlichere Unterschiede rufen hierbei eine größere MMN mit geringerer Latenz hervor (Näätänen 1992).

Da die MMN ein auditorisch evoziertes Potential ist, das auf seltene Deviantreize in einem Oddball-Paradigma evoziert wird, sind für die Stimulation Randbedingungen zu beachten. Die Bedingung des zufälligen Einstreuens der Deviantreize muss zunächst zur Erzeugung einer MMN nicht zwingend erfüllt sein (Scherg *et al.* 1989). Dennoch werden meist randomisierte Stimulusfolgen verwendet. Nach einem Deviant sollten dabei zunächst einige Standardreize folgen, um die Gedächtnisspur wieder zu stabilisieren.

Die MMN-Latenz und -Amplitude verändern sich mit dem Alter der Probanden, was auch mit dem verwendeten SOA zusammenhängt. So lässt sich bei langen SOA (4,5 s) bei über 50jährigen Probanden durch die Verkürzung des Kurzzeitspeichers keine MMN mehr nachweisen, wohingegen sie im Alter von etwa 20 Jahren noch vorhanden ist (Näätänen 2003). Die Latenz hingegen ist nur leicht altersabhängig. Sie nimmt im Alter von 25 bis 85

Jahren um etwa 25 ms ab, was auf eine relative Konstanz der MMN-Latenzen im jungen Erwachsenenalter schließen lässt. Hier sind die interindividuellen Unterschiede größer als die durch den Altersunterschied hervorgerufenen (Näätänen 1992).

Einen weiteren, starken Einfluss auf die Nachweisbarkeit der MMN hat die Vigilanz der Probanden. Wie alle späten AEP unterliegt auch die MMN ihrem Einfluss und ist bei schlafenden Patienten durch ein verändertes EEG (Schlafspindeln) kaum messbar (Näätänen 1992). Somit müssen die Probanden während der Messung wach und möglichst in einem kontrollierten Aufmerksamkeitszustand sein. Dies kann z. B. durch die Präsentation eines Videofilms erreicht werden, auf den die Probanden ihre Aufmerksamkeit richten sollen. Um die Messung der AEP nicht durch undefinierte Störgeräusche zu beeinflussen, sollte dieser lautlos und ggf. mit Untertiteln präsentiert werden.

3.5 Paradigmen mit MMN bei Streaming

Die MMN eignet sich – durch ihre präattentive Wirkungsweise – auch zur Analyse von Streaming und Grouping. Dass diese dabei offenbar auf der Basis von Objekten arbeitet, zeigen die folgenden Ausführungen.

Durch wiederholt präsentierte Sequenzen von fünf Tönen (*AAAAB*) wurde von Sussman *et al.* (1998b) untersucht, ob durch den (immer an fünfter Position auftretenden) Stimulus *B* eine MMN evoziert wird. Dies sollte dann nicht geschehen, wenn die Sequenz zu einem einzigen auditorischen Objekt gruppiert wird. In der Tat wurde bei langsamer Präsentationsrate ($SOA = 1,3$ s) eine MMN evoziert, wohingegen bei einer schnelleren Rate ($SOA = 100$ ms) diese nicht nachzuweisen war. Bei dieser schnellen Rate sind die fünf Töne offensichtlich zu einem Objekt gruppiert worden, was dazu führte, dass der Ton *B* nicht mehr als Deviant erkannt werden konnte.

In einem anderen Experiment im Kontext von Streaming wurden dem rechten und linken Ohr abwechselnd Töne jeweils unterschiedlicher Länge (rechts 300 ms, links 100 ms) sowie Frequenzen und Intensitäten präsentiert (Ritter *et al.* 2000). Auf beiden Seiten wurden Devianttöne mit einer Länge von 200 ms zufallsverteilt eingestreut, wobei die übrigen Stimuluseigenschaften der jeweiligen Seite beibehalten wurden. In zwei Kontrollbedingungen wurden diese Devianttöne jeweils ohne weitere Stimuli auf den entsprechenden Seiten präsentiert, um das durch diese evozierte Potential von dem Deviantpotential zu subtrahieren. Es zeigt sich, dass die MMN auf der linken Seite mit wesentlich kürzerer Latenz als auf der rechten Seite auftritt. Auch evozieren nur die Devianten eine MMN,

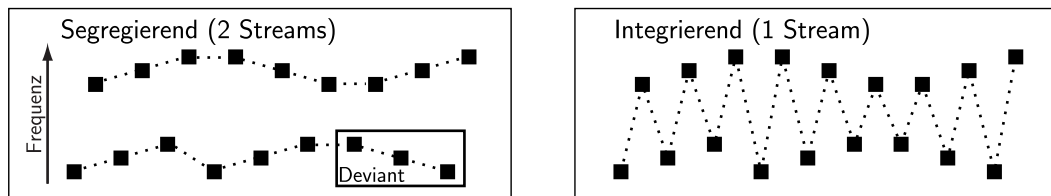


Abb. 3.5: Stimulationsparadigma nach Sussman *et al.* (1999). In zwei Frequenzbereichen wurden je drei Töne verschiedener Frequenz als auf- und absteigende Tripel in einem Oddball-Paradigma präsentiert. Die gepunktete Linie zeigt die erwartete auditorische Organisation: Wenn die Stimuli in zwei Streams segregiert wurden (links), konnten die absteigenden Tripel als Deviant erkannt werden, im Falle der Integration (rechts) hingegen nicht.

die eine direkte Beziehung zu den Standards auf der jeweiligen Präsentationsseite haben. Die Ergebnisse sind ein Indiz dafür, dass die MMN einer Verarbeitungsstufe zugeordnet werden kann, auf der die Objektbildung bereits ausgeführt ist.

Zu dieser Erkenntnis kommen auch Sussman *et al.* (1999, 2005), die ihren Experimenten aus drei Tönen unterschiedlicher Frequenz bestehende auditorische Objekte zu Grunde legten. Diese wurden als auf- oder absteigende Gruppen dreier Töne angeordnet (vgl. Abb. 3.5). Diese Tripel wurden in zwei verschiedenen Frequenzbereichen als Melodielinien angeordnet. Die Folge aufsteigender Tripel wurde dabei in jeder Melodie durch absteigende unterbrochen. Als Deviant-Objekte evozierten sie nur dann eine MMN, wenn die Probanden die Melodielinien (Streams) segregierten. Andernfalls konnten die Muster, und damit auch deren Veränderung von den Probanden, nicht wahrgenommen werden, was zum Ausbleiben der MMN führte.

Mit einem ähnlichen Paradigma unter Verwendung von drei Melodielinien wurde in einem weiteren Experiment innerhalb der tieffrequenten und mittelfrequenten Melodie ebenfalls eine MMN gefunden (Sussman *et al.* 2005). Dieses Resultat erweitert den Raum der Anwendungsmöglichkeiten der MMN auf komplexere Paradigmen erheblich. Auch in der vorliegenden Arbeit werden Objekte aus Drei-Ton-Folgen (Tripel) zur Evozierung einer MMN in Streaming-Experimenten verwendet.

Die Extraktion von abstrakten auditorischen Mustern aus einer Stimulussequenz konnte des Weiteren durch Verwendung von Tonpaaren, deren Töne entweder aufsteigend (Standard) oder absteigend (Deviant) angeordnet waren, belegt werden (Saarinen *et al.* 1992; Paavilainen *et al.* 1995). Das Frequenzverhältnis beider Töne variierte hier zufällig zwischen fünf Werten, so dass die Paare zwar zunächst physikalisch unterschiedlich, aber auf

einer höheren abstrakten Ebene gleich waren. Da die absteigenden Paare eine MMN evozierten, können deren abstrakte Eigenschaften offenbar auch im automatisch geformten Speicher abgelegt werden. Die Verarbeitung dieser Extraktion wurde in diesen Arbeiten hauptsächlich in der linken Hemisphäre lokalisiert. Auch die Möglichkeit, eine MMN bei zufälliger Auslassung von Stimuli in einer regelmäßigen Folge zu evozieren, lässt Rückschlüsse auf die auditorische Verarbeitung zu. Im Kontext eines Streaming-Experiments wurde gezeigt, dass die zeitliche Integration von Stimuli dem Streaming nachgeschaltet ist und auch die Integrationszeit für Streaming bei etwa 170 ms bestimmt (Yabe *et al.* 2001).

Die MMN scheint auch in realitätsnahen auditorischen Szenen ein wichtiges Detektionswerkzeug für plötzliche Änderungen zu sein. In einem Gemisch aus verschiedenen Geräuschen kann sie evoziert werden, wenn eines dieser Geräusche Unregelmäßigkeiten aufweist (Winkler *et al.* 2004). Dies legt nahe, dass eine Repräsentation verschiedener Streams im auditorischen Kortex der weiteren Verarbeitung zu Grunde liegt. Dabei hängt die Evozierung der MMN davon ab, ob nachfolgende Stimuli zu einem Objekt gruppiert werden. Ihre Amplitude nimmt ab, je später ein Deviant im Zeitintervall der (unbewussten) Erwartungshaltung präsentiert wird (Atienza *et al.* 2003). Generell ist die MMN auf das Erkennen einfacher auditorischer Merkmale ausgerichtet und bietet die Basis für eine nachfolgende attentive Verarbeitung, die auf ausgewählte Streams gerichtet ist (Nager *et al.* 2003; Sussman *et al.* 2005).

3.6 Visuelle Mismatch-Negativity

Auch in der visuellen Modalität lassen sich evozierte Potentiale durch reizkorrelierte Mittelung darstellen. Durch optische Reizung wie mit Schachbrettmustern, deren Feldfärbungen wechselt, oder mit Lichtblitzen, werden in der Retina Aktionspotentiale ausgelöst und über die Sehnerven zum Chiasma opticum geleitet. Dort kreuzen sich die Fasern, die die Signale der nasalen Retinahälften transportieren. Von dort aus verlaufen die Wege retinotop zu den primären visuellen Kortizes, die sich bilateral im Occipitallappen befinden. Erst hier fusionieren Informationen aus dem rechten und linken Auge (Altenmüller *et al.* 1996).

Die visuell evozierten Potentiale, deren topologische Zuordnung (abgesehen vom Retinogramm) zu Generatoren auf der Sehbahn nicht eindeutig möglich ist, sind occipital über

dem visuellen Kortex registrierbar. Sie werden nach ihrer Latenz und Polarität analog zu den AEP mit N70, P100 und N140 bezeichnet (Altenmüller *et al.* 1996).

In der auditorischen Modalität ist die Mismatch-Negativity wohldefiniert. Da sie einen Mechanismus reflektiert, der Veränderungen detektiert, ist es naheliegend, auch in der visuellen Modalität einen solchen Mechanismus zu postulieren. In letzter Zeit sind einige Experimente durchgeführt worden, deren Ergebnisse jedoch kontrovers bewertet werden. Besonders die Frage, ob die gefundenen EKP-Komponenten aufmerksamkeitsunabhängig und vom Refraktaritätseffekt zu separieren sind, lässt sich meist nicht eindeutig klären. Übereinstimmend wird jedoch von einem Potential berichtet, das durch Reize evoziert wird, die von einer regelmäßigen visuellen Stimulation abweichen. Dieses ist ebenfalls negativ und wird deshalb analog zur auditorischen MMN als visuelle MMN bezeichnet (für eine Übersicht siehe: Pazo-Alvarez *et al.* 2003).

Es scheint sich zu manifestieren, dass nur dann eine visuelle MMN evoziert wird, wenn die Abweichung nicht im Fokus der Aufmerksamkeit liegt. So konnte durch die Veränderung der Lage zweier Muster zueinander eine Negativität evoziert werden, während die aufmerksam zu verfolgende Farbänderung derselben Stimuli keine Negativierung hervorrief (Berti und Schröger 2006).

Einfache Veränderungen von visuellen Stimuli, wie z. B. das Verformen eines Kreises zu einer Ellipse können ebenfalls zur Evozierung einer visuellen MMN führen: Die Muster wurden in einem Oddball-Paradigma angeordnet und eine Negativierung mit einem Maximum bei etwa 180 ms nach Stimulusbeginn gemessen (Besle *et al.* 2005). Eine visuelle MMN mit einer Latenz von 140–200 ms wurde auch in einem Experiment mit Schachbrettmustern verschiedener Farben nachgewiesen (Czigler *et al.* 2006). Hier evozierten irreguläre Wiederholungen dieser Muster eine Negativierung mit maximaler Amplitude an der Elektrodenposition Oz, also über dem visuellen Kortex.

Interessant ist, dass auch im Kontext des McGurk-McDonald-Effekts (vgl. Kap. 2.3) die MMN als Methode verwendet wurde (Colin *et al.* 2002b). Hier konnte zunächst durch alleinige Präsentation mit visuellen Reizen keine visuelle MMN evoziert werden. Nur wenn die visuellen Stimuli in den Kontext einer auditorischen Szene eingebunden waren, war eine visuelle MMN nachweisbar.

Bei Experimenten mit kombinierter auditorischer und visueller Stimulation treten die EKP (und damit auch die MMN) beider Modalitäten überlagert auf. Sie entstammen verschiedenen Generatoren, werden aber im EEG auch an Elektrodenpositionen gemessen, die nicht in der Nähe der jeweiligen Generatoren sind. Die Trennung beider Potential-

komponenten stellt somit eine Herausforderung dar, der – wie in dieser Arbeit – mit einer Quellenanalyse begegnet werden kann.

3.7 Problemstellung

Die Analyse auditorischer Szenen, wie sie in täglichen Hörsituationen auftreten, stellt besonders für Schwerhörige und Patienten mit zentralen auditiven Verarbeitungsstörungen eine Schwierigkeit dar. Nicht zuletzt deshalb ist das Verständnis der zu Grunde liegenden neuronalen Prozesse von besonderem Interesse. Besonders Schwerhörige sind auf die zusätzlichen visuellen Reize angewiesen, und entnehmen diese z. B. den Lippenbewegungen eines Sprechers. Gerade bei der Wahrnehmung von Musik und Sprache sind die auditorischen Informationen oft mit visuellen verknüpft. Eine solche Kombination kann die auditorische Wahrnehmung beeinflussen und verbessert möglicherweise die Segregation der akustischen Informationen vom Störsignal. Hier ergibt sich der Ansatzpunkt dieser Arbeit. Mit Sinustönen, also einfachen, physikalisch beschreibbaren Stimuli soll eine auditorische Szene konstruiert werden. Die Stimulation soll mit visuell präsentierten Mustern kombiniert werden, um hierdurch eine Beeinflussung der auditorischen Wahrnehmung zu erzielen. Die dieser Arbeit zu Grunde liegenden Hypothesen und Ziele werden im Folgenden zusammengefasst.

- Zunächst soll untersucht werden, inwieweit die Verarbeitung komplexer auditorischer Informationen – und damit deren Perzept – durch visuelle Informationen beeinflusst werden kann. Wenn diese Interaktion auch im Kontext der auditorischen Stream-Segregation wirkt, eignet sich für Experimente hierzu der Parameterbereich des Van-Noorden-Diagramms, in dem das auditorische Perzept ambig ist. Die Indifferenz des Perzepts, das zwischen Integration und Segregation wechselt, sollte durch die synchrone Präsentation visueller Stimuli beeinflussbar sein.
- Dass audiovisuelle Interaktionen auf höheren, kognitiven Ebenen der Reizverarbeitung vorhanden sind, ist leicht nachvollziehbar. Schließlich kann der Text und die Melodie von Liedern in Relation zueinander gesetzt werden. In dieser Arbeit soll untersucht werden, inwieweit sich diese Interaktionen auch schon auf basalen Stufen der Verarbeitung auswirken. Die Verwendung der MMN als präattentives neuronales Korrelat auditorischer Verarbeitung soll Anhaltspunkte für die neuronale Verarbeitungsstufe der Interaktionen liefern.

- Ein weiterer Ansatzpunkt ist die Frage von Aufmerksamkeitseinflüssen auf Streaming. Durch audiovisuelle Stimulationen soll die Modulierbarkeit des auditorischen Perzepts aufgeklärt werden, wenn die Aufmerksamkeit der Probanden allein auf die visuellen Reize gelenkt ist.
- Die neuronalen Quellen der audiovisuellen Integration sind ebenfalls Gegenstand dieser Arbeit. Durch EEG- und MEG-Messungen soll untersucht werden, ob diese allein von auditorischen und visuellen, oder auch anderen Kortexarealen repräsentiert werden.
- Wenn die Beeinflussung des auditorischen Perzepts durch visuelle Reize im ambigen Bereich den Van-Noorden-Diagramms möglich ist, liegt der Gedanke nahe, diese auch in Bereichen mit *per se* eindeutigen Perzept (Integration bzw. Segregation) nachzuweisen. Auch dieses soll in der vorliegenden Arbeit durch eine EEG- sowie eine psychoakustische Messung untersucht werden.

Um den genannten Fragen nachzugehen, werden die von Sussman *et al.* (1999) eingeführten Sequenzen, die entweder als ein oder zwei auditorische Streams wahrgenommen werden können, mit visuellen Reizen gepaart. Der Einfluss der visuellen Information auf die neuronale Repräsentation des ambigen auditorischen Inputs wird dahingehend untersucht, ob durch dessen Präsentation entweder synchron zu einem stream-internen oder einem stream-übergreifenden Muster diese Ambiguität aufgelöst werden kann. Durch Erhöhung oder Verringerung des Frequenzabstands zwischen den Tönen wird im weiteren Verlauf der ambige Bereich verlassen und ebenfalls durch visuelle Stimulation versucht, das automatische auditorische Perzept zu verändern.

Kapitel 4

Versuchsdurchführung und Auswertung

Die vorliegende Arbeit nähert sich der genannten Problemstellung mit verschiedenen, im Folgenden vorgestellten experimentellen Ansätzen. Alle Untersuchungen waren Messungen an normalhörenden Probandenkollektiven, die keine oder geringe Schnittmengen untereinander aufweisen. Der schriftlichen Einwilligungserklärung der Probanden ging eine ausführliche Aufklärung über das Ziel der jeweiligen Studie, die Messmethode und gegebenenfalls den damit verbundenen Risiken voraus. Für alle Studien lag zudem ein positives Votum der Ethik-Kommission der Medizinischen Fakultät der Otto-von-Guericke-Universität in Magdeburg vor.

Als erstes wurde ein elektrophysiologisches Experiment zu Streaming im ambigen Bereich des Van-Noorden-Diagramms mittels EEG-Ableitungen durchgeführt (Rahne *et al.* 2007a). Ausgewertet wurde hier im Wesentlichen die MMN als Korrelat für die Streaming-Zustände. Teilergebnisse dieser Studie wurden in einem psychoakustischen und einem MEG-Experiment verifiziert (Rahne *et al.* 2007b). Das MEG-Experiment, welches eine Reproduktion ausgewählter Stimulationsbedingungen der initialen Studie enthielt, sollte zusätzliche Erkenntnisse über die neuronalen Quellen der evozierten Magnetfelder im Kortex gewinnen. Die Konzipierung des psychoakustischen Experiments wurde dahingehend ausgerichtet, dass durch die Stimulationsbedingungen neben dem ambigen Bereich auch Bereiche mit *per se* eindeutigen auditorischen Perzepten (Integration bzw. Segregation) eingeschlossen sind. Schließlich wurde ein weiteres EEG-Experiment durchgeführt, das in diesen nicht-ambigen Bereichen durch visuelle Stimulation das auditorische Perzept zu beeinflussen versuchte.

4.1 Stimulationsreize und Präsentation

Auditorische Stimuli

Zur Darbietung der auditorischen Reize wurden zwölf Reintöne unterschiedlicher Frequenz erstellt, für die EEG-Studie im ambigen Bereich mit der Software STIM¹ und für alle weiteren Studien mit MATLAB². Die Frequenzen dieser Töne entsprechen den Tönen F#5 bis F6 des abendländischen Tonsystems (vgl. Tab. 4.1). Benachbarte Töne haben einen Abstand von einem Halbtonschritt (Frequenzverhältnis: $1/\sqrt[12]{2}$). Die Töne wurden in einen tiefen (T1–6) und einen hohen (H1–6) Frequenzbereich gruppiert. Zur Vermeidung von Klick-Geräuschen wurde die Signalamplitude der im ersten Experiment verwendeten Töne am Anfang und Ende mit linearen Rampenverläufen versehen (fade in/out), die innerhalb von 7,5 ms die Maximalamplitude zwischen 0 % und 100 % veränderten. Bei den in den weiteren Experimenten verwendeten Tönen wurden diese Rampenverläufe durch die Anwendung einer von-Hann-Fensterfunktion (Blackman und Tukey 1959, S. 98–99) mit der Fensterlänge M und der Amplitude y auf die Stimulusamplitude realisiert:

$$y(n) = \frac{1}{2} \left[1 + \cos \left(\frac{2\pi n}{M} \right) \right] \quad (4.1)$$

Für die Anwendung auf die Stimuli dieser Arbeit wurde $M = 15$ ms gewählt und für den Amplitudenanstieg am Stimulusbeginn die Fensterfunktion im Intervall von $n = -0,5M$ bis $n = 0$ berechnet. Für das Ausblenden wurde die zweite Hälfte der Funktion (Intervall von $n = 0$ bis $n = 0,5M$) mit der Stimulusamplitude multipliziert. Alle Töne wurden binaural mit einem konstanten zeitlichen Abstand (SOA) von 110 ms im Pegelbereich von 70 dB bis 85 dB SPL (Sound Pressure Level) präsentiert. Zur Vermeidung von elektrischen Artefakten wurden Einsteckhörer³ verwendet, bei denen der Schall über einen Schlauch zum Ohr geleitet wird.

Visuelle Stimuli

Die visuellen Stimuli wurden für das erste Experiment ebenfalls mit der Software STIM, und für die weiteren mit handelsüblicher Grafiksoftware erstellt. Es handelt sich hierbei um zwei Kreise (K1, K2) und drei Quadrate (Q1, Q2, Q3). Alle Stimuli sind schwarze Figuren

¹Neuroscan, El Paso, TX, USA

²The MathWorks, Inc., Natick, MN, USA

³E-A-RTONE 3A, Auditory Systems, Indianapolis, IN, USA

Tab. 4.1: Auditorische Stimuli (Stimulus), Frequenzen (f) und tonale Bezeichnungen (TB) im abendländischen Tonsystem.

Stimulus	T1	T2	T3	T4	T5	T6
f (Hz)	739,99	783,99	839,61	880,00	932,33	987,77
TB	F#5	G5	G#5	A5	A#5	B5
Stimulus	H1	H2	H3	H4	H5	H6
f (Hz)	1046,5	1108,7	1174,7	1244,5	1318,5	1396,9
TB	C5	C#5	D5	D#5	E5	F6

Tab. 4.2: Visuelle Stimuli: Kreise (K) und Quadrate (Q) und Sehwinkel für einen Abstand von 1,5 m zwischen Proband und Monitor.

Stimulus	K1	K2	Q1	Q2	Q3
Durchmesser (cm)	3	8			
Kantenlänge (cm)			2,5	5,5	10
Sehwinkel (°)	1	3,1	1,3	2,97	5,4

mit weißem Rand, die vor einem schwarzen Hintergrund über einen Monitor präsentiert wurden. Durch die unterschiedlichen Kantenlängen beziehungsweise Durchmesser ergeben sich die in Tab. 4.2 dargestellten Sehwinkel.

Stimuluspräsentation

Die Bedingungen der durchgeführten Experimente bestanden aus einer Kombination auditorischer und visueller Stimuli. Dabei wurde jeweils eine Auswahl von drei Tönen aus dem hohen und dem tiefen Frequenzbereich mit einem Satz visueller Stimuli kombiniert (z. B. T1-3, H1-3, Q1-3). Die Stimuli beider Frequenzbereiche wurden alternierend präsentiert (z. B. T1–H3–T2–H1–...) und bildeten somit zwei parallele (akustische) Streams. Als visuelle Stimuli wurden zum einen die Quadrate verwendet, synchron mit den Stimuli des tiefen Frequenzbereichs präsentiert und ihre Größe der Frequenz zugeordnet (z. B. T1–Q1, T2–Q2, T3–Q3). Diese stream-interne Zuordnung wird im Folgenden *segregierend* genannt. Die Kreise wurden als visuelle Stimuli verwendet, um bestimmte auditorische

Stimuli aus beiden Frequenzbereichen zu markieren. Eine solche stream-übergreifende Synchronisation wird hier als *integrierend* bezeichnet.

4.2 EEG-Studie im ambigen Bereich

Zur Messung der visuellen Beeinflussung von auditorischem Streaming wurde ein etabliertes experimentelles Setup verwendet und hinsichtlich der visuellen Reize erweitert (Sussman *et al.* 1998a). Nach einigen Pilotmessungen in Magdeburg wurde die Studie in den Neurophysiologie-Labors des Albert Einstein College of Medicine (New York, USA) durchgeführt. Den verwendeten Versuchsaufbau zeigt Abb. 4.1 schematisch.

4.2.1 Probanden

Zehn rechtshändige Erwachsene (davon vier männlich) im Alter von 24–30 Jahren nahmen am Experiment teil. Ihr Hörvermögen wurde zuvor mittels Reintonaudiogramm an den Frequenzen 0,5 kHz, 1 kHz, 2 kHz und 4 kHz untersucht. Bis zu einer maximalen Erhöhung der Reintonhörschwelle um 10 dB wurden die Probanden für das Experiment zugelassen. Zur Auswertung der EEG-Messungen konnten nur die Daten von acht Probanden herangezogen werden. Ein Proband konnte die gestellte Aufgabe nicht erfüllen und ein weiterer hatte zu starke α -Aktivitäten über weite Strecken des EEG.

4.2.2 Stimuli

Diese Studie verwendet auditorische Stimuli, die sowohl als ein (integrierter) Stream oder auch als zwei (segregierte) Streams wahrgenommen werden. Dazu wurden die Töne T3–5 mit den Tönen H3–5 kombiniert. Der Tonhöhenabstand zwischen dem höchsten Ton des tieffrequenten Bereichs (T5) und dem tiefsten Ton des höchsten Bereichs (H3) betrug vier Halbtonschritte. In Kombination mit einem SOA von 110 ms führt dies zu einem ambigen auditorischen Perzept (Bregman 1999; Sussman *et al.* 2007). Die Perzepte *Integration* und *Segregation* treten hierbei mit gleicher Wahrscheinlichkeit (50 %) auf und wechseln im Zeitverlauf häufig.

Abb. 4.2 zeigt die auditorische Stimulationssequenz, die mit den Quadraten und Kreisen kombiniert wurde. Der Pegel der Stimuli betrug 70 dB SPL, wobei stream-übergreifend

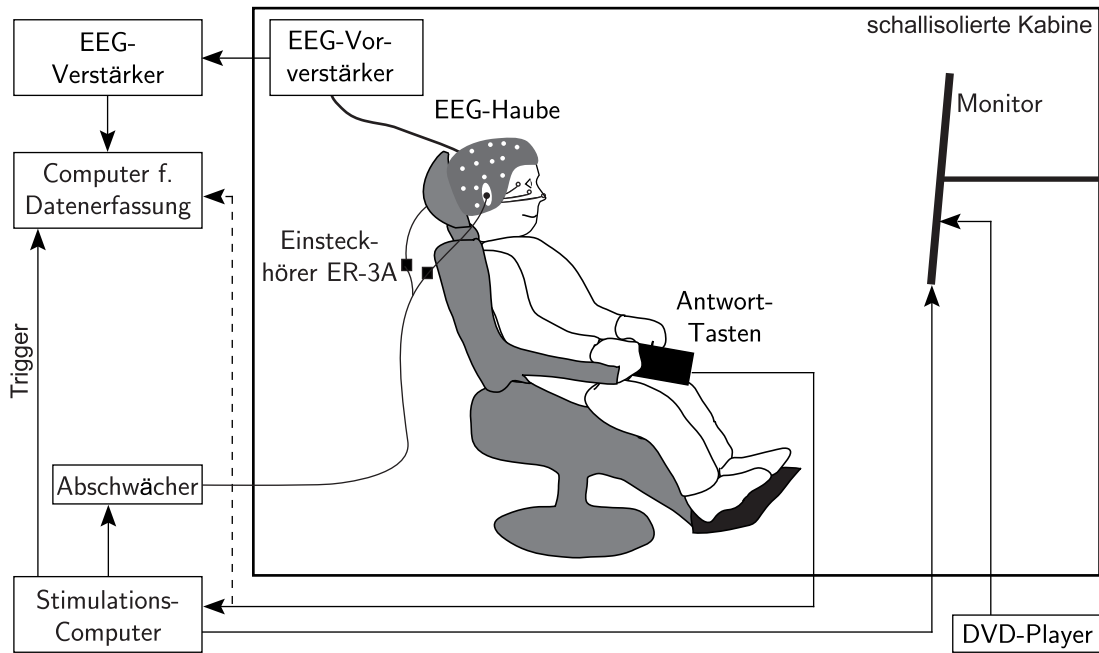


Abb. 4.1: Schematische Versuchsanordnung zur Messung von auditorisch und visuell evozierten Potentialen sowie zur psychometrischen Registrierung von Antworten. Mit einem Computer werden die Stimuli erzeugt und von diesem in einem definierten Paradigma ausgegeben. Die visuellen Stimuli werden zum Monitor geleitet, die auditorischen Reize zunächst zur PegelEinstellung zu einem Abschwächer geleitet und dann über Einsteckhörer appliziert. Der Proband sitzt in einer schallisolierten Kabine. Sein EEG wird über eine Haube abgeleitet und nach Verarbeitung durch den Vorverstärker und EEG-Verstärker mit einem weiteren Computer registriert. Dieser zeichnet auch die Trigger-Signale des Stimulationscomputers auf. Die Antworten der Probanden können sowohl durch den Messcomputer als auch durch den Stimulationscomputer aufgezeichnet werden. In passiven Bedingungen wird ein Videofilm mit einem DVD-Player auf demselben Monitor präsentiert.

jeder dritte Ton mit einem um 15 dB erhöhten Pegel präsentiert wurde. Innerhalb des hohen Frequenzbereichs wurde die Stimulusreihenfolge zufällig gewählt. Dabei war jeder der hohen Töne (H3, H4 und H5) mit einer Wahrscheinlichkeit von 33 % vertreten. Im tieffrequenten Bereich wurden die Töne in aufsteigenden Tripel (T1–T2–T3) angeordnet. Um hier eine Mismatch-Negativity zu evozieren, wurden 15 % der Tripel absteigend präsentiert. Durch die pseudorandomisierte Einstreuung der absteigenden Tripel entstand ein Oddball-Paradigma, das die aufsteigenden Tripel als Standards und die absteigenden Tripel als Devianten enthielt. Es wurde zudem sichergestellt, dass jedem Deviant-Tripel mindestens zwei Standard-Tripel vorangestellt waren.

4.2.3 Bedingungen

Die auditorischen und visuellen Sequenzen wurden, verschieden kombiniert, in fünf verschiedenen Bedingungen präsentiert. Darin enthalten waren drei Kontrollbedingungen. Eine Übersicht zeigt die Tabelle 4.3.

Die *2 Streams–segregierend*-Bedingung enthielt die Töne des hohen und tiefen Frequenzbereichs sowie die Präsentation der Quadrate, deren Seitenlänge mit der Frequenz der Stimuli T3–5 einherging. Durch die Präsentation visueller Stimuli synchron zu der tieffrequenten Tonfolge sollte die Segregation des hochfrequenten und tieffrequenten Streams forciert werden (Abb. 4.2).

In der *2 Streams–integrierend*-Bedingung wurden dieselben auditorischen Stimuli wie in der *2 Streams–segregierend*-Bedingung verwendet. Allein die visuelle Stimulation wurde verändert, indem synchron zu den Tönen mit erhöhtem Pegel die Kreise präsentiert wurden: kleine Kreise, wenn Töne des tieffrequenten Streams präsentiert wurden, große Kreise synchron zu Tönen des hochfrequenten Streams. Hierbei markierte die visuelle Präsentation das stream-übergreifende auditorische Muster, was die Integration der akustischen Information in einen Stream verstärken sollte (Abb. 4.2).

Die *1 Stream*-Bedingung besteht nur aus den Tönen T3–5, die synchron mit den Quadraten präsentiert wurden. Mit dieser Bedingung sollte ein MMN-Potential evoziert werden, wie z. B. von Sussman *et al.* (1999) gezeigt wurde. Durch Kombination mit den visuellen Stimuli Q1–3 sollten während dieser Bedingung ähnliche ereigniskorrelierte Potentiale wie in der *2 Streams–segregierend*-Bedingung evoziert werden. Damit war eine Kontrolle für die *2 Streams–segregierend*-Bedingung gegeben.

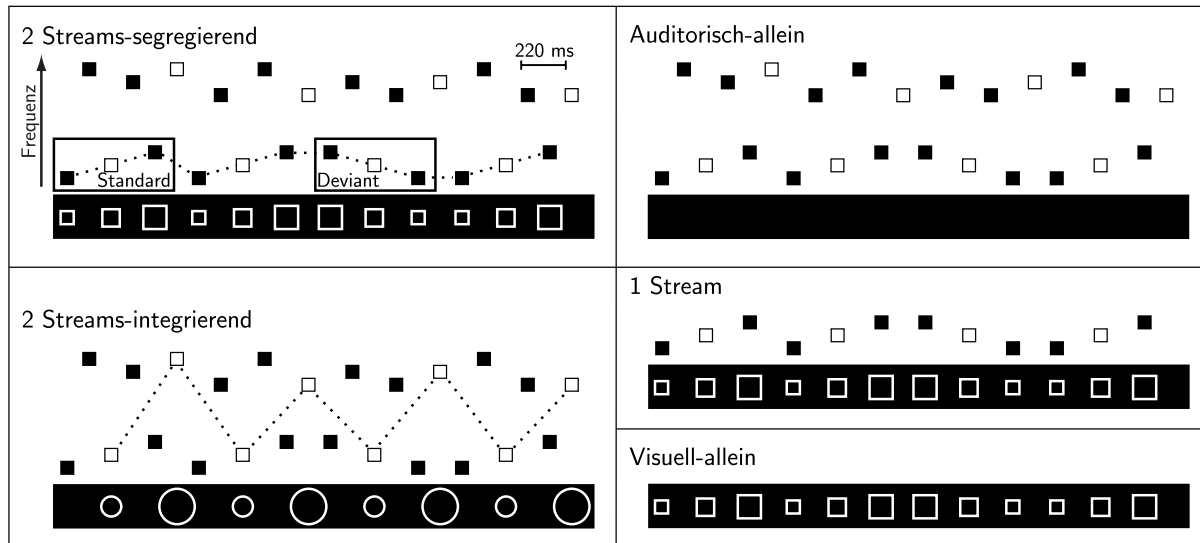


Abb. 4.2: Stimulationsparadigma. Auditorische Stimuli wurden in zwei Frequenzbereichen mit je drei Tönen verschiedener Frequenz präsentiert. Die Darbietung der Töne des hochfrequenten Bereichs erfolgte randomisiert, die des tieffrequenten Bereichs als aufsteigendes oder abfallendes Tripel in einem Oddball-Paradigma. Zusätzlich wurde der Pegel jedes dritten Tones erhöht (weiß: 85 dB, schwarz: 70 dB SPL). Die gepunktete Linie zeigt die erwartete auditorische Organisation: Wenn die visuellen Stimuli synchron zu den Tripel dargeboten werden, wird eine Segregation der Frequenzbereiche erwartet (*2 Streams-segregierend*). Eine Integration in einen auditorischen Stream (*2 Streams-integrierend*) wird erwartet, wenn Synchronität zu dem Intensitätsmuster besteht. Auf der rechten Seite sind die Stimuli der Kontrollbedingungen dargestellt.

Tab. 4.3: Bedingungen des EEG-Experiments mit auditorischen Stimuli innerhalb des ambigen Bereichs und jeweilige visuelle Stimulation.

Bedingung	Auditorische Stimuli		Visuelle Stimuli		Video-Film
	geordnet T1-3	random. H1-3	Q1-3	K1-2	
2 Streams-segregierend	x	x	x		
2 Streams-integrierend	x	x		x	
1 Stream	x		x		
Auditorisch-allein	x	x			x
Visuell-allein			x		

Die *Auditorisch–allein*-Bedingung ist eine Kontrollbedingung für beide *2-Streams*-Bedingungen, in welcher die hohen und tiefen Töne analog zu diesen, allerdings ohne visuelle Stimulation präsentiert wurden. Die elektrophysiologische Antwort auf die auditorischen Stimuli sollte anzeigen, inwieweit beide Formen der Organisation (Integration und Segregation) auftreten, während die Probanden die auditorischen Stimuli ignorierten. Zur Fixierung der Aufmerksamkeit hatten sie einen selbstgewählten Video-Film mit Untertiteln zu verfolgen.

Der Zweck der dritten Kontrollbedingung, der *Visuell–allein*-Bedingung, bestand darin, visuelle EKP in Abwesenheit auditorischer Stimuli zu messen. Die visuelle Präsentation bestand hierbei aus den Stimuli Q1–3, die in derselben Weise präsentiert wurden wie in der *2 Streams–segregierend*-Bedingung.

In allen Bedingungen sollten die Probanden die auditorischen Stimuli ignorieren. Sie erhielten keine Information über die Stimuli. Die Aufmerksamkeit auf die visuellen Stimuli wurde durch folgende Aufgabe kontrolliert: 1 % der visuellen Stimuli wurden als rote anstelle von weißen Mustern dargeboten. Ihr Auftreten sollte von den Probanden durch einen Tastendruck angezeigt werden.

4.2.4 Messablauf

Die Probanden nahmen, bequem sitzend, in einer akustisch gedämmten Kabine am Experiment teil. Die *Auditorisch–allein*-Bedingung wurde in allen Fällen als erste präsentiert, gefolgt von den vier verbleibenden Bedingungen in zufälliger, ausbalancierter Reihenfolge. Der Grund für den Beginn mit der Bedingung ohne visuelle Stimulation war, zunächst einen Überblick über das rein auditorisch evozierte Perzept zu erhalten, bevor irgendeine Beeinflussung erfolgen konnte. Jede Bedingung wurde in vier zusammenhängende Blöcke von etwa vier Minuten Dauer unterteilt, die jeweils 55 Devianten enthielten. Insgesamt wurden 220 Devianten pro Bedingung präsentiert.

Bei der Hälfte der Probanden wurden die Frequenzbereiche der Streams invertiert, so dass die Tripel im hochfrequenten Bereich enthalten waren. Die visuellen Stimuli wurden derart adaptiert, dass die Quadrate auch hier die Tripel markierten. Einschließlich der Präparationszeit und der Pausen zwischen den Blöcken und Bedingungen betrug der Zeitaufwand einer Messung etwa drei Stunden.

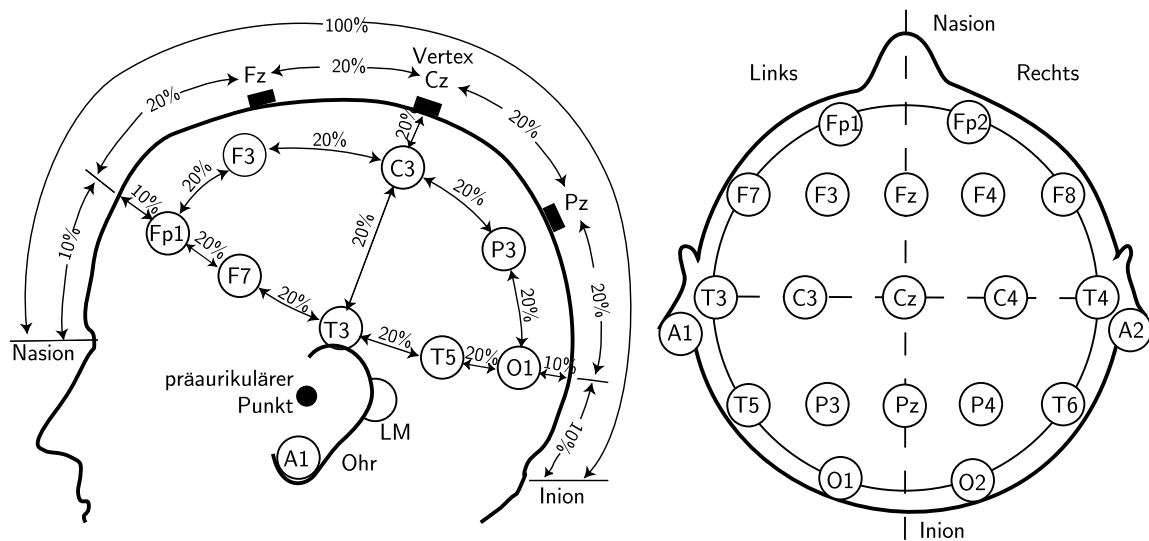


Abb. 4.3: Internationales 10-20-System zur Elektrodenanordnung auf der menschlichen Schädeloberfläche (links: Blick auf die linke Schädelseite, rechts: Blick von oben). Ankerpunkte für dieses relative System sind die Knochenstrukturen Nasion und Inion sowie die präaurikulären Punkte (nach Jasper 1958).

4.2.5 EEG-Messungen

Das EEG wurde kontinuierlich unter Verwendung eines 32-kanaligen Verstärkers⁴ aufgezeichnet. Dabei wurde ein analoger Bandpass von 0,05–200 Hz (12 dB/Oktave) und eine A/D-Rate von 1 kHz verwendet. Die Elektroden wurden gemäß des internationalen 10-20-Systems (Jasper 1958) mit einer Haube⁵ an 29 Positionen⁶ angelegt (vgl. Abb. 4.3). Zusätzlich wurden am linken und rechten Mastoid Elektroden befestigt. Die Referenzelektrode befand sich, in Übereinstimmung mit den meisten MMN-Studien, auf der Nasenspitze. Das vertikale Elektrookulogramm (VEOG) wurde mit einer bipolaren Elektrodenkonfiguration zwischen Fp2 und einer unter dem rechten Auge befindlichen Elektrode aufgezeichnet. Das horizontale Elektrookulogramm (HEOG) wurde zwischen den Positionen F7 und F8 registriert. Alle Elektrodenimpedanzen betragen weniger als 5 k Ω .

⁴SynAmps, Neuroscan, El Paso, TX, USA

⁵Electro-Cap International, Inc., Eaton, OH, USA

⁶Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F7, F8, F3, F4, Fc5, Fc6, Fc1, Fc2, T7, T8, C3, C4, Cp5, Cp6, Cp1, Cp2, P7, P8, P3, P4, O1, O2

4.2.6 Analyse der EEG-Daten

Das EEG an Cz, F3, F4, Fz und den EOG-Kanälen wurde automatisch auf Bewegungsartefakte untersucht: Wenn der Spannungsunterschied zwischen zwei aufeinander folgenden Messpunkten mehr als $20 \mu\text{V}$ betrug, oder die Differenz in einem Intervall mit einer Länge von 100 ms größer als $100 \mu\text{V}$ war, wurde dieser Bereich als artefaktbehaftet markiert und von der späteren Mittelung ausgeschlossen. Danach erfolgte die digitale Filterung des EEG in den Grenzen von 1–15 Hz (24 dB/Oktave). Das gefilterte EEG wurde separat für jede Bedingung relativ zu einem Stimulus-Marker (Standard und Deviant) in Epochen unterteilt, die 100 ms vor Beginn des Markers begannen und eine Länge von 600 ms hatten. Die Epochen der Standard-Stimuli beziehen sich auf den ersten Ton der aufsteigenden Tripel (T1), die der Deviantstimuli auf den ersten der absteigenden Folge (T3). Nachdem eine Nulllinien-Korrektur (–100 bis 0 ms) angewendet wurde, sind die Epochen, deren Amplituden $\pm 50 \mu\text{V}$ überstiegen, entfernt worden. Bei den ausgewerteten Datensätzen führte die Artefaktreduktion, zu einem Ausschluss von etwa 15 % der Epochen. Somit blieben für die Berechnung der individuellen MMN etwa 185 Deviant-Epochen pro Bedingung übrig. Die EKP wurden als Mittelwerte der Standard- und Deviant-Epochen berechnet. Anschließend erfolgte eine Differenzbildung dieser EKP (Deviant minus Standard) für die Beurteilung der MMN. Für beide Epochentypen wurden separat auch die Gruppenmittelwerte berechnet.

Zur Berechnung der MMN-Amplitude wurden im Fenster von ± 20 ms um einen sich aus dem Gruppenmittelwert ergebenden Zeitpunkt die individuellen Amplituden gemittelt. In den Bedingungen *2 Streams-segregierend*, *2 Streams-integrierend*, *Auditorisch-allein* und *Visuell-allein* war dies das Amplitudenmaximum im Latenzbereich von 100–200 ms der *2 Streams-segregierend*-Bedingung, da in den anderen Bedingungen keine auditorische MMN zu erwarten war und die Stimuli (sofern vorhanden) identisch waren. Zur Bestimmung der MMN-Amplitude in der *1 Stream*-Bedingung wurde im gleichen Latenzbereich das Amplitudenmaximum separat bestimmt.

Zur weiteren Auswertung wurden Elektrodenpositionen ausgewählt, die das Auftreten der auditorisch und visuell evozierten Potentiale charakterisieren. Neben den Standardpositionen Fz, Cz, und Pz wurden zusätzlich die frontalen Positionen F3, F4, FC1 und FC2 verwendet, da hier das Maximum für die MMN zu erwarten ist. Die Mittelwerte wurde mit Hilfe einer Varianzanalyse (analysis of variance, ANOVA) untersucht. Dieses Verfahren ist für den Vergleich von Mittelwerten, die sich durch mehrere unabhängige Einflussgrößen (Faktoren) ergeben, geeignet. Es überprüft als Nullhypothese, ob alle Mittelwerte gleich

sind. Die ANOVA ist ein parametrischer Test, setzt also normalverteilte Mittelwerte voraus. Zudem sollten die Varianzen der Mittelwerte und die Stichprobengrößen homogen sein. Es existiert auch eine ANOVA-Variante für wiederholte Messungen, die z. B. an den gleichen Probanden durchgeführt wurden. Diese setzt zusätzlich voraus, dass die Kovarianz zwischen Bedingungs-paaren etwa gleich ist. Diese *Sphärizität* genannte Voraussetzung kann mit dem *Mauchly-Test* überprüft werden. Er testet, ob die Varianzen der Differenzen zwischen den Bedingungen gleich sind. Bei Verletzungen der Sphärizität muss eine Korrektur (z. B. nach Greenhouse-Geisser) verwendet werden. Hierbei werden die durch die ANOVA ermittelten F -Werte mit einem vom Freiheitsgrad abhängigen Faktor korrigiert.

Die ANOVA beruht auf F -Tests der Varianzen. Für die verschiedenen Einflussfaktoren werden Gruppen der Messwerte gebildet und getestet, ob die Varianz zwischen den Gruppen größer ist als die Varianz innerhalb der Gruppen. Die ANOVA zeigt also nur, *ob* Effekte von Faktoren vorhanden sind oder *ob* Interaktionen von verschiedenen Faktoren signifikant sind. *Wie* sich der jeweilige Einfluss darstellt, muss durch *post-hoc*-Test ermittelt werden (Field 2005). Diese bestehen aus paarweisen Vergleichen (z. B. t -Tests), die jede mögliche Kombination von Gruppen enthält. Dabei werden Stichproben mehrfach getestet. Um die Wahrscheinlichkeit der fälschlichen Ablehnung der Nullhypothese (α , Typ-I-Fehler) auf dem gewünschten Niveau zu halten, haben sich Korrekturverfahren etabliert, die auch in dieser Arbeit verwendet wurden.

- *Bonferroni-Korrektur*. Der einfachste, aber auch konservativste Weg, den Typ-I-Fehler zu kontrollieren ist die Bonferroni-Korrektur. Hier wird α durch die Anzahl der Vergleiche dividiert und das so ermittelte α' als Signifikanzniveau für die Einzeltests verwendet.
- *Tukey-Test*. Dieser Test führt alle paarweisen Vergleiche auf Basis der t -Verteilung durch. Er ist ebenfalls konservativ, aber für eine größere Anzahl von paarweisen Vergleichen als die Bonferroni-Korrektur geeignet. Er setzt die Fehlerrate für das Experiment gleich der Fehlerrate für die Gesamtheit aller paarweisen Vergleiche.

Das Auftreten einer MMN wurde in der vorliegenden Arbeit durch eine ANOVA mit den drei Faktoren *Elektrode* (Stufen: Fz, F3, F4, FC1, FC2, Cz, Pz), *Stimulustyp* (Stufen: Standard, Deviant) und *Bedingung* (5 Stufen) durchgeführt. Die Differenzamplituden wurden unter Verwendung einer ANOVA mit den Faktoren *Elektrode* (Stufen: Cz, Fz, Pz, Oz, F3, F4, FC1, FC2, LM, RM) und *Bedingung* (5 Stufen) verglichen. *Post-hoc*-Tests

mit Bonferroni-Korrektur für Mehrfachvergleiche wurden angewendet, um die gefundenen Effekte zu spezifizieren.

Für jede Bedingung wurde auf der Basis der Gruppenmittelwerte sowohl die Verteilungen der Differenzpotentiale auf der Kopfhaut als auch ihre Quellendichten (Source Current Density, SCD) berechnet und im Latenzbereich der MMN graphisch dargestellt. Als Ausgangspunkt wird hierbei die Potentialverteilung auf der Kopfhaut verwendet. Um von der Potentialverteilung auf die Stromflüsse schließen zu können, wird die elektrische Stromdichte \vec{j} über das Ohmsche Gesetz mit der elektrischen Feldstärke \vec{E} verknüpft (σ : elektrische Leitfähigkeit):

$$\vec{j} = \sigma \vec{E} \quad (4.2)$$

Bildet man die Divergenz dieser Gleichung und substituiert $\vec{E} = -\nabla\phi$, so folgt

$$\nabla \bullet \vec{j} = -\sigma \nabla \bullet (\nabla\phi) \quad (4.3)$$

und damit der Zusammenhang zwischen Stromquellendichte und zweiter räumlicher Ableitung der Potentialverteilung auf der Kopfhaut:

$$\nabla \bullet \vec{j} = -\sigma \nabla^2 \phi \quad (4.4)$$

Die Empfindlichkeit der SCD ist für oberflächennahe Quellen am höchsten und fällt mit dem Abstand r zur Quelle ($\sim r^{-4}$). Zur Beschreibung von tiefergelegenen Quellen ist sie somit nicht geeignet. Die Potentialverteilungen und die SCD wurden mit der Software BESA⁷ berechnet.

Die über alle Probanden gemittelten Differenzkurven wurden einer Quellenanalyse zu Grunde gelegt. Hierbei wurde zunächst mittels *Hauptkomponentenzerlegung* (PCA) das Potential in Eigenvektoren zerlegt und die Anzahl der möglichen Generatoren ermittelt. Unter der Annahme von symmetrisch in beiden Hemisphären lokalisierten Generatorenpaaren wurde ihre genaue Lage unter Verwendung der Software BESA ermittelt. Die Lösung wurde akzeptiert, wenn durch das Modell mehr als 95 % des EKP-Verlaufes erklärt werden konnten. Erwartet wird hierbei, dass der zu Grunde liegende iterative Algorithmus Lösungen generiert, die Regionen des auditorischen und visuellen Kortex zugeordnet werden können. Die Analyse beruhte auf einem ellipsoiden 4-Schalen-Kopfmodell nach Geddes und Baker (1967) mit unterschiedlichen relativen elektrischen Leitfähigkeiten für Gehirn (0,3300), zerebrospinale Flüssigkeit (1,0000), Schädelknochen (0,0042) und Haut (0,3300) sowie den Radien für Kopf (85 mm), Haut (6 mm), Schädelknochen (7 mm), und zerebrospinale Flüssigkeitskammer (1 mm).

⁷MEGIS Software GmbH, Gräfelfing, Deutschland

4.3 Psychoakustische Studie

4.3.1 Probanden

Vierzehn Probanden im Alter von 18–33 Jahren (davon acht männlichen Geschlechts) nahmen an diesem zweiten Experiment teil und unterzogen sich zuvor einem audiometrischen Test (Reintonaudiogramm (Luftleitung) mit den Frequenzen 0,125 kHz, 0,25 kHz, 0,5 kHz, 1 kHz, 1,5 kHz, 2 kHz, 3 kHz, 4 kHz, 6 kHz und 8 kHz). Sie konnten an der Studie teilnehmen, wenn ihre Ruhehörschwelle um nicht mehr als 10 dB erhöht war.

4.3.2 Stimuli

Für diese Studie wurden die auditorischen Stimuli T1–6 und H1–6 verwendet (vgl. Tab. 4.1) und abwechselnd ein Ton aus beiden Frequenzbereichen mit Hilfe der Software PRESENTATION⁸ präsentiert. Wie im EEG-Experiment (vgl. Kap. 4.2.2) wurden die Stimuli des jeweiligen tieffrequenten Bereichs in aufsteigenden Tripel angeordnet, so dass ein Frequenzmuster entstand. Im hochfrequenten Bereich wurden die Töne zufällig verteilt. Der Pegel betrug 65 dB SPL, bei jedem dritten Ton 80 dB SPL, so dass auch in dieser Studie ein stream-übergreifendes Intensitätsmuster entstand.

Die visuellen Stimuli waren die Figuren Q1–3 sowie K1–2 (vgl. Tab. 4.2), die auf einem LCD-Monitor synchron zu den auditorischen Stimuli (wie im EEG-Experiment) in den Bedingungen *integrierend* und *segregierend* angezeigt wurden.

4.3.3 Bedingungen

Insgesamt bestand das Experiment aus achtzehn verschiedenen Kombinationen von jeweils einer auditorischen und einer visuellen Stimulation. Die auditorischen Stimulationen unterschieden sich in dem Abstand der beiden Frequenzbereiche in *weit*, *mittel* und *eng*, was unterschiedlichen Wahrnehmungsbereichen nach van Noorden (Bregman 1999, S. 58ff) entspricht (vgl. Abb. 4.4). Demnach werden die Töne mit weitem Frequenzabstand (T1–3, H4–6) *per se* als zwei segregierte Objekte wahrgenommen, die des engen Abstands (T4–6, H1–3) stets in ein Objekt integriert. Der mittlere Abstand wurde durch die gleichen Stimuli erzeugt, wie sie im EEG-Experiment verwendet wurden (T3–5, H3–5) und

⁸Neurobehavioral Systems Inc., Albany, USA

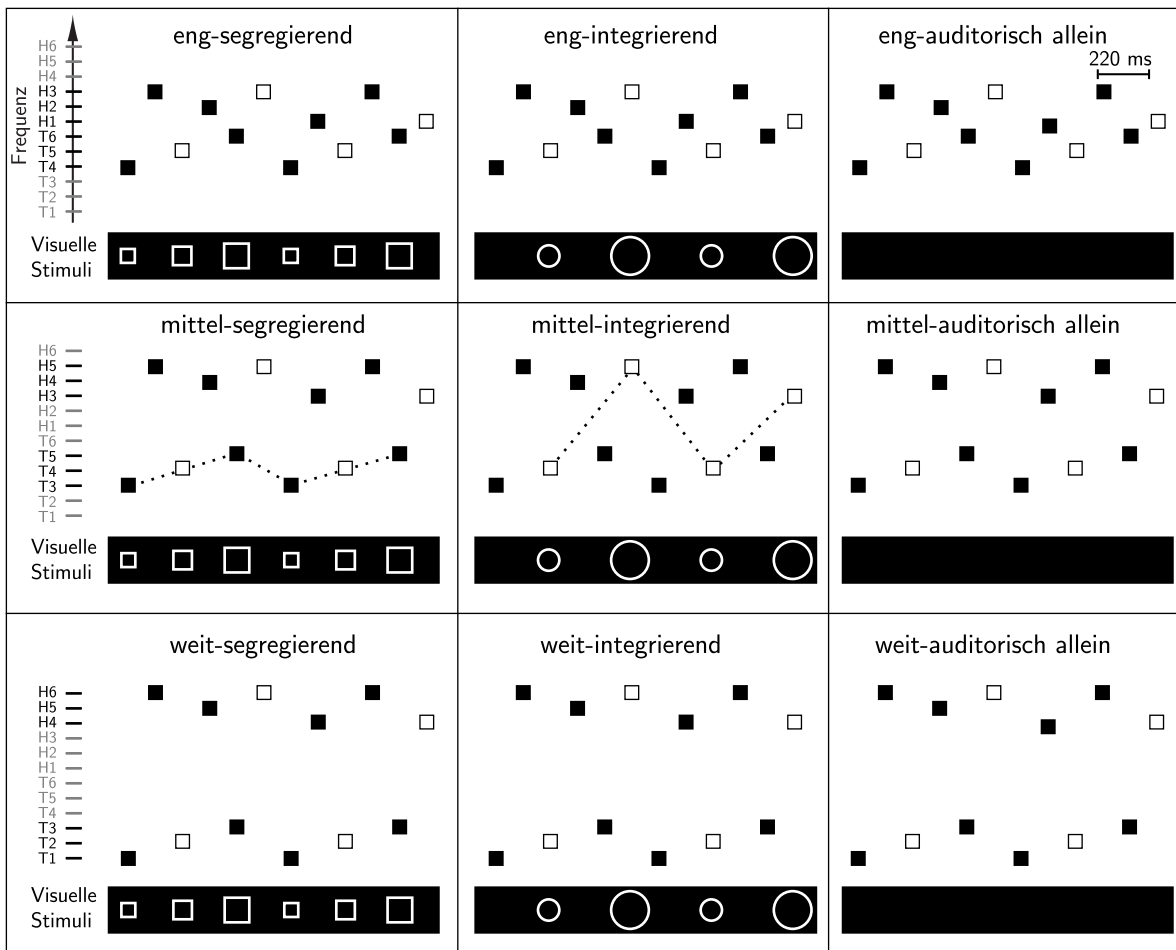


Abb. 4.4: Stimulationsparadigmata des psychoakustischen Experiments. Die auditorischen Reize sind jeweils drei Töne mit Halbtonabstand aus zwei verschiedenen Frequenzbereichen (vgl. Tab. 4.1). Im hochfrequenten Bereich sind die Töne zufällig angeordnet, im tieffrequenten Stream folgen sie aufsteigenden Tripel. Der Pegel jedes dritten Tones ist erhöht (weiß: 65 dB, schwarz: 80 dB SPL), so dass auch in diesem Experiment ein stream-übergreifendes Intensitätsmuster entsteht. Zusätzlich wird der Abstand der Frequenzbereiche zwischen *eng* (oben), *mittel* (Mitte) und *weit* (unten) variiert und jede der auditorischen Stimuluskonfigurationen mit verschiedenen visuellen Stimuli gekoppelt. Die Quadrate sollen dabei ein Zwei-Stream-Perzept (*segregiert*), die Kreise ein Ein-Stream-Perzept (*integriert*) verstärken (gepunktete Linie). Unter Einschluss der auditorischen Präsentationen ohne visuelle Stimuli (rechts) entstehen neun experimentelle Bedingungen.

entspricht dem ambigen Bereich. Hier können die Stimuli, abhängig von der Aufmerksamkeit und Intention der Probanden, entweder als ein oder zwei auditorische Objekte wahrgenommen werden.

Jede der beschriebenen Stimulationsbedingungen enthielt die Tripel im tieffrequenten Stream. Um Unterschiede zwischen dem Erkennen des Drei-Ton-Musters in beiden Frequenzlagen auswerten zu können, wurde jede der Kombinationen verdoppelt, indem die Töne des tieffrequenten Streams randomisiert und die des hochfrequenten Streams als Tripel präsentiert wurden. Jede Kombination enthielt 1620 auditorische Stimuli, die eine Präsentationszeit von drei Minuten benötigten.

4.3.4 Messablauf

Die Probanden hatten die Aufgabe, sowohl die visuelle als auch die auditorische Stimulation aufmerksam zu verfolgen und ihren vordergründigen Höreindruck durch einen Tastendruck anzuzeigen. In einem 2-AFC (Alternative Forced Choice)-Verfahren konnte zwischen dem Perzept einer Drei-Ton-Folge (segregiert) und einer Zwei-Ton-Folge (integriert) gewählt werden. Wiederholte Tastendrucke wurden erlaubt. Wenn sich der Höreindruck während der Stimuluspräsentation änderte, sollten die Probanden zur entsprechend anderen Antworttaste wechseln. Zu Beginn der Messung wurden zwei Trainingsdurchläufe mit den Bedingungen *weit-segregierend* und *eng-integrierend* durchgeführt, also Bedingungen, die jeweils ein eindeutiges auditorisches Perzept, das mit den visuellen Stimuli übereinstimmte, erzeugen sollten. Wenn die Probanden die korrekte Taste betätigten, wurde mit dem Experiment fortgefahren. Anderenfalls wurde die Trainingseinheit einmalig wiederholt und die Probanden bei erneuter Falschantwort von der Studie ausgeschlossen. Für die Aufzeichnung der Entscheidungen wurde ebenfalls PRESENTATION verwendet.

Die Probanden wurden in einem akustisch abgeschirmten Raum bequem platziert. Nach den zwei Trainingsdurchläufen wurden die Stimulus-Kombinationen präsentiert. Dabei wurde jede Kombination in drei Teile von je einer Minute Dauer unterteilt und alle 54 Teile zufällig gemischt. Zwischen jeder Präsentation wurden kurze Pausen eingefügt, deren Länge die Probanden selbst bestimmen konnten. Eine längere Pause von etwa zehn Minuten wurde eingefügt, nachdem die Hälfte der Präsentationen erfolgt war. Einschließlich der Pausen betrug die gesamte Messzeit etwa 80 Minuten.

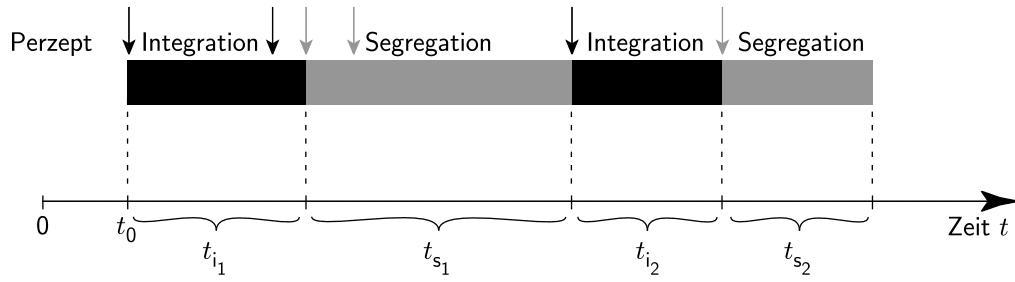


Abb. 4.5: Antwortschema psychoakustischen Messung. Wenn sich das Perzept der Probanden änderte, betätigten sie die entsprechende Taste (Pfeile). Die Zeitintervalle zwischen der Betätigung *verschiedener* Tasten wurden dem jeweiligen Perzept zugeordnet.

4.3.5 Datenverarbeitung und -analyse

Die Antworten der Probanden wurden offline analysiert. Dabei wurde die Dauer eines jeden Perzepts, *segregiert* oder *integriert*, für jede Kombination als Zeit zwischen zwei verschiedenartigen Antworten gemessen. Diese Zeitintervalle t_{i_m} bzw. t_{s_n} bilden – separat aufsummiert – die Gesamtzeit, in der jedes der beiden Perzepte angezeigt wurde:

$$t_i = \sum_m t_{i_m} \quad (4.5)$$

$$t_s = \sum_n t_{s_n} \quad (4.6)$$

Die Gesamtstimulationszeit (Messzeit) pro Kombination sei t_M , die Gesamtantwortzeit t_R ergibt sich aus der Summe der perzeptspezifischen Antwortzeiten

$$t_R = t_s + t_i \quad (4.7)$$

Die relativen Antwortzeiten τ_s und τ_i wurden aus t_i und t_s berechnet:

$$\tau_s = \frac{t_s}{t_R} \quad (4.8)$$

$$\tau_i = \frac{t_i}{t_R} \quad (4.9)$$

Auch die Zeit bis zur ersten Betätigung einer Taste (Initialzeit) t_0 wurde bestimmt und mit einer ANOVA mit den Faktoren *visuelle Stimulation* (Stufen: integrierend, segregierend, auditorisch allein) und *Frequenzabstand* (Stufen: eng, mittel, weit) untersucht.

Durch eine ANOVA bei Messwiederholungen wurde dann ermittelt, ob die Lage des Tripels (hoch- oder tieffrequenter Stream) einen Einfluss auf die Zeit t_i hat. Da sich herausstellte, dass kein Einfluss der Lage vorliegt, wurden für die folgenden Auswertungen die Zeiten t_i und t_s für gleiche Kombinationen, die sich nur in der Lage des Tripels unterschieden, zusammengefasst und weiter analysiert.

Für weitere Analysen wurden die relativen Zeiten τ_i und τ_s berechnet, jedoch nur τ_i weiter verwendet, da beide Größen komplementär sind.

Die Unterschiede zwischen den Kombinationen aus Frequenzabstand und visueller Stimulation wurden zunächst mit einer Zwei-Wege-ANOVA mit den Faktoren *Frequenzabstand* (eng, mittel, weit) und *visuelle Stimulationsbedingung* (integrierend, segregierend, auditorisch allein) analysiert. Daran anschließend wurden geplante Vergleiche verwendet, um die Unterschiede von τ_i zwischen den visuellen Stimulationen und Frequenzabständen zu untersuchen. Hierzu wurden zweiseitige t -Tests für verbundene Stichproben angewendet, deren Signifikanzniveau durch Bonferroni-Korrektur für Mehrfachvergleiche angepasst wurde.

4.4 MEG-Studie im ambigen Bereich

Diese Studie ist eine Ergänzung der EEG-Studie im ambigen Bereich (vgl. Kap. 4.2). Es soll die dort gewonnenen Ergebnisse reproduzieren und mit der Methode des MEG eine genauere Quellenlokalisierung, als das bei der EEG-Messung möglich ist, herbeiführen. Die Messungen fanden in den Labors der Abteilung für Neurologie II der Otto-von-Guericke-Universität Magdeburg und in Kooperation mit der Abteilung für Nicht-Invasive Bildgebung des Leibniz-Instituts für Neurobiologie, Magdeburg statt.

4.4.1 Probanden

Vierzehn normalhörende, rechtshändige Erwachsene (sieben davon männlich) im Alter von 18–36 Jahren nahmen am Experiment teil und unterzogen sich zuvor einem audiometrischen Test (Reintonaudiogramm (Luftleitung) mit den Frequenzen 0,125 kHz, 0,25 kHz, 0,5 kHz, 1 kHz, 1,5 kHz, 2 kHz, 3 kHz, 4 kHz, 6 kHz und 8 kHz). Sie konnten an der Studie teilnehmen, wenn ihre Ruhehörschwelle um nicht mehr als 10 dB erhöht war. Zur Auswertung konnten nur die Daten von dreizehn Probanden herangezogen werden, da bei einem Proband starkes Rauschen im MEG-Signal auftrat.

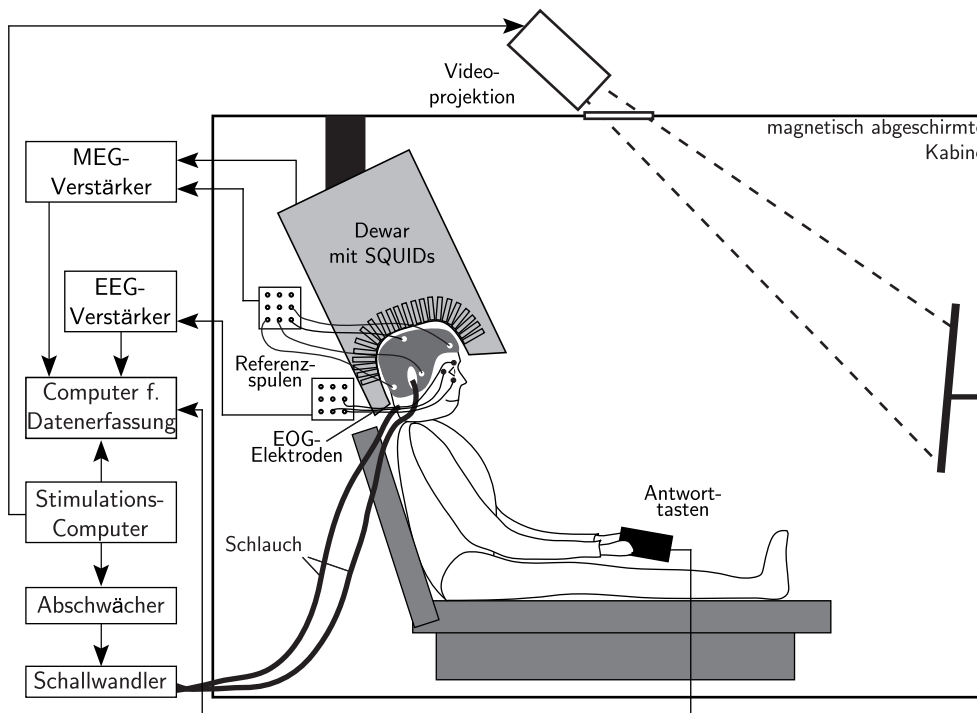


Abb. 4.6: Schematische Versuchsanordnung zur Messung von auditorisch und visuell evozierten Magnetfeldern sowie zur psychometrischen Registrierung von Antworten. Mit einem Computer werden die Stimuli erzeugt und in einem definierten Paradigma ausgegeben. Die visuellen Stimuli werden direkt zur Videoprojektion, die auditorischen Reize zunächst zu einem Abschwächer geleitet (zur Pegeljustierung) und dann über Schläuche und Ohrpasstücke appliziert. Der Proband sitzt in einer schallisolierten, magnetisch abgeschirmten Kabine. Sein MEG wird über die in einem Dewar befindlichen SQUIDs gemessen und mit einem Computer registriert. Dieser zeichnet auch die Trigger-Signale auf. Zusätzlich werden Referenzspulen am Kopf der Probanden befestigt, um die Lage des Kopfes relativ zur SQUID-Matrix bestimmen zu können. Um die Augenartefakte kontrollieren zu können, wird auch das EOG abgeleitet.

4.4.2 Stimuli

Wie in der EEG-Studie (vgl. Kap. 4.2) wurden die auditorischen Stimuli T3–5 in Tripel und H3–5 pseudorandomisiert präsentiert, um ein *per se* ambiges Perzept zu erzeugen. Kombiniert mit den im Abstand von 1 m auf eine Aluminiumplatte projizierten visuellen Stimuli Q1–3 (*segregierend*) beziehungsweise K1–2 (*integrierend*) wurden sie den Probanden mit der Software PRESENTATION⁹ dargeboten. Die MEG-Messungen wurden in einer magnetisch abgeschirmten Kabine durchgeführt. Um Störungen durch die von den Schallwandlern generierten Magnetfelder zu vermeiden, wurden die akustischen Stimuli außerhalb der Kabine generiert und über etwa 6 m lange Kunststoffschläuche zu dem Probanden geleitet. In der Nähe der Ohren wurde der Schlauchradius verringert und der dünnere Schlauch jeweils über Adapter mit dem Gehörgang verbunden. Die durch die Schläuche verursachte Zeitverzögerung der Stimuluspräsentation betrug 20 ms und wurde bei der Datenanalyse berücksichtigt.

Der Pegel der Stimuli betrug 80 dB SPL, wobei jeder dritte Ton (stream-übergreifend) mit einer um 15 dB erhöhten Intensität präsentiert wurde. Bei einigen Probanden wurden, da sie die Standardpegel als unangenehm empfanden, reduzierte Schallintensitäten von 70 dB SPL beziehungsweise 85 dB SPL verwendet.

4.4.3 Bedingungen

Die fixierte Lagerung der Probanden unter der MEG-Haube limitierte die zur Verfügung stehende Messzeit, so dass nur zwei Bedingungen aus der EEG-Studie (vgl. Kap. 4.2) verwendet wurden. Dies waren die Kombinationen von T3–5 mit H3–5 und entweder Q1–3 (*segregierend*) oder K1–2 (*integrierend*). Auch hier sollten die unterschiedlichen visuellen Stimuli die automatische auditorische Organisation (und damit die MMN) unterschiedlich beeinflussen.

Ohne Information über die Struktur der auditorischen und visuellen Stimuli sollten die Probanden die visuelle Stimulation verfolgen und gelegentlich präsentierte Target-Stimuli (rote Kreise bzw. Quadrate) durch Betätigen einer Antworttaste mit ihrem rechten Zeigefinger anzeigen.

⁹Neurobehavioral Systems Inc., Albany, USA

4.4.4 Messablauf

Jede Bedingung wurde in vier zusammenhängende Blöcke unterteilt, die jeweils 55 Devianten enthielten. Bei der Hälfte der Probanden wurden die Blöcke der Bedingung *integrierend* zuerst präsentiert, bei der anderen Hälfte umgekehrt. Da jeder Block eine Dauer von vier Minuten hatte, betrug die Netto-Messzeit etwa 32 Minuten. Durch Pausen zwischen den Blöcken und die Präparation der Probanden war ein Zeitaufwand von etwa 1,5 Stunden erforderlich.

4.4.5 MEG-Messungen

Die Probanden wurden in einer elektromagnetisch abgeschirmten, schallisolierten Kabine auf einer Liege platziert. Vor der Messung wurde die Kopfform der Probanden unter Verwendung eines 3SPACE FASTRAK-SYSTEMS¹⁰ digitalisiert. Dabei wurden zunächst die anatomischen Punkte Nasion, Inion, linker und rechter präaurikulärer Punkt sowie die Elektrodenposition Cz vermessen, dann die Position der Referenzspulen und schließlich die Kopfform durch Abfahren derselben mit dem Sender des Messsystems digitalisiert (vgl. Abb. 4.3).

Das MEG-Signal wurde kontinuierlich unter Verwendung eines 148-Sensoren Ganzkopf-Magnetometers¹¹ mit DC-Magnetometern aufgezeichnet. Es wurde analog gefiltert (100 Hz-Tiefpass) und mit einer A/D-Rate von 1017 Hz gespeichert. Bereits während der Messung wurde eine Rauschreduktion angewendet: An acht Referenzkanälen, die hinreichend weit vom Kopf des Probanden entfernt sind, wurde das von der Umwelt eingestreuete Signal gemessen und gewichtet vom an der Kopfhaut gemessenen Signal subtrahiert (Robinson 1989). Simultan wurde auch das HEOG und VEOG mit einem EEG-Verstärker¹² aufgezeichnet, die Elektrodenimpedanzen betragen weniger als 5 k Ω .

4.4.6 Analyse der MEG-Daten

Die offline-Datenanalyse wurde mit der Software BESA¹³ durchgeführt. Zunächst wurde das MEG automatisch auf Artefakte untersucht und diese entfernt, wenn die dort auftre-

¹⁰Polhemus, Colchester, VT, USA

¹¹BTi Magnes 2500 WH, 4-D Neuroimaging, San Diego, USA

¹²SynAmps, Neuroscan, El Paso, TX, USA

¹³MEGIS Software GmbH, Gräfelfing, Deutschland

tende Signalamplitude auffällig erhöht war (10–15 % über der mittleren Signalamplitude). Dies führte zur Verwerfung von etwa 15 % der Epochen. Das Signal wurde digital gefiltert (1–15 Hz Bandpass, 24 dB/Oktave) und dann für jeden MEG-Kanal und beide Stimulationsbedingungen relativ zu einem Stimulus-Marker (Standard und Deviant) in Epochen unterteilt, die 100 ms vor Beginn des Markers begannen und eine Länge von 600 ms hatten. Standard- und Deviantepochen wurden wie beim EEG-Experiment gemittelt und daraus die Differenzwerte ermittelt.

Um auditorisch und visuell evozierte Magnetfelder abgrenzen zu können, wurde eine zusammenhängende Auswahl aus den zur Verfügung stehenden 148 Sensoren getroffen, die die unterschiedlichen Modalitäten repräsentieren. Dies waren je zwölf temporal gelegene Sensoren über dem auditorischen Kortex der linken bzw. rechten Hemisphäre und zur Analyse der visuell evozierten Magnetfelder ein occipital über dem visuellen Kortex gelegenes Ensemble aus 27 Sensoren, bei denen deutliche visuell evozierte Magnetfelder erkennbar waren (vgl. Abb. 4.7).

Mit *t*-Tests für verbundene Stichproben wurde mit der getroffenen Auswahl aus den Sensoren nach Latenzbereichen gesucht, an denen Standard- und Deviantstimuli unterschiedliche Magnetfelder evozierten. In diesen Latenzintervallen wurden in einem Zeitfenster von 40 ms um das Amplitudenmaximum die individuellen Amplituden von Standard- und Deviantverläufen gemittelt und für die weitere Analyse verwendet. Separat für die gefundenen Intervalle wurde durch Varianzanalysen mit den Faktoren *Sensor* (24 bzw. 27 Stufen), *Stimulustyp* (Stufen: Standard, Deviant) und *Bedingung* (Stufen: integrierend, segregierend) überprüft, in welcher Modalität signifikante Unterschiede auftraten.

Die signifikanten Differenzen zwischen Standard- und Deviantantwort (Gruppenmittel) der *segregierend*-Bedingungen wurden mit einer Quellenanalyse weiter untersucht. Die Analyse unter Verwendung aller gemessenen MEG-Kanäle mittels BESA beruht auf einem sphärischen Kopfmodell mit einem Radius von 84,33 mm. Da die Zeitintervalle mit signifikanten Unterschieden zwischen Standard- und Deviantantwort der temporalen und occipitalen Sensorengruppe vergleichbar waren, wurde das in der temporalen Gruppe bestimmte Zeitintervall der Quellenanalyse zu Grunde gelegt. Als Quellen wurden ECD-Paare (equivalent current dipoles) postuliert, die symmetrisch zur Medianebene in beiden Hemisphären liegen. Die genaue Positionsbestimmung erfolgte danach durch Anwendung eines iterativen Algorithmus, bis Quellenmodell und Messwerte signifikant übereinstimmten. Auch für die individuellen Differenzkurven wurde eine Quellenanalyse durchgeführt. Für das Kopfmodell wurden die gemessenen Punkte auf der Kopfhaut verwendet. Die

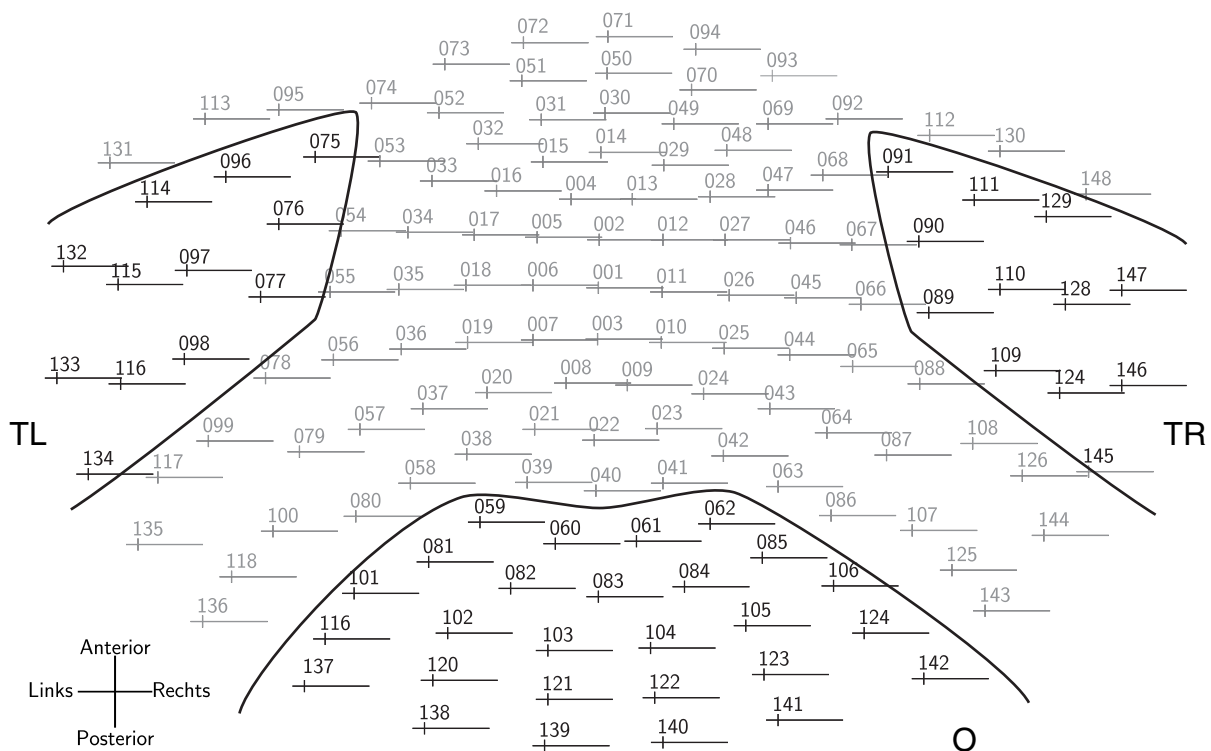


Abb. 4.7: Sensoren des Magnetometers. Für die Analyse der auditorisch evozierten Magnetfelder wurde eine temporal links bzw. rechts gelegene Sensorengruppe (TL, TR) und für die Analyse der visuell evozierten Magnetfelder eine occipital gelegene Auswahl (O) der Sensoren verwendet.

Anzahl der ECD und ihre Symmetrie wurde vom Gruppenmittelwert übernommen und für jeden Probanden die Lokalisation der Quellen iterativ bestimmt.

4.5 EEG-Studie im nicht-ambigen Bereich

Um die Sensitivität der auditorischen Organisation für visuelle Stimuli in Bereichen mit einem *per se* eindeutigen auditorischen Perzept zu untersuchen, wurde eine weitere EEG-Studie mit auditorischen und visuellen Stimuli durchgeführt. Alle Messungen fanden in den Labors der Abteilung für Experimentelle Audiologie und Medizinische Physik an der Medizinischen Fakultät der Otto-von-Guericke-Universität Magdeburg statt.

Wesentliche Teile der Methodik dieser Studie entsprechen der EEG-Studie im ambigen Bereich (vgl. Kap. 4.2) und der psychoakustischen Studie (vgl. Kap. 4.3). Im Folgenden sind deshalb nur die Eckdaten und wesentlichen Veränderungen zusammengefasst.

4.5.1 Probanden

Fünfzehn normalhörende Erwachsene (fünf davon männlich), einer davon linkshändig, im Alter von 18–28 Jahren nahmen am Experiment teil und unterzogen sich zuvor einem audiometrischen Test (Reintonaudiogramm (Luftleitung) mit den Frequenzen 0,125 kHz, 0,25 kHz, 0,5 kHz, 1 kHz, 1,5 kHz, 2 kHz, 3 kHz, 4 kHz, 6 kHz und 8 kHz). Sie konnten an der Studie teilnehmen, wenn ihre Ruhehörschwelle um nicht mehr als 10 dB erhöht war. Die Daten eines Probanden wurden auf Grund eines unverhältnismäßig hohen Störanteils im EEG von der Auswertung ausgeschlossen.

4.5.2 Stimuli

In diesem Experiment wurden die auditorischen Stimuli T1–6 und H1–6 verwendet und verschieden mit den visuellen Stimuli Q1–3 und K1–2 kombiniert (vgl. Tab. 4.4). Die auditorischen Stimuli wurden mit einem konstanten zeitlichen Abstand (SOA) von 110 ms abwechselnd zwischen einem hoch- und einem tiefrequenten Ton über Einsteckhörer¹⁴ mit der Software PRESENTATION¹⁵ dargeboten. Auch in diesem Experiment waren die Töne

¹⁴E-A-RTONE 3A, Auditory Systems, Indianapolis, IN, USA

¹⁵Neurobehavioral Systems Inc., Albany, USA

Tab. 4.4: Bedingungen des EEG-Experiments mit Stimuli außerhalb des ambigen Bereichs.

Bedingung	Auditorische Stimuli					Visuelle Stimuli		
	geordnet			random.		Q	K	Video-film
	T1-3	T3-5	T4-6	H1-3	H4-6			
eng-auditorisch allein			x	x				x
eng-segregierend			x	x		x		
weit-auditorisch allein	x				x			x
weit-integrierend	x				x		x	
1 Stream-auditorisch allein		x						x

der tieffrequenten Stimulationen in aufsteigenden bzw. fallenden Tripel angeordnet, um eine MMN auszulösen. Zudem wurde stream-übergreifend jeder dritte Ton mit einem um 15 dB erhöhten Pegel präsentiert.

4.5.3 Bedingungen

Auditorische und visuelle Sequenzen wurden, verschieden kombiniert, in fünf verschiedenen Bedingungen präsentiert. Darin enthalten waren drei Kontrollbedingungen. Eine Übersicht zeigt die Tabelle 4.4.

In den Bedingungen ohne visuelle Stimulation wurde entsprechend des Frequenzabstands die automatische Integration *eng-auditorisch allein* beziehungsweise Segregation *weit-auditorisch allein* erwartet. Der Nachweis über das automatische Perzept sollte auch hierbei über die MMN erfolgen, was einen Rückschluss auf das Erkennen der umgekehrten Tripel zulässt.

In den Bedingen mit visueller Stimulation sollte, bei erfolgreicher Beeinflussung des auditorischen Perzepts durch die visuellen Stimuli, im Gegensatz zu den *auditorisch allein*-Bedingungen die MMN verschwinden (*weit-integrierend*) oder evoziert werden (*eng-segregierend*).

Als Kontrollbedingung für die MMN diente die *1 Stream-auditorisch allein*-Bedingung. Da hier keine visuellen Stimuli enthalten waren und zudem nur ein Stream präsentiert wurde, sollte von den absteigenden Devianten eine deutliche MMN evoziert werden.

Auch in diesem Experiment hatten die Probanden keine Information über die Struktur der auditorischen und visuellen Stimuli. Sofern visuelle Stimuli präsentiert wurden, sollten die Probanden seltene Target-Stimuli (rote Kreise bzw. Quadrate) durch Betätigen einer Antworttaste anzeigen. Wurden keine visuellen Stimuli präsentiert, bestand die Aufgabe darin, einen lautlos präsentierten Videofilm mit Untertiteln zu verfolgen.

4.5.4 Messablauf

Die *auditorisch allein*-Bedingungen wurde in allen Fällen als erste präsentiert, gefolgt von den verbleibenden zwei Bedingungen mit visueller Stimulation in über die Probanden wechselnder Reihenfolge. Jede Bedingung wurde in vier zusammenhängende Blöcke unterteilt, die jeweils 55 Devianten enthielten. Ein Block hatte eine Dauer von etwa vier Minuten, insgesamt wurden 220 Devianten pro Bedingung präsentiert.

Einschließlich der Präparationszeit und der Pausen zwischen den Blöcken und Bedingungen betrug der Zeitaufwand der gesamten Messung etwa 2,5–3 Stunden.

4.5.5 EEG-Messungen

Das EEG wurde kontinuierlich unter Verwendung eines 32-kanaligen Verstärkers¹⁶ aufgezeichnet. Dabei wurde ein analoger Bandpass von 0,3–200 Hz (12 dB/Oktave) und eine A/D-Rate von 1 kHz verwendet. Die Elektroden wurden gemäß des internationalen 10-20-Systems (vgl. Kap. 4.2.5) mit einer Elektrodenhaube¹⁷ an neunzehn Positionen¹⁸ der Schädeloberfläche angebracht. Zusätzlich wurden am linken und rechten Mastoid Elektroden befestigt, die Referenzelektrode befand sich auf der Nasenspitze. Das VEOG wurde mit einer bipolaren Elektrodenkonfiguration zwischen Fp2 und einer unter dem rechten Auge befindlichen Elektrode aufgezeichnet. Die Elektrodenimpedanzen betragen weniger als 5 k Ω , allerdings konnten bei zwei Probanden an einigen Elektroden Impedanzen von 11 k Ω nicht unterschritten werden.

¹⁶SynAmps, Neuroscan, El Paso, TX, USA

¹⁷Electro-Cap International, Inc., Eaton, OH, USA

¹⁸Fz, Cz, Pz, Fp1, Fp2, F7, F8, F3, F4, C3, C4, T3, T4, T5, T6, P3, P4, O1, O2

4.5.6 Analyse der EEG-Daten

Die Analyse dieser EEG-Daten erfolgte weitgehend analog zur ersten EEG-Messung (vgl. Kap. 4.2.6). Nach Artefaktdetektion, -entfernung und Filterung wurde das EEG relativ zu den ersten Stimuli der aufsteigenden beziehungsweise fallenden Tripel segmentiert. Die Artefakte im EEG, die meist durch Augenbewegungen verursacht wurden, führten bei ausgewerteten Datensätzen zu einem Ausschluss von etwa 17 % aller Epochen. Die EKP wurden als Mittelwerte der Standard- und Deviant-Epochen berechnet, anschließend erfolgte deren Differenzbildung für die Beurteilung der MMN. Für beide Epochentypen wurden die Gruppenmittelwerte separat berechnet.

Zur Berechnung der MMN-Amplitude wurden im Fenster von ± 40 ms um einen sich aus dem Gruppenmittelwert ergebenden Amplitudenmaximum im Intervall von 100–200 ms nach Reizeinsatz die individuellen Amplituden gemittelt. In den Bedingungen *eng-auditorisch allein*, *weit-auditorisch allein*, und *eng-segregierend* war dies das Amplitudenmaximum der Bedingung *weit-auditorisch allein* (195 ms), in der Bedingung *1 Stream-auditorisch allein* die Latenz von 155 ms und in der Bedingung *weit-integrierend* die Latenz von 163 ms.

Eine Drei-Wege-ANOVA mit den Faktoren *Elektrode* (Stufen: Fz, Cz, Pz, LM, RM), *Stimulustyp* (Stufen: Standard, Deviant) und *Bedingung* (5 Stufen) wurde durchgeführt, um deren Einfluss auf die MMN-Amplitude zu testen. *Post-hoc*-Tests mit Bonferroni-Korrektur für Mehrfachvergleiche wurden angewendet, um die gefundenen Effekte zu spezifizieren. Schließlich wurden die Potentialverteilungen auf der Kopfhaut sowie die SCD berechnet und grafisch dargestellt. Die Quellenanalyse wurde für das Differenzpotential aller Bedingungen (außer *eng-auditorisch allein*) analog zu Kap. 4.2.6 durchgeführt.

4.6 Analyse der Aufmerksamkeit

Für die Experimente mit visueller Stimulation wurden die Reaktionen der Probanden auf den visuellen Targetreiz mit Methoden der Signalentdeckungstheorie analysiert. Sie ist zur Messung von schwachen Signalen, oder – wie in dieser Arbeit – der Beurteilung von Detektionsleistungen geeignet (Green und Swets 1966; Swets 1964). Das zu entdeckende Signal kann auch ein Targetreiz in einer Folge anderer (Standard-) Reize sein. Wenn Probanden diesen detektieren sollen, ergeben sich vier Antwortkombinationen, die wie folgt klassifiziert werden:

	Target-Reiz	Standard-Reiz
Geantwortet	Treffer (Hit)	Fehlalarm (False Alarm),
Nicht geantwortet	Verpasser (Miss)	korrekte Ablehnung (Correct Rejection)

Ein Maß für die Detektionsleistung ist die *Sensitivität* d' . Sie wird aus den relativen Häufigkeiten der Treffer p_h und Fehlalarme p_{fa} berechnet. Nach Anwendung der z -Transformation erhält man die Sensitivität d' durch Differenzbildung:

$$d' = z(p_h) - z(p_{fa}) \quad (4.10)$$

Man geht bei der Signalentdeckungstheorie von einer normalverteilten Wahrscheinlichkeit aus, ein nicht vorhandenes Signal im Rauschen zu detektieren (Fehlalarm). Bei Vorhandensein eines Signals verschiebt sich die Normalverteilung nach rechts. Die z -Transformation hierbei liefert zu den gegebenen Raten (Treffer und Fehlalarm) die entsprechenden z -Werte der Standardnormalverteilung. Auftretende Trefferraten $p_h = 0$ werden ersetzt durch:

$$p_h = 2^{-\frac{1}{n}} \quad (4.11)$$

Dabei ist n die Anzahl der möglichen Treffer (Targets). Fehlalarm-Raten $p_{fa} = 1$ werden mit Hilfe der Anzahl der möglichen Fehlalarme m ersetzt durch:

$$p_{fa} = 1 - 2^{-\frac{1}{m}} \quad (4.12)$$

In den Experimenten der vorliegenden Arbeit wurde eine Antwort wurde als Treffer gewertet, wenn sie in den Zeitbereich von 0,1–1 s nach Beginn des Targetreizes fiel. Der Sensitivitätsindex d' wurde in den entsprechenden Experimenten für jeden Probanden berechnet.

Zudem wurde die mittlere Antwortzeit t_r der Probanden für jeden Reiz ermittelt. Durch eine Varianzanalyse (ANOVA) mit dem Faktor *Experiment* wurde überprüft, ob sich die Mittelwerte der Antwortzeiten zwischen den Experimenten unterscheiden. Die gefundenen Unterschiede wurden durch *post hoc* durchgeführte Tukey-Tests (vgl. Kap. 4.2.6) spezifiziert.

Kapitel 5

Experimentelle Ergebnisse

5.1 EEG-Studie im ambigen Bereich

Abb. 5.1 zeigt an ausgewählten Elektrodenpositionen die Gruppenmittelwerte der Amplituden für Standard- und Deviantpotentiale. Diese Mittelwerte und die dazugehörigen Differenzen zwischen Standard- und Deviantpotential im für die MMN relevanten Latenzbereich sind in Tab. 5.1 dargestellt. Generell ist hier die Amplitude der ereigniskorrelierten Potentiale (EKP) in Folge der hohen Stimulationsrate verhältnismäßig klein und ihre Zuordnung zu den einzelnen Stimuli nicht immer möglich (Javitt *et al.* 1993; Sussman *et al.* 1999). Zudem tritt eine Überlagerung auditorischer und visuell evozierter EKP auf. Dennoch kann die von den Standard- und Deviantstimuli evozierte auditorische P1-Welle (aP1) in allen Bedingungen mit auditorischer Stimulation klar bei einer Latenz von etwa 50 ms identifiziert werden. Ein visuell evozierter N70-P100-Komplex (vN70-vP100, vgl. Kap. 3.6) kann im Latenzbereich um 50–150 ms nach Stimulus-Einsatz in allen Bedingungen mit visueller Stimulation an occipitalen Elektroden gefunden werden.

Die ANOVA an ausgewählten Positionen mit den Faktoren *Elektrodenposition* (Fz, Cz, Pz, Oz, F3, F4, FC1, FC2), *Bedingung* (5 Bedingungen) und *Stimulustyp* (Standard, Deviant) ergab einen Haupteffekt des *Stimulustyps* ($F(1,7) = 21,14, p < 0,01$) und eine Interaktion der Faktoren *Stimulustyp* und *Bedingung* ($F(4,28) = 4,24, p < 0,05$). Andere Haupteffekte und Interaktionen wurden nicht gefunden. Die angegebenen Signifikanzwerte verwenden stets die Greenhouse-Geisser-Korrektur, weil der Mauchly-Test eine signifikante Verletzung der Sphärizitätsvoraussetzung für den Faktor *Elektrodenposition* ($W(27) = 0,00, p < 0,01$) und die Interaktion von *Elektrodenposition* und *Stimulustyp* ($W(27) = 0,00,$

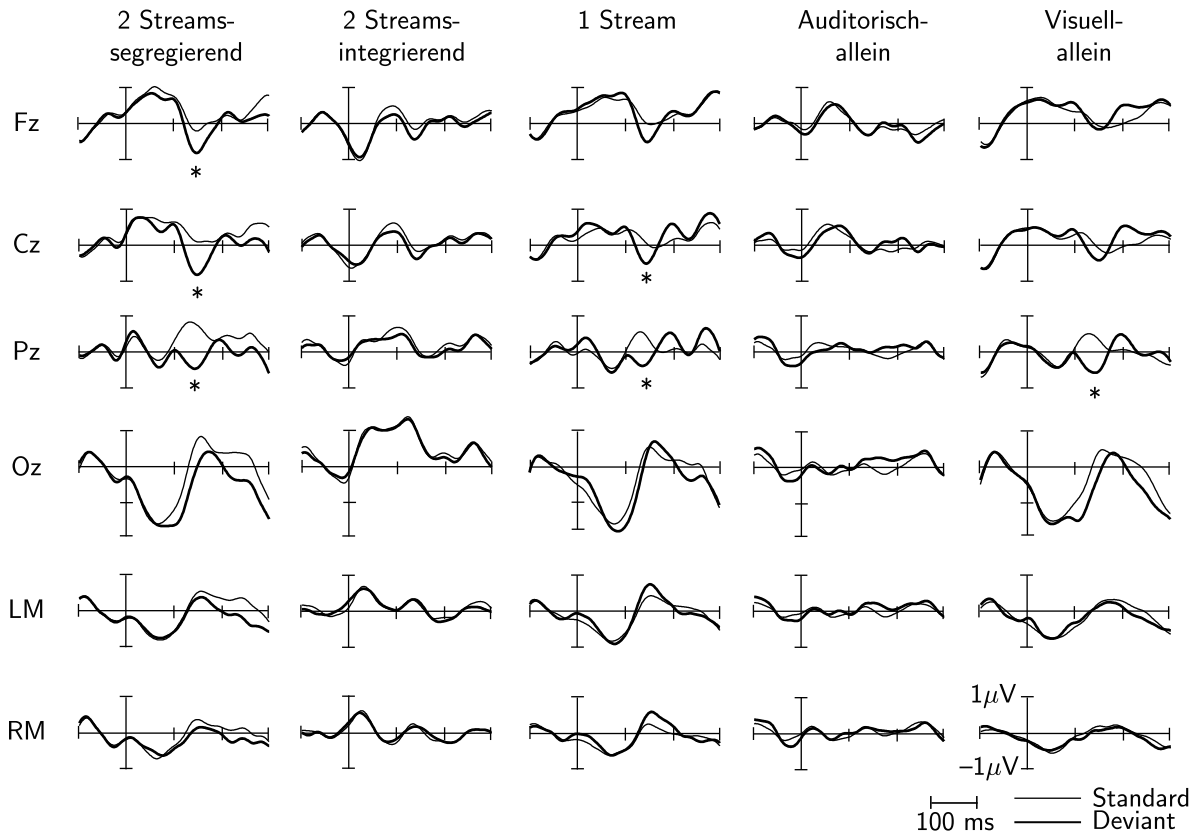


Abb. 5.1: Gruppenmittelwert ($n = 8$) der EKP, die durch den ersten Stimulus eines Tripels evoziert wurden, gemessen an den Elektrodenpositionen Fz, Cz, Pz, Oz und den Mastoiden. Signifikante Unterschiede zwischen Standard- (dünn) und Deviantpotential (dick) an der MMN-Latenz sind markiert (Bonferroni-korrigiertes $\alpha = 0,95$ für Mehrfachvergleiche).

Tab. 5.1: Gruppenmittelwerte (MW) und Standardabweichung (SD) für Standard-, Deviant- und Differenzpotential (in μV).

		Standard		Deviant		Differenz	
		MW	SD	MW	SD	MW	SD
2 Streams– segregierend	Fz	–0,04	0,753	–1,17	0,74	–1,13	0,67
	Cz	–0,29	0,71	–1,18	0,71	–1,47	0,95
	Pz	0,99	0,88	–0,72	0,78	–1,71	1,05
	Oz	0,36	1,24	–0,16	1,18	–0,52	1,49
	LM	0,38	0,68	0,37	0,91	–0,01	0,49
	RM	0,32	0,63	0,17	0,66	–0,15	0,71
2 Streams– integrierend	Fz	0,19	0,96	–0,53	0,98	–0,34	0,55
	Cz	–0,10	0,73	–0,40	1,08	–0,29	0,56
	Pz	0,28	1,01	–0,03	0,71	–0,32	0,62
	Oz	1,39	2,25	1,19	1,72	–0,20	0,65
	LM	0,39	0,89	0,39	0,66	0,00	0,71
	RM	0,29	0,88	0,24	0,47	–0,06	0,54
1 Stream	Fz	0,42	0,75	–0,65	1,21	–1,07	0,99
	Cz	0,24	0,71	–0,72	1,06	–0,96	0,78
	Pz	0,40	0,88	–0,65	0,92	–1,06	0,91
	Oz	–0,08	1,65	–0,05	1,19	0,13	1,26
	LM	0,11	0,83	0,57	0,84	0,46	0,60
	RM	–0,05	0,75	0,43	0,82	0,49	0,55
Auditorisch– allein	Fz	–0,22	0,68	–0,11	0,78	0,11	0,61
	Cz	–0,13	0,67	–0,01	0,94	0,12	1,05
	Pz	–0,05	0,46	0,15	0,80	0,20	1,21
	Oz	–0,12	0,57	0,35	0,73	0,47	1,13
	LM	0,02	0,38	0,21	0,57	0,19	0,58
	RM	0,07	0,45	0,19	0,74	0,12	0,77
Visuell– allein	Fz	0,08	0,82	–0,24	0,75	–0,31	0,80
	Cz	0,03	0,78	–0,50	1,09	–0,53	0,97
	Pz	0,45	0,96	–0,65	1,36	–1,10	0,97
	Oz	–0,08	1,33	–0,32	1,31	–0,24	1,74
	LM	0,02	0,41	0,23	0,56	0,20	0,55
	RM	0,17	0,54	0,17	0,74	0,00	0,58

Tab. 5.2: Ergebnisse der t -Tests zwischen Standard- und Deviantpotentialen an ausgewählten Elektrodenpositionen. Signifikante Unterschiede ($\alpha = 0,95$) sind mit (*) gekennzeichnet.

Elektrodenposition	2 Streams-segregierend	2 Streams-integrierend	1 Stream	Auditorisch-allein	Visuell-allein
Fz	-4,18 (*)	-1,73	-3,06	0,49	-1,16
Cz	-4,41 (*)	-1,47	-3,48 (*)	0,32	-1,56
Pz	-4,61	-1,46	-3,26 (*)	0,46	-3,20 (*)
F3	-6,04 (*)	-1,70	-3,00	0,76	-0,76
F4	-4,43 (*)	-1,12	-3,64 (*)	0,61	-1,47
FC1	-5,05 (*)	-1,64	-3,34 (*)	0,66	-0,82
FC2	-4,49 (*)	-1,54	-3,53 (*)	0,38	-1,16

$p < 0,01$) anzeigt. Der Haupteffekt des Stimulustyps zeigt einen Amplitudenunterschied zwischen Standard- und Deviantpotential an, der aber nur dann als MMN identifiziert werden kann, wenn die Deviant-Amplitude kleiner als die Standard-Amplitude ist. Um dies zu verifizieren, wurden einseitige t -Tests für verbundene Stichproben als geplante Vergleiche zwischen Standard- und Deviantpotential *post hoc* durchgeführt. Sie ergaben die in Tab. 5.2 angegebenen Resultate. Signifikante Unterschiede, die eine MMN anzeigen, treten in den Bedingungen *2 Streams-segregierend* und *1 Stream* an frontalen Elektroden auf, was hypothesenkonform ist. Die Bedingung *Visuell-allein* zeigt nur an der occipitalen Elektrode Pz signifikant verschiedene Potentiale, was auf eine visuell evozierte MMN schließen lässt. In den anderen Bedingungen sind keine Unterschiede zwischen Standard- und Deviantpotentialen signifikant. Wie erwartet, tritt dort keine MMN auf.

Die Differenzpotentiale (Deviant minus Standard) sind als Gruppenmittelwerte für jede Bedingung in Abb. 5.2 dargestellt. Die Negativierung (bzw. Positivierung an den Mastoid-Elektroden) im Latenzbereich um 140 ms wird als MMN identifiziert. Durch die audiovisuelle Stimulation kommt es zu einer Überlagerung von auditorischen und visuellen EKP. In der visuellen Kontrollbedingung (*Visuell-allein*) wird eine occipital ausgeprägte Negativierung gefunden, die jedoch frontal nicht sichtbar wird. In den Bedingungen *1 Stream* und *2 Streams-segregierend* ist die MMN deutlich ausgeprägt, erkennbar an negativen Potentialen an den frontalen Elektroden. Die übrigen Bedingungen hingegen weisen keine klare MMN auf. Nur in der *1 Stream*-Bedingung ist eine Inversion der Polarität an den Mastoiden erkennbar.

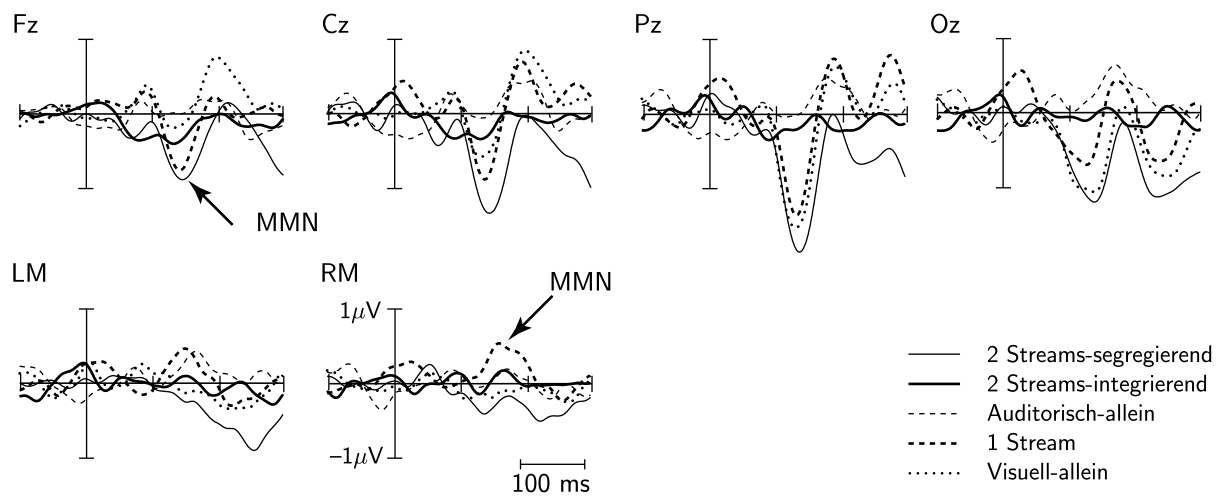


Abb. 5.2: Gruppenmittelwert ($n = 8$) der Differenzpotentiale der Stimulationsbedingungen *2 Streams-segregierend* (dünne Linie), *2 Streams-integrierend* (dicke Linie), *1 Stream* (dick gestrichelt), *Visuell-alone* (gepunktet) und *Auditorisch-alone* (dünn gestrichelt). Eine Negativierung bei einer Latenz um 140 ms wurde als MMN interpretiert. Eine Inversion der Polarität *1 Stream*-Bedingung an den Mastoiden in den Bedingungen *1 Stream* und *2 Streams-segregierend* erkennbar.

Tab. 5.3: Mittlere Differenz der Differenzpotentiale (in μV) zwischen den Bedingungen. Signifikante Unterschiede (*Post-hoc*-Test, Bonferroni-korrigiertes $\alpha = 0,95$) sind mit (*) gekennzeichnet.

Bedingung	2 Streams-segregierend	2 Streams-integrierend	1 Stream	Auditorisch-alone	Visuell-alone
2 Str.-seg.	—	0,73 (*)	0,37	1,15 (*)	0,65 (*)
2 Str.-int.	-0,73 (*)	—	-0,36	0,43 (*)	-0,07
1 Stream	-0,37	0,36	—	0,78 (*)	0,28
Aud.-allein	-1,15 (*)	-0,43 (*)	-0,78 (*)	—	-0,50 (*)
Vis.-allein	-0,65 (*)	0,07	-0,28	0,50 (*)	—

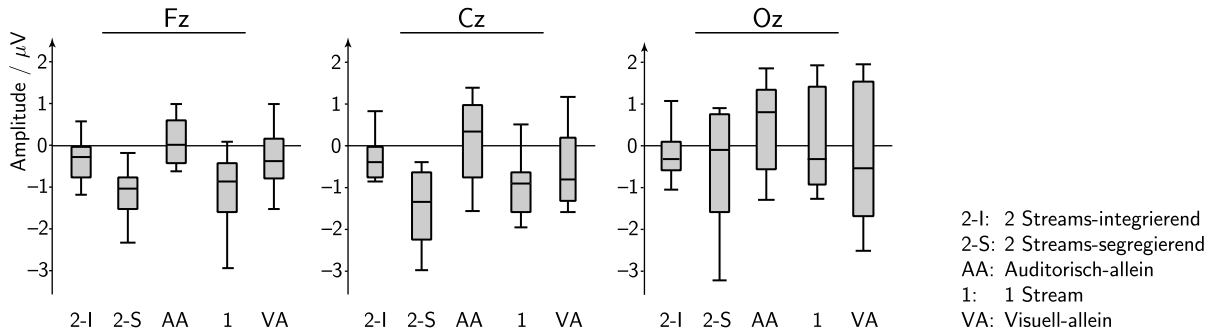


Abb. 5.3: Boxplots der MMN-Amplituden an den Elektrodenpositionen Fz (links), Cz (Mitte) und Oz (rechts) für alle Stimulationsbedingungen.

Die Analyse der MMN-Amplitude des Differenzpotentials wurde mit einer Zwei-Wege-ANOVA durchgeführt. Als Faktoren wurden die *Elektrodenposition* (Pz, Fz, Cz, Oz, F3, F4, Fc1, Fc2, LM, RM) und die Stimulationsbedingung verwendet. Beide rufen Haupteffekte hervor (Elektrodenposition: $F(1,9) = 6,10$, $p < 0,001$; Bedingung: $F(1,4) = 21,17$, $p < 0,001$) und weisen keine Interaktion auf.

Die *post-hoc*-Tests für den Haupteffekt der Elektrodenposition ergaben signifikante Unterschiede der MMN-Amplitude zwischen den frontalen Elektroden und den Mastoiden sowie keine signifikanten Unterschiede innerhalb beider Gruppen. Dadurch wird die Invertierung der Polarität zwischen frontalen Elektroden und den Mastoiden bestätigt. Ebenso sind die Unterschiede im Differenzpotential zwischen den linken und rechten Elektroden nicht signifikant. Zwischen den Elektrodenpositionen Pz und Oz wurde hingegen ein signifikanter Amplitudenunterschied gefunden. Homogene Untergruppen (nach Tukey, vgl. Kap. 4.2.6) sind die Gruppe der frontalen und zentralen Elektroden (Pz, Fz, Cz, Fc1, Fc2, F3, F4) sowie die Gruppe der Mastoid-Elektroden einschließlich Oz.

Der Effekt der Stimulationsbedingung wurde ebenfalls *post hoc* untersucht (vgl. Tab. 5.3). Dabei zeigt sich vor allem, dass der Amplitudenunterschied zwischen den Bedingungen *1 Stream* und *2 Streams-segregierend* nicht signifikant ist, wohingegen sich die Bedingung *2 Streams-segregierend* signifikant von den Bedingungen *2 Streams-integrierend* und *Visuell-allein* unterscheidet. Die Bedingung *Auditorisch-allein* unterscheidet sich von allen anderen Bedingungen signifikant. Abb. 5.3 zeigt exemplarisch die MMN-Amplituden an Fz, Cz, und Oz für alle Stimulationsbedingungen. Auch diese Ergebnisse bestätigen die Hypothese, wonach eine MMN evoziert wird, wenn eine Invertierung des Tripels detektiert wird.

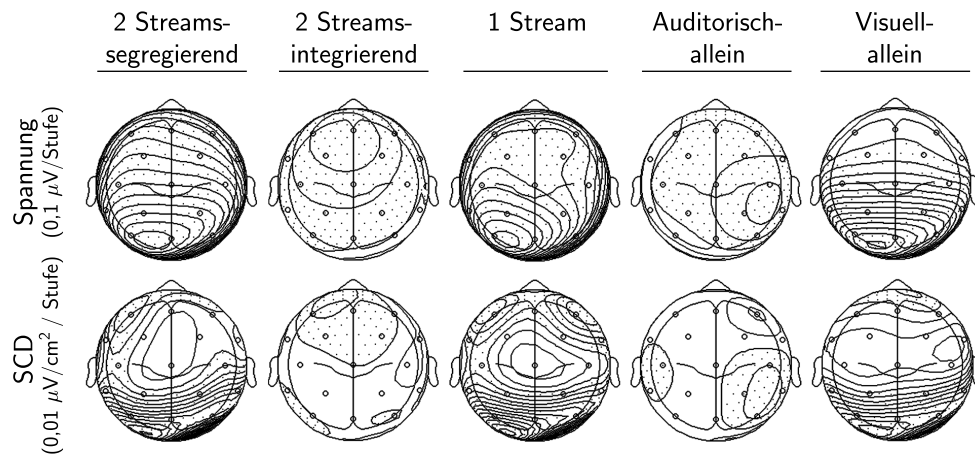


Abb. 5.4: Spannungsverteilung (Gruppenmittelwert) auf der Schädeloberfläche (oben) und SCD-Karten (unten) für alle Stimulationsbedingungen bei einer Latenz von 142 ms. Dargestellt sind die Äquipotentiallinien bzw. Linien gleicher Quellendichte. Gepunktete Flächen kennzeichnen negative Werte. Die SCD-Karten der Bedingungen *2 Streams-segregierend* und *1 Stream* weisen frontale und occipitale Quellen auf. Ausschließlich frontal gelegene Quellen sind in der Bedingung *Visuell-allein* erkennbar. Die übrigen Bedingungen weisen eine breite Verteilung ohne Foci auf.

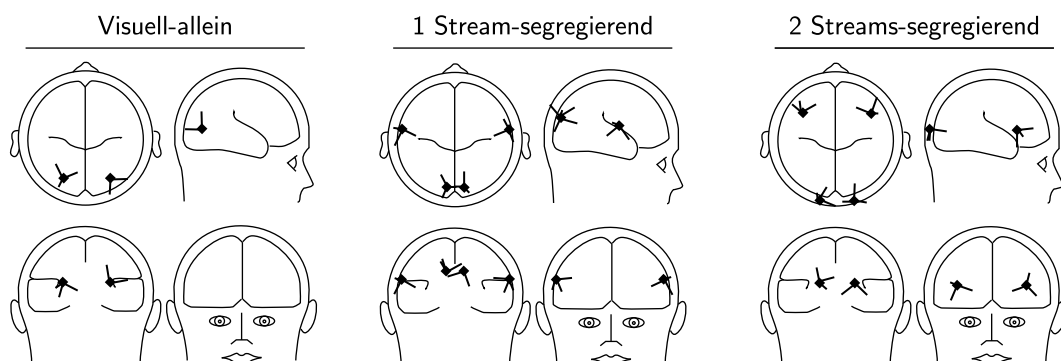


Abb. 5.5: Ergebnisse der Quellenanalyse mit BESA. Dargestellt sind Lage und Orientierung der regionalen Quellen, die das Differenzpotential der Bedingungen *1 Stream*, *Visuell-allein* und *2 Streams-segregierend* ergeben. In der Bedingung mit ausschließlicher visueller Stimulation sind occipitale Quellen erkennbar. Das Differenzpotential in den Bedingungen zusätzlicher auditorischer Stimulation kann unter zusätzlicher Annahme eines bilateral im auditorischen Kortex gelegenen Quellenpaares erklärt werden.

In Abb. 5.4 sind die mit BESA berechneten Spannungsverteilungen und SCD-Karten der Schädeloberfläche für jede Stimulationsbedingung im Differenzpotential 142 ms nach Stimulationsbeginn dargestellt. Die Bedingungen, in denen eine MMN auftrat (*2 Streams-segregierend* und *1 Stream*), weisen frontale Quellen mit Foci nahe den Elektrodenpositionen F3 und F4 auf. Diese lassen sich bilateralen Generatoren der auditorischen MMN zuordnen. Zusätzlich wurden occipital gelegene Foci der Quellendichte nahe den Positionen O1 und O2 gefunden. In der Bedingung *Visuell-allein* sind nur die occipital gelegenen Foci erkennbar, was auf visuell evozierte Potentiale schließen lässt. Die Bedingungen *Auditorisch-allein* und *2 Streams-segregierend*, in denen keine MMN evoziert wurde, weisen eine relativ gleichmäßige Verteilung der SCD über der gesamten Kopfhaut auf, die keine Foci erkennen lässt. Hierdurch wird die MMN in den Bedingungen *2 Streams-segregierend* und *1 Stream-segregierend* klar unterstützt und ihr Fehlen in den übrigen Bedingungen bestätigt.

Durch die Quellenanalyse (vgl. Abb. 5.5) wird die Identifizierung des negativen Differenzpotentials als auditorisch beziehungsweise visuell evozierte MMN ebenfalls unterstützt. In der *Visuell-allein*-Bedingung ergibt die Analyse der Hauptkomponenten (PCA) zwei Generatoren, die bilateral occipital im visuellen Kortex lokalisiert sind. Zusätzlich zu diesen Quellen tritt, wenn auch auditorisch stimuliert wurde, ein Generatorenpaar bilateral im Bereich der auditorischen Kortex auf (Bedingungen *2 Streams-segregierend* und *1 Stream*). In den Bedingungen *2 Streams-integrierend* und *Auditorisch-allein* konnten im Differenzpotential keine Generatoren identifiziert werden, da die Amplitude des Differenzpotentials $V = 0$ betrug.

5.2 Psychoakustische Studie

Alle Probanden, die an dieser Studie teilnahmen, waren nach maximal zwei Trainingsdurchläufen in der Lage, die gestellte Aufgabe zu erfüllen. Während der Stimuluspräsentation wurde in jedem Block von allen Probanden mindestens einmal eine Antworttaste betätigt. Die erste Antwort wurde im Mittel ($5,3 \pm 3,4$) s nach Blockbeginn registriert. In Abb. 5.6 sind die Mediane dieser absoluten Initialzeiten für die verschiedenen Bedingungen dargestellt. Sie unterschieden sich zwischen den Stimulationsbedingungen nicht. Für die *auditorisch allein*-Bedingungen wurde zusätzlich das erste von den Probanden angegebene Perzept ausgewertet. In der Bedingung mit weitem Abstand der Frequenzbereiche war dies zu 84 % ein 2-Stream-Perzept, bei engem Abstand wurde bei 80 %

der Erstantworten das 1-Stream-Perzept angegeben. In der Bedingung *mittel-auditorisch allein* wurde zu 74 % ein segregiertes Perzept als Erstantwort angegeben.

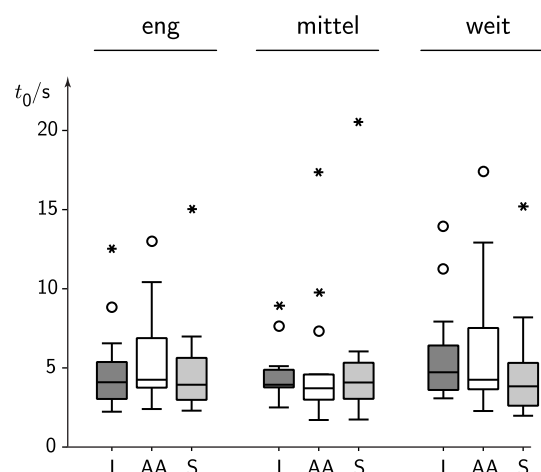
Die absoluten Zeiten für die Perzepte *integriert* und *segregiert* sind in Abb. 5.7 dargestellt. Obwohl die interindividuelle Streubreite meist groß ist, wird eine Abhängigkeit der Zeiten von der Stimulationsbedingung deutlich. Bei allen Bedingungen mit visueller Stimulation nimmt t_i mit zunehmendem Frequenzabstand der Bereiche ab. Besonders die Bedingungen mit ambigem Frequenzabstand führen zu weit variierenden Werten von t_i . Am größten ist diese bei integrierenden visuellen Stimuli, am kleinsten bei segregierenden Stimuli.

Eine ANOVA mit dem Faktor *Lage* (Tripel im hohen beziehungsweise tiefen Frequenzbereich) ergab keinen Effekt ($F(1,13) = 0,80$, $p = 0,395$). Somit konnten die Daten beider Lage-Bedingungen zusammengefasst weiter verwendet werden. Die folgenden Auswertungen beschränken sich auf die Analyse der relativen Zeit für ein integriertes Perzept τ_i , da die Dauer des segregierten Perzepts τ_s zu τ_i komplementär ist.

Die weitere Untersuchung von τ_i mittels einer Zwei-Wege-ANOVA mit den Faktoren *Frequenzabstand* (eng, mittel, weit) und *visuelle Stimulationsbedingung* (integrierend, segregierend, auditorisch allein) ergab hoch signifikante Haupteffekte des Frequenzabstands ($F(2,26) = 31,96$, $p < 0,001$) und der visuellen Stimulation ($F(2,26) = 14,04$, $p < 0,001$). Die Interaktion zwischen beiden Faktoren erwies sich ebenfalls als signifikant ($F(4,52) = 5,42$, $p < 0,01$).

Die Gruppenmittelwerte von τ_i für alle Stimulationsbedingungen sind in Abb. 5.8 dargestellt. Es zeigt sich, dass sowohl bei weitem und mittlerem Frequenzabstand abhängig von der visuellen Stimulation Unterschiede in der Wahrnehmung auftreten, wohingegen bei

Abb. 5.6: Absolute Initialzeiten für die Stimulationsbedingungen *integrierend* (I), *auditorisch allein* (AA) und *segregierend* (S). Dargestellt sind Median und Interquartilsabstand in den drei Bereichen des Frequenzabstands. Werte, die mehr als das 1,5-fache des Interquartilsabstands vom Median entfernt sind, sind als Ausreißer (Kreise), und bei Abweichungen von mehr als dem dreifachen des Interquartilsabstands als Extremfälle (Sterne) gekennzeichnet.



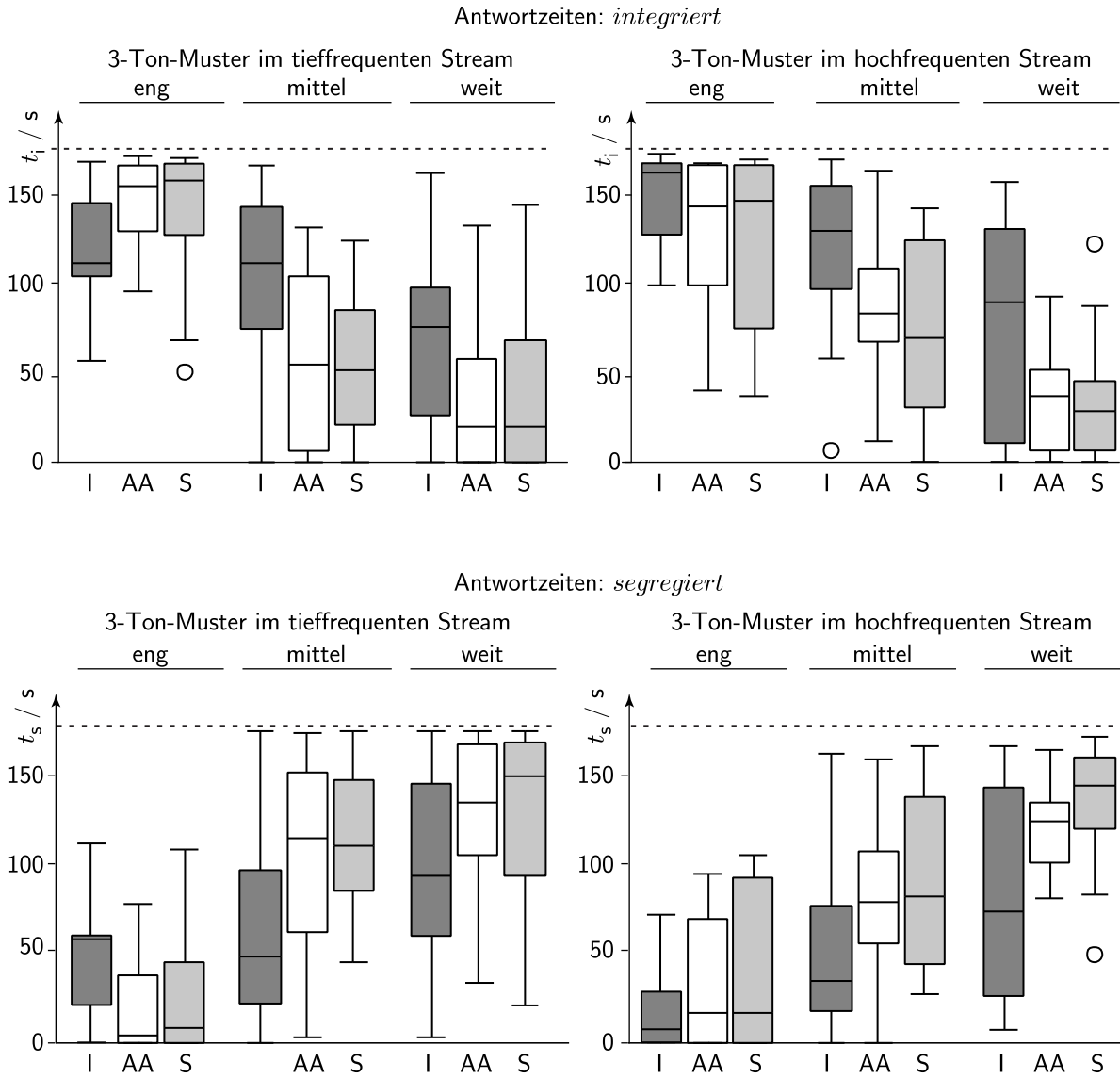


Abb. 5.7: Absolute Zeiten, in denen die Perzepte *integriert* t_i (oben) und *segregiert* t_s (unten) von den Probanden angezeigt worden sind, dargestellt als Boxplots für die Stimulationsbedingungen *integrierend* (I, dunkelgrau), *segregierend* (S, hellgrau) und *auditorisch allein* (AA, weiß) in den unterschiedlichen Frequenzabständen unter Berücksichtigung der Lage der Tripel (links: tieffrequenter Stream, rechts: hochfrequenter Stream). Die gepunktete Linie zeigt die maximal mögliche Zeit (Messzeit t_M) an.

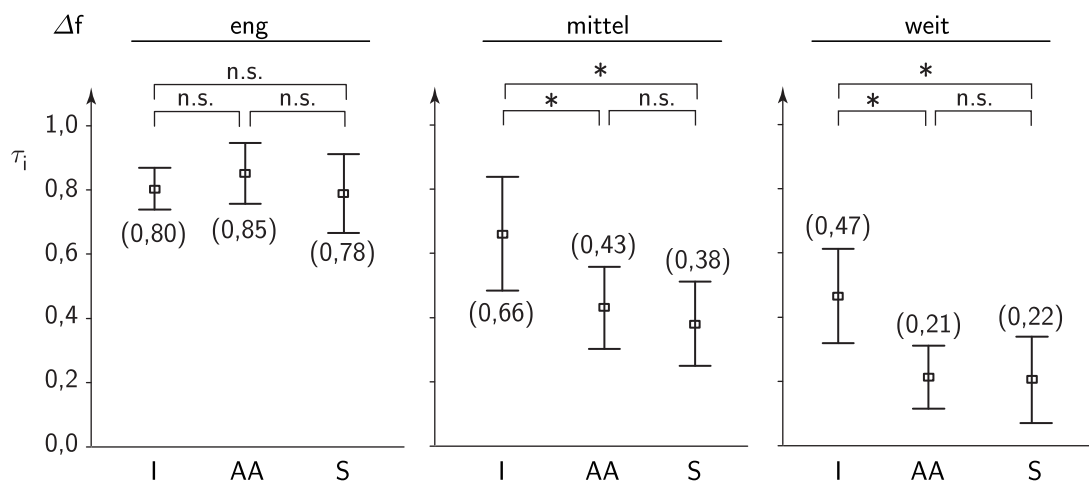


Abb. 5.8: Gruppenmittelwerte ($n = 14$) und Konfidenzintervall (95 %) der relativen Zeiten τ_i für alle Stimulationen, in denen die Stimuli *integriert* wahrgenommen wurden. Bei engem Frequenzabstand der Bereiche (links) wurde keine Abhängigkeit des auditorischen Perzepts von der visuellen Stimulation (I: *integrierend*, AA: *auditorisch allein*, S: *segregierend*) gefunden. Bei mittlerem (Mitte) und weitem Frequenzabstand (rechts) hingegen nimmt das Perzept eines integrierten Streams bei integrierender visueller Stimulation signifikant zu (*: $p < 0,05$).

Stimulation mit engem Frequenzabstand das Perzept unabhängig von der Art der visuellen Stimulation ist. Diese Unterschiede entsprechen der in der Varianzanalyse gefundenen Interaktion zwischen den Faktoren *Frequenzabstand* und *visueller Stimulationsbedingung*.

Mit zweiseitigen *t*-Tests, deren Signifikanzniveau $\alpha = 0,95$ durch Bonferroni-Korrektur für Mehrfachvergleiche angepasst wurde, wurden die Antwortzeiten τ_i innerhalb der Frequenzbereiche zwischen den visuellen Stimulationsbedingungen verglichen. Desweiteren wurde τ_i bei rein auditorischer Stimulation zwischen den Frequenzbereichen verglichen. Die Ergebnisse sind in Tab. 5.4 dargestellt. τ_i erwies sich hierbei zwischen den Frequenzabständen *eng* und *mittel* sowie zwischen *eng* und *weit* als signifikant verschieden. Zwischen den Bedingungen *mittel* und *weit* verfehlte τ_i auf einem Niveau von $\alpha = 0,90$ die Signifikanz ($p > 0,1$). Betrachtet man die Präsentationen mit engem Frequenzabstand (vgl. Abb. 5.8), so zeigten sich zwischen keinem Paar visueller Bedingungen signifikante Unterschiede in den Antwortzeiten τ_i . Bei mittlerem und weitem Frequenzabstand war die Bedingung *integrierend* verschieden von *segregierend* und *auditorisch allein*.

Wie erwartet, werden bei mittlerem Frequenzabstand die Stimuli jeweils etwa zur Hälfte segregiert oder integriert wahrgenommen. In den *per se* eindeutigen Bereichen (*weit* und *eng*) ist das auditorische Perzept ohne visuelle Stimulation erwartungsgemäß weitgehend Integration oder Segregation. Die visuelle Stimulation beeinflusst das auditorische Perzept im ambigen Bereich deutlich hin zur Wahrnehmung eines integrierten Streams.

5.3 MEG-Studie im ambigen Bereich

Die Gruppenmittelwerte der durch Standard- und Deviantstimuli evozierten Magnetfelder sind in Abb. 5.9 dargestellt. In beiden Bedingungen wurde durch diese Stimuli über temporalen und occipitalen Regionen des Kortex eine deutliche Aktivität messbar. In den Differenzkurven hingegen (Abb. 5.10) zeigte sich nur in der *segregierend*-Bedingung eine deutliche Aktivität, die als Unterschied zwischen Standard und Deviant zu werten ist. Dieser ist sowohl an den temporalen als auch occipitalen Sensorengruppen erkennbar.

Betrachtet man die Verläufe in der Überlagerung einer Auswahl aus den Sensoren (Abb. 5.11), so zeigen sich in der *segregierend*-Bedingung keine klar abgrenzbaren Komponenten evozierter Felder, was durch die hohe Stimulationsrate verursacht ist. In der *integrierend*-Bedingung sind in den Latenzbereichen um 0 ms und 330 ms an den temporalen Sensoren Wellenformen sichtbar. Diese können als auditorisch evozierte N1 der lautereren Töne zugeordnet werden. Die Latenz von 0 ms entspricht dabei der N1 des

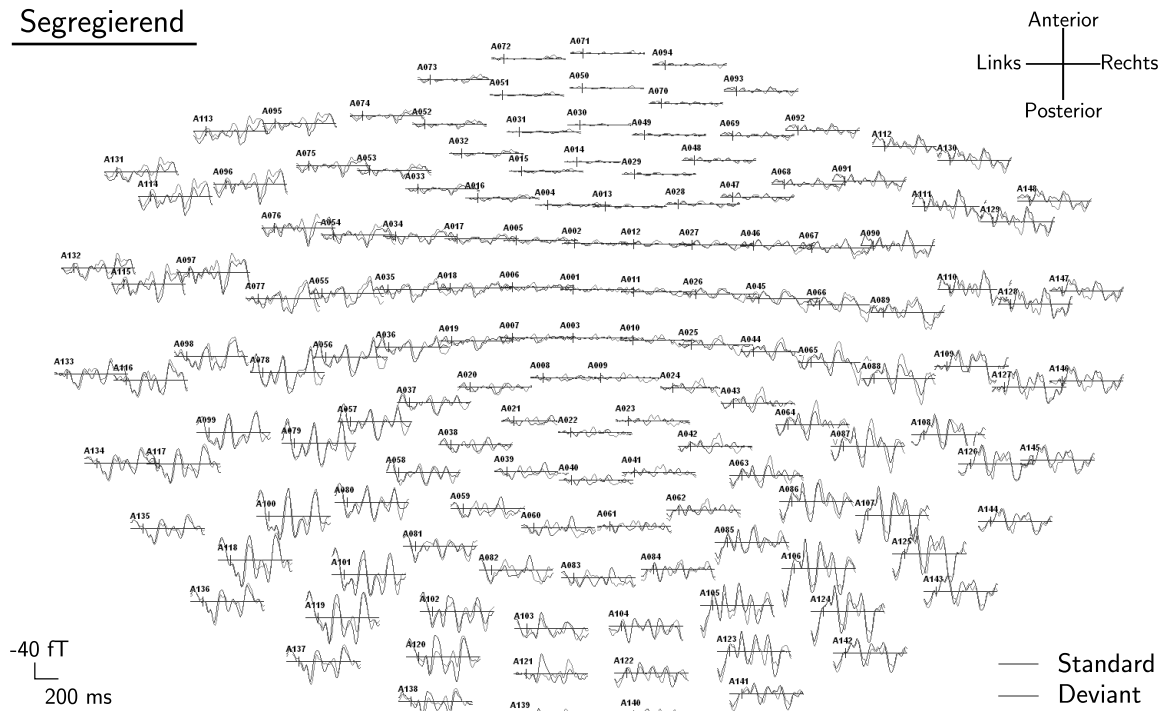
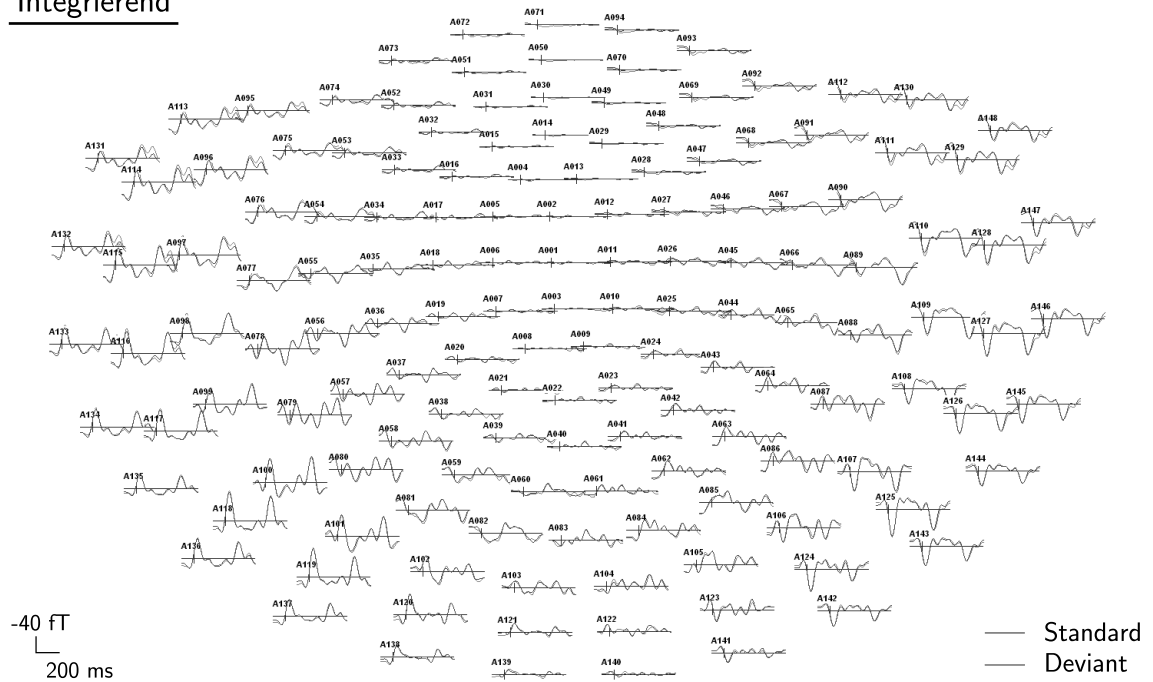
SegregierendIntegrierend

Abb. 5.9: Gruppenmittelwerte ($n = 13$) der durch Standard- und Deviantstimuli evozierten Magnetfelder in den Bedingungen *segregierend* (oben) und *integrierend* (unten).

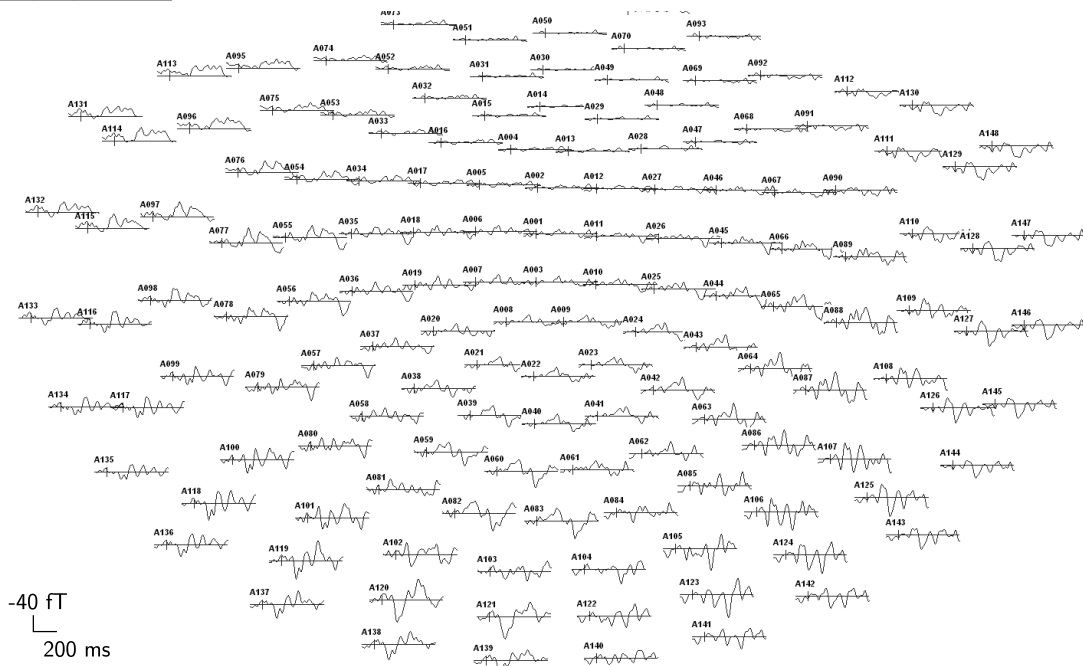
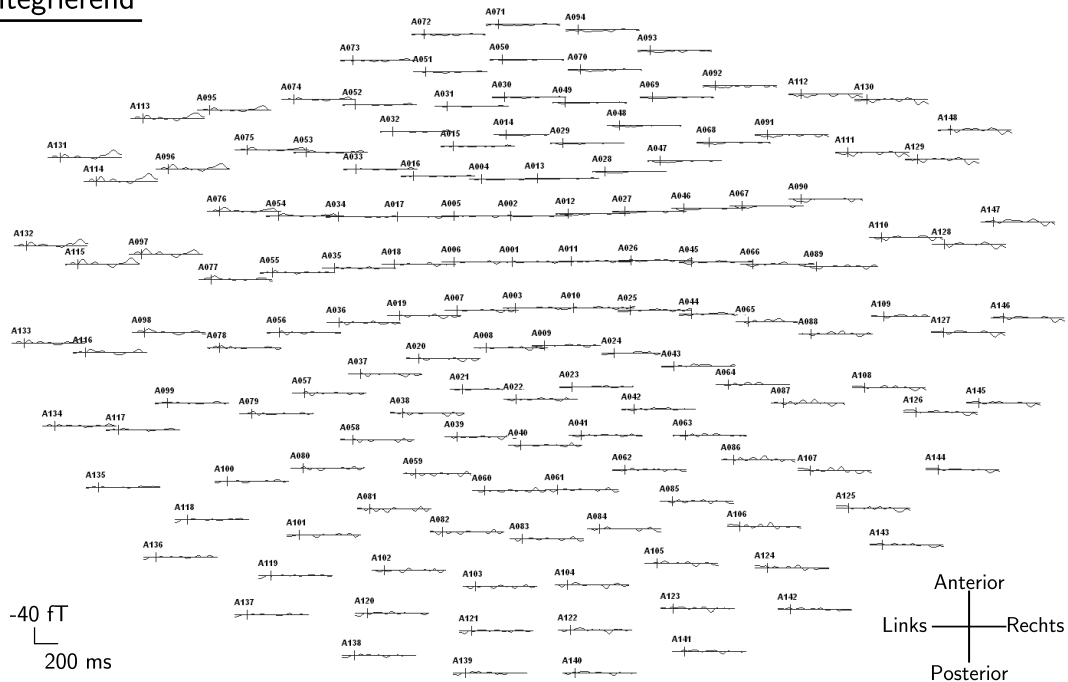
SegregierendIntegrierend

Abb. 5.10: Gruppenmittelwerte ($n = 13$) der Differenzkurven (Deviant- minus Standardantwort) in den Bedingungen *segregierend* (oben) und *integrierend* (unten).

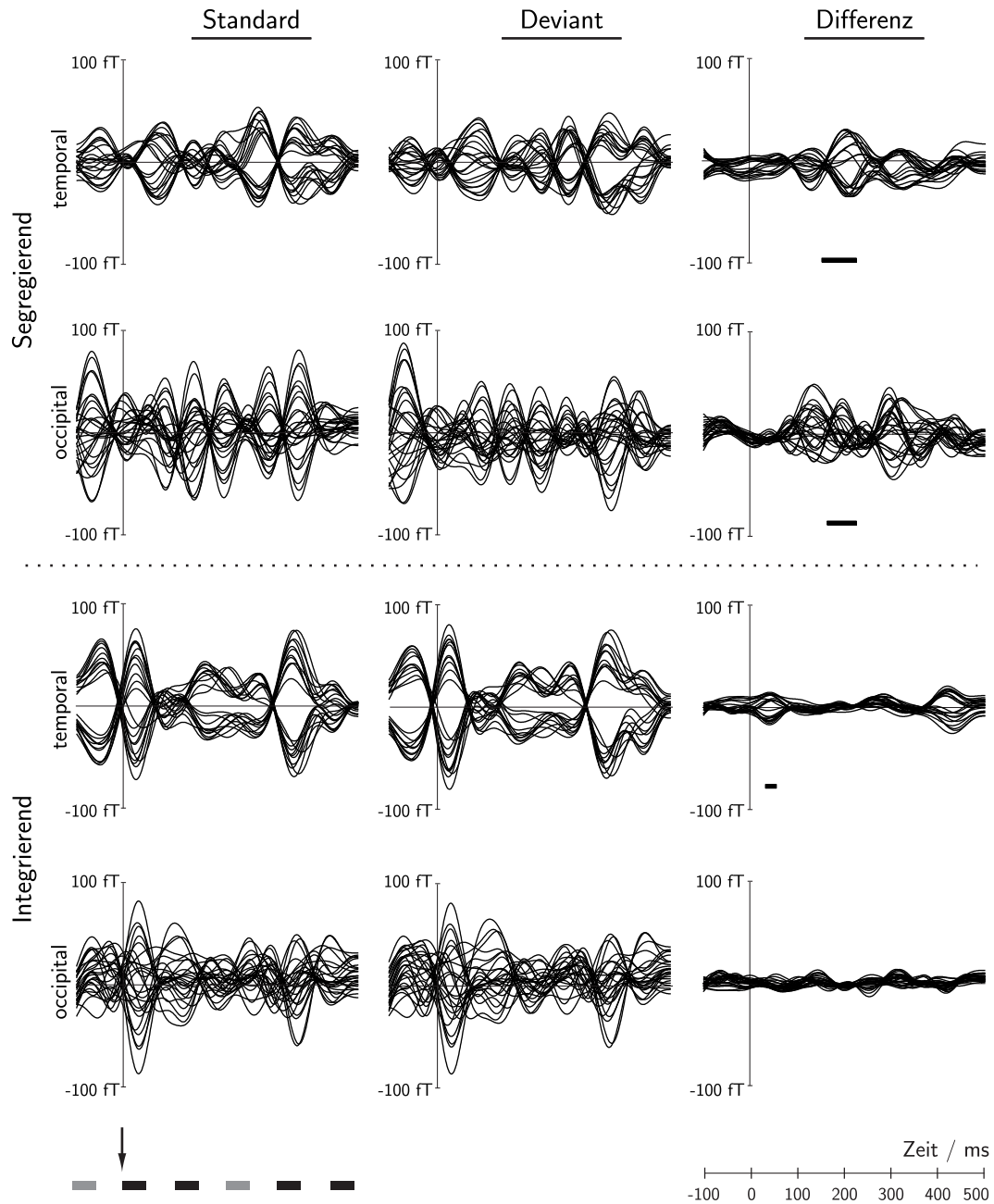


Abb. 5.11: Gruppenmittelwerte ($n = 13$) der im MEG gemessenen Magnetfelder, die durch die Standard- und Deviantstimuli evoziert wurden sowie deren Differenzkurven für eine modalitätsbezogene Auswahl aus den Sensoren in den Bedingung *segregierend* (oben) und *integrierend* (unten). Unten ist der zeitliche Bezug zum Stimulus schematisch dargestellt (grau: lautere Töne). Die Latenzintervalle mit significantem Unterschied ($p < 0,05$) zwischen den Standard- und Deviantantworten sind in den Differenzkurven markiert.

Tab. 5.4: Vergleiche der Antwortzeiten τ_1 (Gruppenmittel) zwischen den Stimulationsbedingungen gruppiert nach Vergleichen mit dem selben Frequenzabstand bzw. visueller Stimulation (t -Tests, Bonferroni-korrigiertes $\alpha = 0,95$).

Bedingung 1	—	Bedingung 2	$t(13)$	
mittel-integrierend	—	mittel-aud. allein	3,55	*
mittel-integrierend	—	mittel-segregierend	3,94	*
mittel-aud. allein	—	mittel-segregierend	0,68	
eng-integrierend	—	eng-aud. allein	-1,20	
eng-integrierend	—	eng-segregierend	0,33	
eng-aud. allein	—	eng-segregierend	1,98	
weit-integrierend	—	weit-aud. allein	3,81	*
weit-integrierend	—	weit-segregierend	4,13	*
weit-aud. allein	—	weit-segregierend	0,16	
eng-aud. allein	—	weit-aud. allein	7,68	**
mittel-aud. allein	—	eng-aud. allein	-4,96	**
mittel-aud. allein	—	weit-aud. allein	3,11	(*)

(*): $p < 0,1$; *: $p < 0,05$; **: $p < 0,01$

110 ms vor Beginn des ersten Tripeltons auftretenden lauterem Tones, die bei 330 ms der N1 des bei 220 ms auftretenden zweiten Tripeltones. Die dargestellten Differenzkurven verdeutlichen den Aktivitätsunterschied zwischen beiden Bedingungen und zeigen Latenzintervalle, in denen die von Standard- und Deviantstimuli evozierten Magnetfelder signifikant verschieden sind. Für die Sensoren über dem auditorischen Kortex ist dies in der *segregierend*-Bedingung das Intervall von 172 ms bis 247 ms sowie in der *integrierend*-Bedingung das Intervall von 39 ms bis 60 ms. Über dem visuellen Kortex wird nur in der *segregierend*-Bedingung ein Intervall mit signifikanten Unterschieden gefunden, und zwar im Latenzintervall von 173 ms bis 234 ms.

Zunächst liefert die ANOVA mit den Faktoren *Sensorgruppe*, *Stimulustyp* und *Bedingung* einen Effekt der *Sensorgruppe* ($F(50,650) = 6,47$, $p < 0,01$) und Interaktionen der Faktoren *Stimulustyp* und *Bedingung* ($F(50,650) = 6,11$, $p < 0,01$), *Bedingung* und *Sensorgruppe* ($F(50,650) = 8,34$, $p < 0,01$) sowie *Bedingung*, *Sensorgruppe* und *Stimulustyp* ($F(50,650) = 2,54$, $p < 0,05$). Für das in der *integrierend*-Bedingung gefundene Latenzintervall liefert die ANOVA mit den Faktoren *Sensor* und *Stimulustyp* für occipital gelegene Sensoren einen Haupteffekt des Sensors ($F(26,338) = 5,04$, $p < 0,01$), aber keine Haupteffekte und Interaktion beider Faktoren. Für die über dem auditorischen Kortex gelegenen

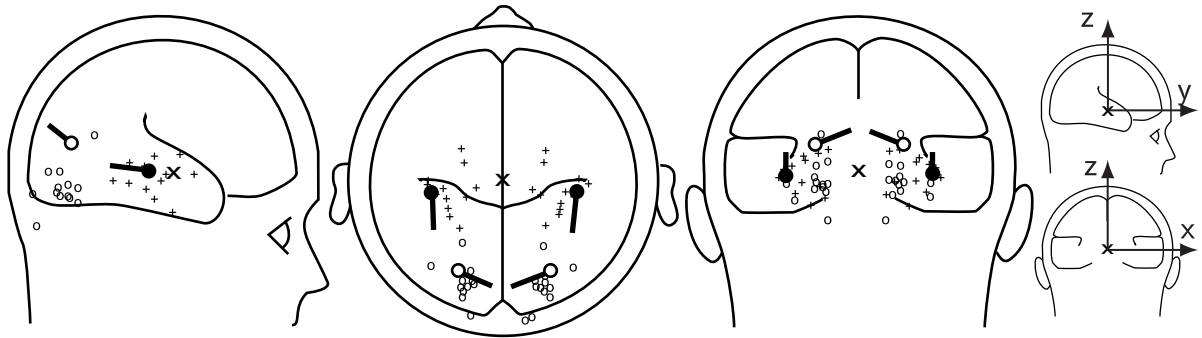


Abb. 5.12: Generatorenpaare (ECD), die das individuelle Differenzmagnetfeld der Bedingung *segregierend* erklären. Diese wurden sowohl temporal (+), als auch occipital (o) lokalisiert. Die ECD, die das über alle Probanden gemittelte Differenzfeld erklären, sind ebenfalls dargestellt (x: Ursprung des Koordinatensystems).

Sensoren ist der Haupteffekt des Faktors *Sensor* ebenfalls signifikant ($F(23,299) = 17,85$, $p < 0,001$), aber keine weiteren Haupteffekte und Interaktionen. Somit wurde kein MMF in der *integrierend*-Bedingung evoziert. In der *segregierend*-Bedingung liefert die ANOVA für die occipitale Sensorengruppe einen Haupteffekt des Faktors *Sensor* ($F(26,338) = 5,04$, $p < 0,01$) und einen Trend der Interaktion von *Sensor* und *Stimulustyp* ($F(26,338) = 2,25$, $p < 0,1$). Bei den temporal gemessenen Daten ist kein Haupteffekt, aber die Interaktion der Faktoren *Sensor* und *Stimulustyp* signifikant ($F(23,299) = 10,62$, $p < 0,05$). Die Interaktion beruht auf der Polaritätsumkehr zwischen rechten und linken Sensoren, womit in dieser Bedingung somit an beiden Sensorengruppen ein MMF evoziert wurde.

Für das in den Differenzverläufen (Gruppenmittel) der *segregierend*-Bedingung gefundene MMF wurde eine Quellenanalyse durchgeführt. Zunächst wurde ein symmetrisch zur Medianebene gelegenes Paar eines Stromdipols (ECD) angenommen. Hiermit wurde keine akzeptable Anpassungsgüte erzielt und die ermittelten Positionen lagen in unerklärlichen Regionen des Kortex, teilweise sogar außerhalb. Daher wurde ein zweites ECD-Paar postuliert, das für die visuelle und auditorische Modalität separate Generatoren zulässt. Hiermit lieferte das Modell ein temporal und ein occipital gelegenes Paar, die mit dem MMF übereinstimmen (Anpassungsgüte: 95 %). In Abb. 5.12 sind die Positionen und Orientierungen dieser ECD dargestellt. Im Gruppenmittel befinden sich diese vier Quellen an folgenden Positionen (x; y; z) im kartesischen Koordinatensystem (in mm): Temporal links ($-47; -8; -3$), Temporal rechts ($47; -8; -3$), Occipital links ($-30; -60; 17$) und Occipital rechts ($30; -60; 17$). Unter der Annahme von vier ECD wurde versucht, auch die MMF der individuellen Probandendaten zu modellieren. Hier zeigt sich zumindest bei

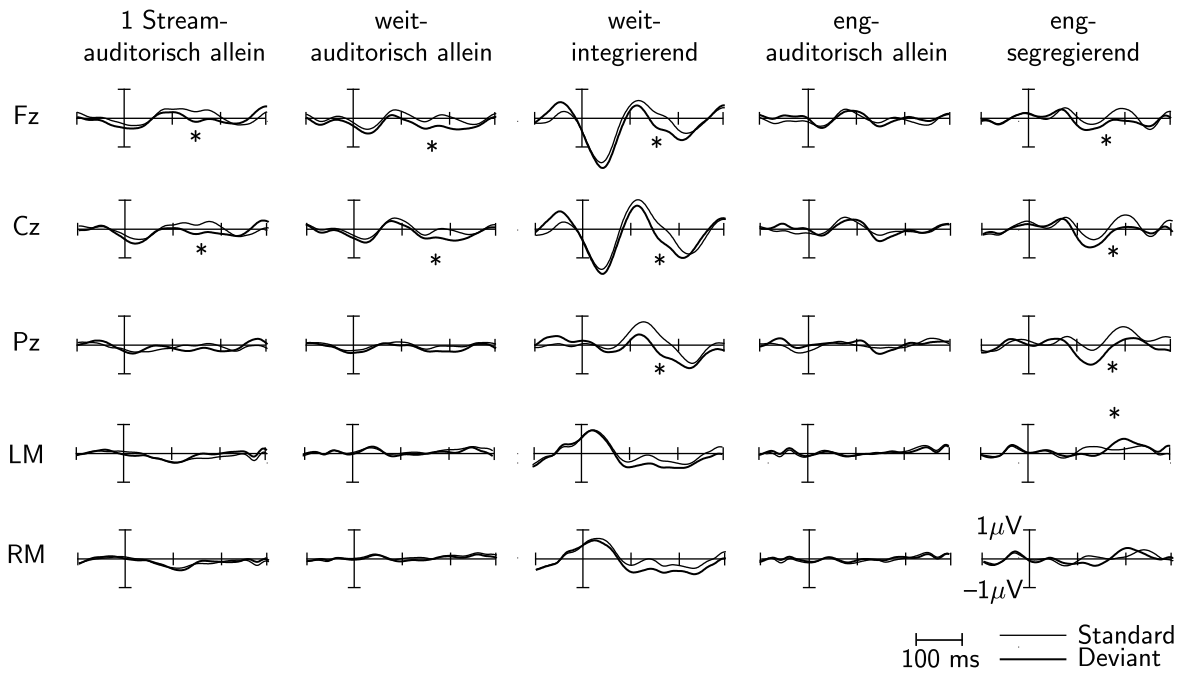


Abb. 5.13: Gruppenmittelwerte ($n = 15$) der EKP, die durch den ersten Stimulus eines Tripels evoziert wurden, gemessen an den Elektrodenpositionen Fz, Cz, Pz und den Mastoiden. Signifikante Unterschiede zwischen Standard- (dünn) und Deviantpotential (dick) sind markiert (Bonferroni-korrigiertes $\alpha = 0,95$ für Mehrfachvergleiche).

den meisten Probanden, dass sich die vier ECD in temporalen und occipitalen Bereichen des Kortex befinden. Somit muss das MMF auf verschiedenen Generatoren beruhen: Das temporal lokalisierte Paar kann einer auditorischen Komponente und das occipitale Paar einer visuellen Komponente zugeschrieben werden.

5.4 EEG-Studie im nicht-ambigen Bereich

Abb. 5.13 zeigt für eine Auswahl der EEG-Elektrodenpositionen die Gruppenmittelwerte für Standard- und Deviantpotentiale. Die Amplituden dieser Mittelwerte und die dazugehörigen Differenzen im für die MMN relevanten Latenzbereich sind in Tab. 5.5 dargestellt. Generell war, wie in der ersten EEG-Studie (vgl. Kap. 5.1), die Amplitude der EKP in Folge der hohen Stimulationsrate verhältnismäßig klein. Die von den Standard- und Deviantstimuli evozierten auditorischen aN1- und aP2-Wellen können trotzdem noch in allen Bedingungen mit ausschließlich auditorischer Stimulation identifiziert werden, wobei die Identifikation der aN1- und aP2-Wellen aller Stimuli in Folge der hohen Stimulationsrate

Tab. 5.5: Gruppenmittelwerte (MW) mit Standardabweichung (SD) der Amplituden (in μV) für Standard-, Deviant- und Differenzpotential.

		Standard		Deviant		Differenz	
		MW	SD	MW	SD	MW	SD
1 Stream- auditorisch allein	Fz	0,54	0,73	-0,13	0,80	-0,67	0,71
	Cz	0,38	0,78	-0,27	1,05	-0,65	0,97
	Pz	0,03	0,64	-0,25	0,94	-0,29	0,95
	LM	-0,42	0,42	-0,20	0,66	0,23	0,58
	RM	-0,33	0,37	-0,24	0,73	0,09	0,70
weit- auditorisch allein	Fz	0,16	0,37	-0,61	0,85	-0,78	0,86
	Cz	-0,02	0,52	-0,70	0,92	-0,68	0,92
	Pz	-0,04	0,55	-0,41	0,84	-0,37	0,85
	LM	0,15	0,36	0,04	0,64	-0,11	0,61
	RM	0,22	0,60	0,17	0,67	-0,06	0,48
weit- integrierend	Fz	-0,54	1,16	-1,24	1,15	-0,70	0,66
	Cz	-0,99	1,80	-1,69	1,45	-0,70	0,99
	Pz	-0,28	1,67	-1,11	1,12	-0,83	1,31
	LM	-0,69	0,81	-1,05	0,98	-0,36	0,72
	RM	-0,45	0,86	-0,84	1,15	-0,39	0,67
eng- auditorisch allein	Fz	0,00	0,55	-0,02	0,66	-0,02	0,53
	Cz	-0,15	0,53	-0,24	0,60	-0,09	0,61
	Pz	-0,13	0,48	-0,07	0,61	0,06	0,63
	LM	0,04	0,48	0,02	0,60	-0,03	0,42
	RM	0,01	0,61	0,13	0,60	0,12	0,44
eng- segregierend	Fz	0,63	0,98	-0,12	0,86	-0,75	1,01
	Cz	0,87	1,14	0,11	1,17	-0,76	1,23
	Pz	1,19	1,42	0,32	1,30	-0,87	1,50
	LM	0,23	0,64	0,83	0,84	0,60	0,96
	RM	0,18	0,72	0,65	0,84	0,47	0,94

nicht möglich ist (Javitt *et al.* 1993; Sussman *et al.* 1999). In den Bedingungen mit visueller Stimulation treten zusätzlich visuell evozierte Potentiale auf, die insbesondere in der Bedingung *weit-integrierend* deutlich ausgeprägt sind.

Die Drei-Wege-ANOVA mit den Faktoren *Elektrodenposition* (Fz, Cz, Pz, LM, RM), *Bedingung* (alle) und *Stimulustyp* (Standard, Deviant) ergab einen Effekt des *Stimulustyps* ($F(1,14) = 18,61$, $p < 0,01$) und der *Bedingung* ($F(4,56) = 4,21$, $p < 0,05$). In-

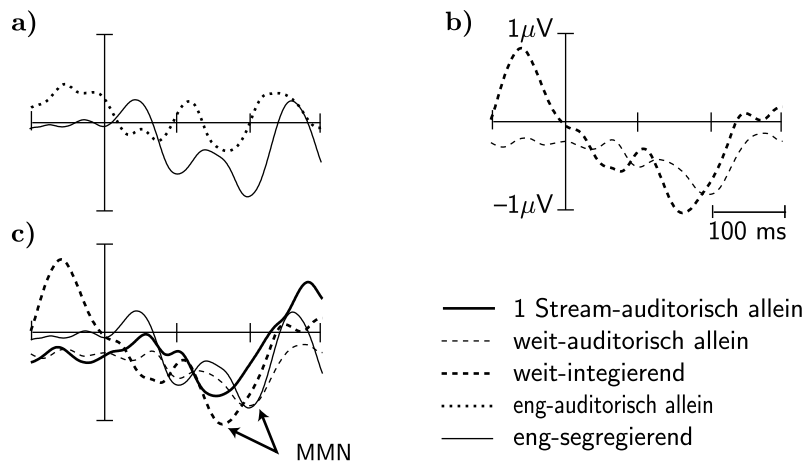


Abb. 5.14: Gruppenmittelwert ($n = 15$) der Differenzpotentiale an der Elektrode Fz in den Bedingungen mit engem Tonhöhenabstand (a) und mit weitem Tonhöhenabstand (b). Zusätzlich sind alle Differenzpotentiale, die eine MMN aufweisen, im Vergleich dargestellt (c).

teraktionen waren zwischen den Faktoren *Stimulustyp* und *Bedingung* ($F(4,56) = 3,89$, $p < 0,05$), *Stimulustyp* und *Elektrodenposition* ($F(4,56) = 20,18$, $p < 0,001$) sowie *Stimulustyp*, *Bedingung* und *Elektrodenposition* signifikant ($F(16,224) = 3,45$, $p < 0,01$). Es wurden die p -Werte nach Greenhouse-Geisser-Korrektur verwendet. Andere Haupteffekte und Interaktionen wurden nicht gefunden. *Post hoc* durchgeführte geplante Vergleiche (zweiseitige t -Tests für verbundene Stichproben) zwischen den Peak-Amplituden von Standard- und Deviantpotential ergaben nach Anwendung der Bonferroni-Korrektur signifikante Differenzen zwischen Standard- und Deviantpotential an Fz in den Bedingungen *1 Stream-auditorisch allein* ($t(14) = -3,70$), *weit-auditorisch allein* ($t(14) = -3,51$), *weit-integrierend* ($t(14) = -6,59$) und *eng-segregierend* ($t(14) = -2,89$). An Cz sind die Differenzen ebenfalls in diesen Bedingungen signifikant (*1 Stream-auditorisch allein*: $t(14) = -2,59$, *weit-auditorisch allein*: $t(14) = -2,84$, *weit-integrierend*: $t(14) = -4,57$, *eng-segregierend*: $t(14) = -2,39$), an Pz in den Bedingungen *weit-integrierend* ($t(14) = -3,97$) und *eng-segregierend* ($t(14) = -2,25$). Am linken Mastoid findet sich in der Bedingung *eng-segregierend* eine signifikant positive Differenz ($t(14) = 2,41$).

Abb. 5.14 zeigt die Differenzkurven zwischen Standard- und Deviantpotential. Eine Negativierung (Positivierung an den Mastoiden) im Latenzbereich um 150–200 ms wurde in den Differenzpotentialen als MMN identifiziert. Die MMN ist in den Bedingungen *1 Stream-auditorisch allein*, *weit-auditorisch allein* und *eng-segregierend* deutlich ausgeprägt. Die

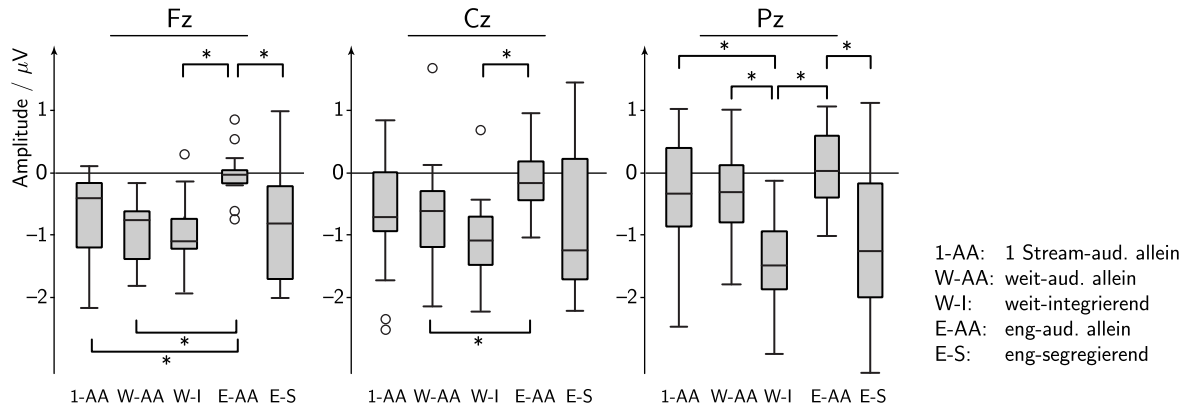


Abb. 5.15: Boxplots der MMN-Amplituden an den Elektrodenpositionen Fz (links), Cz (Mitte) und Pz (rechts) für alle Stimulationsbedingungen. Signifikante Unterschiede zwischen den Amplituden sind markiert (Bonferroni-korrigiertes $\alpha = 0,95$ für Mehrfachvergleiche).

Inversion der Polarität an den Mastoiden ist in den Bedingungen *1 Stream-auditorisch allein* und *eng-segregierend* erkennbar.

Die individuellen MMN-Amplituden sind in Abb. 5.15 zusammengefasst. *Post-hoc*-Tests (Bonferroni-korrigiertes $\alpha = 0,95$) der Interaktion von *Stimulustyp*, *Bedingung* und *Elektrodenposition* ergaben unterschiedliche Differenzpotentiale zwischen den Bedingungen *1 Stream-auditorisch allein* und *eng-auditorisch allein* (Fz: $t(14) = -2,78$), *eng-segregierend* und *eng-auditorisch allein* (Fz: $t(14) = -2,33$), *eng-auditorisch allein* und *weit-integrierend* (Fz: $t(14) = -5,70$, Cz: $t(14) = -4,15$) sowie zwischen *eng-auditorisch allein* und *weit-auditorisch allein* (Fz: $t(14) = -2,83$, Cz: $t(14) = -2,38$).

In Abb. 5.16 sind die Spannungsverteilungen und SCD-Karten der Schädeloberfläche für jede Stimulationsbedingung dargestellt. Hier zeigt sich zum einen, dass in der Kontrollbedingung (*1 Stream-auditorisch allein*) die für die MMN typischen Foci der Negativität an Cz und Fz auftreten. Diese typische Verteilung tritt zudem in der Bedingung *weit-auditorisch allein* auf. In der Bedingung *eng-segregierend* tritt zusätzlich zu den frontozentralen Foci ein occipital lokalisierter Focus auf. Keine klare Verteilung der Foci ist in den Bedingungen *eng-auditorisch allein* und *weit-integrierend* erkennbar.

Das Auftreten der MMN wurde durch eine Analyse der neuronalen Quellen verifiziert. Wie in Abb. 5.17 dargestellt, ergibt sich in der Kontrollbedingung *1 Stream-auditorisch allein* ein bilaterales temporal gelegenes Generatorenpaar, das die Amplituden des Differenzpotentials erklärt. Der Ort dieser Generatoren wurde auch in den Bedingungen *weit-*

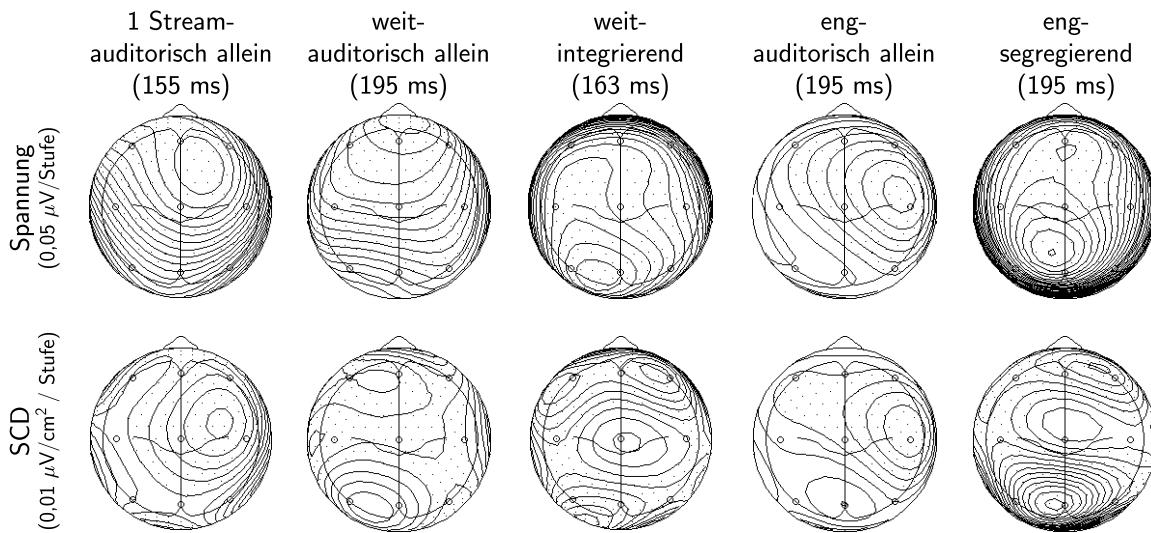


Abb. 5.16: Spannungsverteilung (Gruppenmittelwert) auf der Schädeloberfläche (oben) und SCD-Karten (unten) für alle Stimulationsbedingungen bei den Latenzen mit größter negativer Amplitude an Fz im Differenzpotential. Gepunktete Flächen kennzeichnen negative Werte. Alle Bedingungen außer *eng-auditorisch allein* weisen frontale Quellen und zusätzlich posterior gelegene Quellen auf. In der Bedingung *eng-segregierend* ist zusätzlich eine starke posterior gelegene Quelle erkennbar.

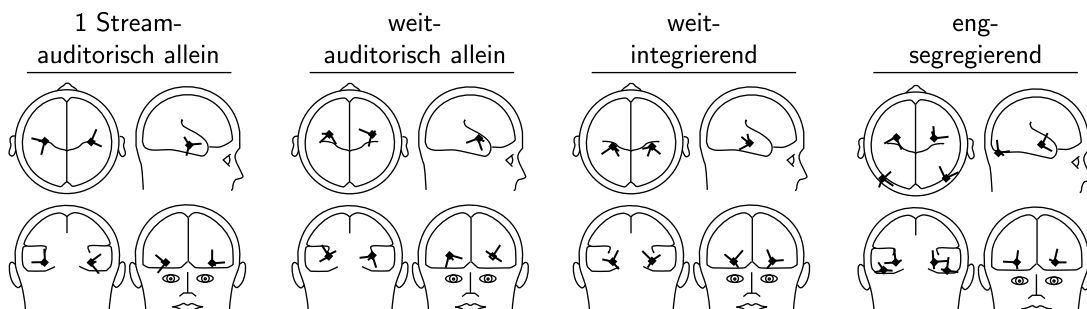


Abb. 5.17: Lage und Orientierung der Generatorenpaare, die das Differenzpotential der Bedingungen *1 Stream-auditorisch allein*, *weit-auditorisch allein*, *weit-integrierend* und *eng-segregierend* ergeben. Die temporal gelegenen Generatoren werden in der Bedingung *eng-segregierend* um ein occipital gelegenes Paar ergänzt.

auditorisch allein und *eng-segregierend* bestimmt. In der Bedingung *eng-segregierend* kann das Differenzpotential durch Annahme eines weiteren, occipital gelegenen Generatorenpaares besser erklärt werden. Die Restvarianz beträgt dann 9 %.

5.5 Kontrolle der Aufmerksamkeit

In dem MEG- und beiden EEG-Experimenten wurde die Aufmerksamkeit der Probanden auf die visuellen Stimuli durch Targetreize kontrolliert. Die Detektionsrate war in allen Experimenten mit einem Sensitivitätsindex von $d' > 3$ sehr hoch. Somit wurden die Targetreize in allen Bedingungen sicher und ohne Mühe detektiert. Die Werte des Sensitivitätsindex sind in Abb. 5.18 dargestellt.

Die Auswertung der Antwortzeiten für die visuellen Targetreize ergab in den EEG-Experimenten etwa gleiche Werte (vgl. Abb. 5.19), die insgesamt gegenüber denen des MEG-Experiments verringert sind. Eine ANOVA des Sensitivitätsindex mit dem Faktor *Experiment* ergab einen signifikanten Effekt ($F(2,3339) = 811, p < 0,001$). *Post hoc* durchgeführte Vergleiche (Tukey, vgl. Kap. 4.2.6) ergaben keinen Unterschied zwischen den Mittelwerten der EEG-Experimente, jedoch gegenüber beiden EEG-Experimenten signifikant größere Antwortzeiten im MEG-Experiment.

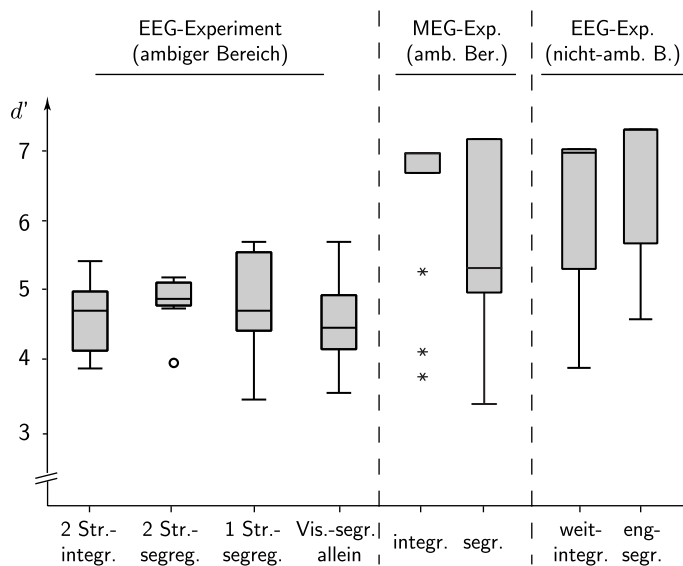


Abb. 5.18: Boxplot der Sensitivität d' für die Bedingungen mit visueller Stimulation im MEG- und beiden EEG-Experimenten. Dargestellt sind Median und Interquartilsabstand. Werte, die mehr als das 1,5-fache des Interquartilsabstands vom Median entfernt sind, sind als Ausreißer (Kreise), und bei Abweichungen von mehr als dem dreifachen des Interquartilsabstands als Extremfälle (Sterne) gekennzeichnet.

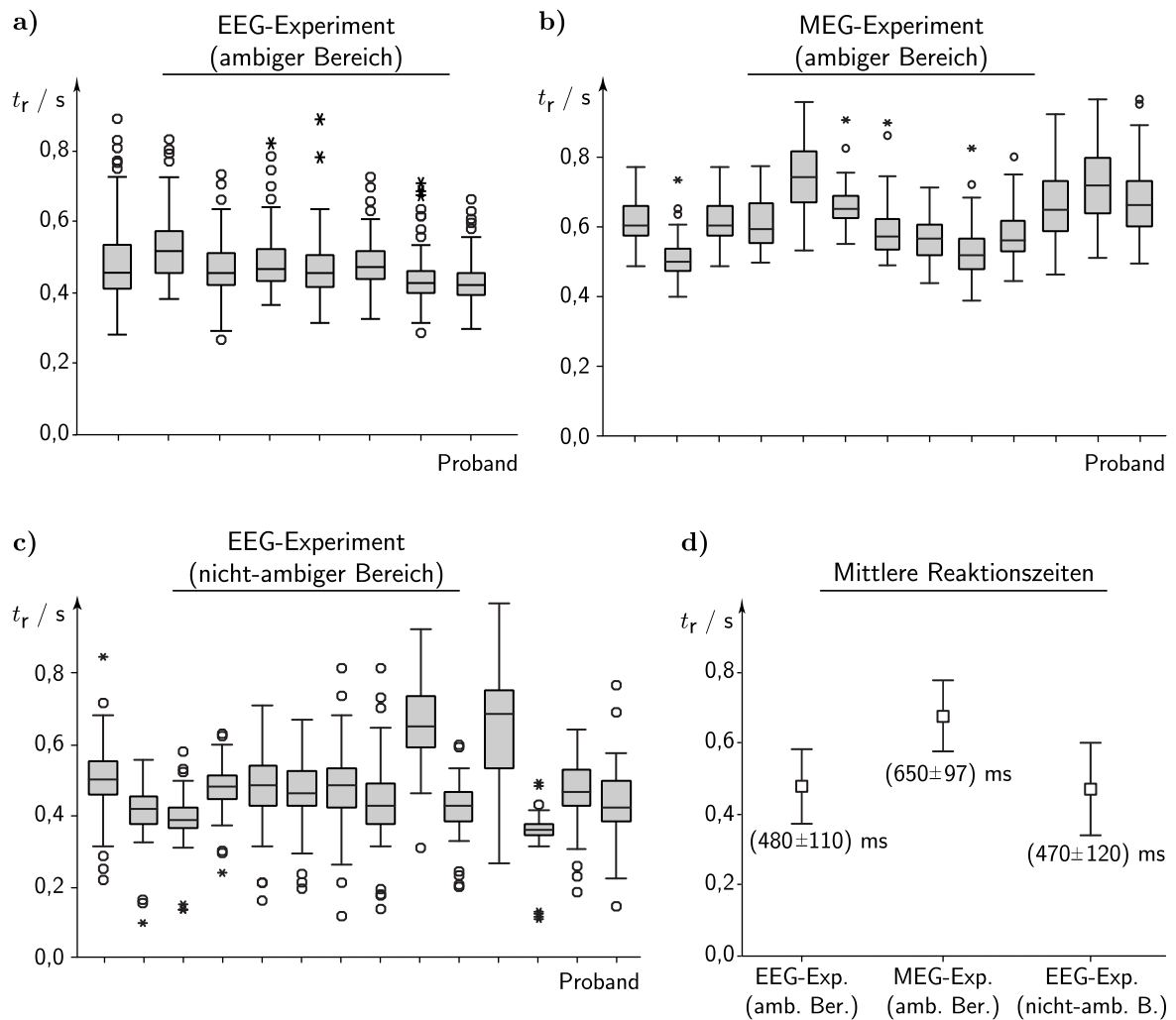


Abb. 5.19: **a-c)** Boxplot der Reaktionszeiten t_r für alle Bedingungen mit visueller Stimulation. Dargestellt sind Median und Interquartilsabstand. Werte, die mehr als das 1,5-fache des Interquartilsabstands vom Median entfernt sind, sind als Ausreißer (Kreise), und bei Abweichungen von mehr als dem dreifachen des Interquartilsabstands als Extremfälle (Sterne) gekennzeichnet. **d)** Mittlere Reaktionszeiten mit Standardabweichung für alle Experimente mit visueller Stimulation.

Kapitel 6

Diskussion der Ergebnisse

Die Ergebnisse dieser Arbeit zeigen einen deutlichen Einfluss der visuellen Stimulation auf die kortikale Organisation auditorischer Stimuli. Dieser modalitätsübergreifende Effekt konnte in allen durchgeführten Experimenten gefunden werden. Die auditorische Organisation bzw. das auditorische Perzept wurde dabei nicht nur in *per se* ambigen Stimuluskonfigurationen durch synchrone visuelle Stimulation deutlich modifiziert, sondern auch dann, wenn durch den Frequenzabstand der Bereiche und die Stimulationsrate die auditorischen Stimuli immer integriert oder segregiert wahrgenommen werden sollten. Zunächst werden die Ergebnisse im Hinblick auf die unterschiedlichen Stimuluskonfigurationen betrachtet und im Weiteren die audiovisuellen Interaktionen diskutiert.

6.1 Audiovisuelle Stimulation

6.1.1 Auflösung der Ambiguität

In der EEG- und MEG Studie mit ambiger Stimuluskonfiguration wurden sowohl auditorische als auch visuelle Stimulusfolgen verwendet, die jeweils Deviantreize enthielten. Demzufolge war eine Überlagerung auditorisch und visuell evozierter Potentiale, wie sie die Ergebnisse zeigen, zu erwarten. Die Amplitude der auditorisch evozierten Potentiale war im Vergleich zu anderen Studien gering. Dies liegt in der vorliegenden Arbeit an der hohen Stimulationsrate, die auch die Zuordnung der EKP zu den einzelnen Stimuli nicht immer möglich macht (Javitt *et al.* 1993; Sussman *et al.* 1999). Zudem führte die hohe

Stimulationsrate in den hier durchgeführten Experimenten zu einer Habituation der auditorisch evozierten Komponenten N1 und P2. Die MMN jedoch, die in der Differenzantwort (Deviant minus Standard) auftritt, ist von dieser hohen Stimulationsrate durch den großen zeitlichen Abstand der Deviantreize weniger beeinflusst (Sussman *et al.* 1999). Allerdings tritt sie in beiden Modalitäten auf und muss deshalb in den Ergebnissen diesen wieder eindeutig zugeordnet werden.

Das EEG-Experiment mit den zu einer ambigen Stream-Formation führenden auditorischen Stimuli sollte zeigen, dass die Mehrdeutigkeit durch visuelle Stimulation aufgelöst werden kann. In der *2 Streams-segregierend*-Bedingung wurde ein stream-internes Tripel-Muster durch die visuellen Stimuli nachempfunden, in der *2 Streams-integrierend*-Bedingung ein stream-übergreifendes Intensitätsmuster. Dadurch sollte eine Veränderung der Objektbildung in Richtung der Formation eines oder zweier auditorischer Streams bewirkt werden. Der experimentelle Nachweis wurde über die MMN geführt, die, wie die Ergebnisse zeigen, in der *2 Streams-segregierend*-Bedingung, aber nicht in der *2 Streams-integrierend*-Bedingung evoziert wurde. Dieser Unterschied ist bemerkenswert, denn die auditorischen Stimuli beider Bedingungen waren identisch. Die Tripel in der Deviantfunktion evozierten nur dann eine MMN, wenn die visuelle Präsentation mit dem stream-internen Muster synchronisiert war und hierdurch die tieffrequenten Stimuli in einen Stream segregiert wurden. Offenbar wurde erst durch die visuelle Stimulation die Regularität im auditorischen Stream und somit auch die Abweichung davon registriert, die sich in der Evozierung einer MMN manifestiert. Dahingegen wurde in der *2 Streams-integrierend*-Bedingung die Regularität durch die integrierte Wahrnehmung aller Töne nicht mehr erkannt und somit auch keine MMN evoziert. Die Modifikation der MMN-Amplitude bei gleicher auditorischer Stimulation lässt auf eine veränderte Repräsentation der auditorischen Reizfolgen im Kortex (Sussman *et al.* 1999) und, da die visuelle Stimulation variierte, auf einen deutlichen Effekt der visuellen Stimulation auf die auditorische Objektbildung schließen.

Die Messwerte der Ausgangsbedingung (ohne visuelle Stimulation) unterstützen die gefundenen Ergebnisse. Insbesondere zeigen die Ergebnisse bei der rein auditorischen Präsentation (*Auditorisch-allein*-Bedingung) keine signifikante MMN. Dies ist u. a. durch einen mehrfachen Wechsel der auditorischen Organisation zwischen Integration und Segregation in dieser ambigen Stimuluskonfiguration erklärbar. Bei optimaler Wahl der auditorischen Stimuli lägen bei der Hälfte der gemessenen Deviantantworten segregierte Streams vor. Da aber zudem erst einige Standardstimuli präsentiert werden müssen, bis Deviantstimuli eine MMN evozieren, beträgt die erwartete MMN-Amplitude deutlich unter 50 % der MMN-

Amplitude in der Kontrollbedingung. Damit ist sie bei den vorliegenden Messwerten im Bereich der Reststörung (Picton *et al.* 2000) und kann von dieser nicht abgegrenzt werden. Die Kontrollbedingungen *1 Stream* und *Visuell-allein* zeigen eine auditorisch bzw. visuell evozierte MMN, die vergleichbar mit der MMN in der *2 Streams-segregierend*-Bedingung ist. Folglich unterstützten sie die Identifizierung der Negativität als MMN.

Die MMN liefert Informationen über die automatische neuronale Organisation der Stimuli. Sie ist zwar die Basis für die Bildung des bewussten Perzepts, jedoch sind weitere, aufmerksamkeitsgesteuerte *top-down*-Prozesse denkbar, durch welche die automatische, basale Verarbeitung und das Werden des Perzepts beeinflusst wird. Um einen experimentellen Zugang zum bewussten Perzept zu erhalten, wurde das psychoakustische Experiment durchgeführt, bei dem die Versuchspersonen bei Präsentation der selben Töne, aber unterschiedlichen visuellen Stimuli einschätzen sollten, ob sie einen oder zwei Streams hören. Bei rein auditorischer Stimulation lagen die Perzepte *Integration* und *Segregation* im zeitlichen Verlauf etwa gleich häufig vor. Dies zeigt, dass die Wahl der auditorischen Parameter *Frequenzabstand* und *SOA* zu einem ambigen auditorischen Perzept geführt hat. Hierzu sei bemerkt, dass im Gegensatz zu klassischen *ABAB*-Paradigmen in beiden Frequenzbereichen jeweils drei unterschiedliche Frequenzen verwendet wurden. Auch bei diesem erweiterten *ABAB*-Paradigma liefert offenbar, wenn der Frequenzabstand der mittleren Töne als Vergleich verwendet wird, das Van-Noorden-Diagramm für *ABAB*-Paradigmen (nach Bregman 1999, S. 60) einen guten Anhaltspunkt für die zu erwartende auditorische Organisation. In diesem Experiment weisen die gemessenen Perzepte eine große interindividuelle Variabilität auf, die jedoch auf die interindividuelle Unschärfe von *fission boundary* (FB) und *temporal coherence boundary* (TCB) zurückgeführt werden kann (van Noorden 1975).

Das psychoakustische Experiment zeigt einen deutlichen Einfluss der simultanen visuellen Stimulation auf das auditorische Perzept der ambigen Stimuluskonfiguration. Bei mittlerem Frequenzabstand der Sequenzen (vgl. Abb. 5.8) wurde durch die Synchronisierung der visuellen Stimuli mit dem stream-übergreifenden Intensitätsmuster eine signifikant verstärkte Wahrnehmung der Töne als ein einziger Stream beobachtet. Demzufolge bewirkt diese stream-übergreifende visuelle Information eine perzeptuelle Integration der präsentierten Töne. Der generelle Einfluss von visuellen Stimuli auf die auditorische Wahrnehmung ist bereits von der Beurteilung der Richtung von Schallquellen (Bertelson *et al.* 2000) sowie der Kategorisierung von Konsonanten bekannt (McGurk und McDonald 1976). O'Leary und Rhodes (1984) zeigten, dass die Anzahl der visuell wahrgenommenen Objekte die Anzahl der auditorisch wahrgenommenen Streams beeinflussen kann. Die Er-

gebnisse der psychoakustischen Studie bestätigen den Einfluss der visuellen Stimulation auf die Wahrnehmung einer ambigen Stimuluskonfiguration.

Es ist allerdings auch festzustellen, dass die segregierenden visuellen Stimuli das Perzept segregierter Streams nicht signifikant verstärkten. Eine mögliche Erklärung hierfür ist die relativ komplizierte Aufgabe, die die Probanden während der Messung zu erfüllen hatten: Sie sollten gleichzeitig visuelle Targetreize finden und die auditorische Wahrnehmung beurteilen. Der Vergleich der visuellen Bedingungen *segregierend* und *integrierend* zeigt eine schnellere Präsentationsrate der Quadrate (SOA = 220 ms) gegenüber der der Kreise (SOA = 330 ms). Durch die schnelle Abfolge der visuellen Stimuli in der *segregierend*-Bedingung könnten die Probanden überfordert gewesen sein, was die Zuordnung von visuellen und auditorischen Stimuli eingeschränkt haben könnte (vgl. Kap. 6.2.2). Auch erscheint es möglich, dass neben der Frequenz die Pegelunterschiede zwischen den auditorischen Stimuli selbst segregierend gewirkt haben. Die Stimuli mit erhöhtem Pegel wären dann eine Folge mit Tönen aus alternierenden Frequenzbereichen. Dieser würde somit bei den Probanden das Perzept eines integrierten Streams hervorrufen. Da die Segregation auf der Basis des Pegels zu einer erhöhten relativen Integrationszeit führen müsste, aber in allen Bedingungen nicht beobachtet wurde, erscheint diese Erklärung als nicht plausibel. Auch wurde nicht explizit nach den auditorischen Perzepten *1 Stream* oder *2 Streams*, sondern nach den Perzepten *Drei-Ton-Sägezahnmuster* oder *Zwei-Ton-Folge (hoch-tief)* gefragt. Diese indirekte Bewertung des auditorischen Streaming-Perzepts schafft, weil der Proband in Bezug auf das 2-Stream-Konzept naiv bleibt, eine objektivere Methode zur Psychometrie von audiovisuellen Stimulationsparadigmen als das direkte Abfragen von Streaming-Perzepten. Die Pegelunterschiede waren somit geeignet gewählt und wirkten nicht vordergründig segregierend.

Im psychoakustischen Experiment waren auch zufällig eingestreute Bedingungen enthalten, deren visuelle Reizfolgen dem erwarteten eindeutigen auditorischen Perzept widersprachen. Dadurch sollte eine Beurteilung der audiovisuellen Szene allein durch Auswertung der visuellen Stimuli ausgeschlossen werden. Die Auswertung zeigt, dass bei gleicher visueller Stimulation in Abhängigkeit vom Frequenzabstand der Streams in den eindeutigen Bereichen die erwarteten auditorischen Wahrnehmungen angegeben wurden. Somit wurde tatsächlich die auditorische Wahrnehmung von den Probanden beurteilt.

Die relativen Initialzeiten bis zur ersten Antwort des Probanden betragen etwas über fünf Sekunden. Dies entspricht den üblichen beobachteten Zeiten bis zur Etablierung eines eindeutigen Perzepts (Anstis und Saida 1985; Pressnitzer und Hupé 2006). Geht man davon aus, dass zunächst immer das 1-Stream-Perzept vorliegt, müsste dieses in der Bedingung

weit-auditorisch allein vermehrt als erste Antwort auftreten. Beobachtet wurde hingegen, dass in beiden Bedingungen (*eng-auditorisch allein* und *weit-auditorisch allein*) etwa gleiche Anteile von dem erwarteten automatischen Perzept entgegengesetzten Erstantworten auftraten. Auch waren die Zeiten bis zur Entscheidung für das erste jeweilige Perzept etwa gleich. Möglicherweise verursachte das Wissen der Probanden um die möglichen Perzpte eine abwartende Haltung, bis eine klare Entscheidung möglich wurde. Dann wurde mit der ersten Entscheidung nicht das initiale Perzept gemessen und stellt somit auch keinen Widerspruch zu den o. g. Studien dar.

Auch im MEG-Experiment wurde in einer ambigen auditorischen Stimuluskonfiguration durch die visuelle Stimulation die auditorische Organisation verändert, was konsistent mit den Ergebnissen der beiden vorgenannten Studien ist. Das hierbei untersuchte Mismatch-Field (MMF) als Komponente der evozierten Magnetfelder ermöglicht, analog zur MMN im EEG-Experiment, den Zugang zur automatischen Detektion der Unregelmäßigkeiten und damit zur Stream-Formation (Yabe *et al.* 2001; Gutschalk *et al.* 2005). Es wurde nur in der *segregierend*-Bedingung evoziert, wenngleich seine Latenz größer als in der EEG-Messung ist. Zur Abgrenzung der auditorisch und visuell evozierten Differenzantwort voneinander und deren Interpretation sei auf die Ausführungen zu den neuronalen Quellen verwiesen (vgl. Kap. 6.2.3).

Im Gegensatz zum EEG-Experiment kann in den MEG-Daten durch die bessere räumliche Auflösung auch die N1-Amplitude ausgewertet werden. Cowan (1995) sowie May und Tiitinen (2004) beschreiben, dass durch die biomagnetische N1 oder das Gleichspannungspotential diejenigen Formen des auditorischen sensorischen Gedächtnisses repräsentiert werden, die mit der Wahrnehmung von Tönen als singuläre Elemente oder Streams korreliert sind. Auch in einem EEG-Experiment wurde eine Korrelation der P1-N1-P2-Amplitude mit dem psychoakustisch gemessenen Streaming-Perzept nachgewiesen (Snyder *et al.* 2006). Dabei ist die Amplitude des Komplexes kleiner, weil durch die integrierte Wahrnehmung der Töne das SOA kürzer ist als das SOA eines segregierten Streams. In der durchgeführten MEG-Studie ist in der *segregierend*-Bedingung die N1-Amplitude klein, was auf das kurze SOA der auditorischen Stimulation zurückgeführt werden kann. Wenn die Wahrnehmung der lauterer Stimuli als ein integrierter Stream erfolgen würde, hätte das dabei wahrgenommene SOA eine Länge von 330 ms und würde somit zu einer Vergrößerung der P1-N1-Amplitude führen. Tatsächlich ist in der *integrierend*-Bedingung an den temporalen Sensoren eine gegenüber der *segregierend*-Bedingung veränderte Wellenform erkennbar (vgl. Abb. 5.11). Ihre Latenz kann der P1 oder auch der N1 des letzten lauterer Stimulus vor Beginn der Tripel zugeordnet werden. Da diese nur in der *integrie-*

rend-Bedingung auftritt, spricht der Befund für eine integrierte Wahrnehmung aller Töne als ein Stream, wobei die N1 durch die Pegelerhöhung jedes dritten Tones hervorgerufen wird. Generell kann im vorliegenden Fall die zwischen den Bedingungen veränderte N1-Amplitude als Veränderung der auditorischen Organisation interpretiert werden.

6.1.2 Nicht-ambiger Bereich

In dieser Arbeit wurden in zwei Experimenten Stimuluskonfigurationen verwendet, die einen so großen (bzw. kleinen) Abstand zwischen den akustischen Streams hatten, so dass diese automatisch als zwei segregierte Streams (bzw. als ein integrierter Stream) wahrgenommen werden sollten (vgl. Abb. 6.1). Wie bei ambiger Stimuluskonfiguration wurde auch in diesen Bereichen versucht, durch synchrone visuelle Präsentation eine Veränderung der auditorischen neuronalen Organisation herbeizuführen.

Zunächst wurde in einem psychoakustischen Experiment das auditorische Perzept von Probanden beurteilt, das durch gezielte Aufmerksamkeit auf eine Kombination auditorischer und visueller Reize entstanden ist. In den Bedingungen ohne visuelle Stimulation (enger bzw. weiter Frequenzabstand) ist das auditorische Perzept in Abhängigkeit vom Frequenzabstand der akustischen Streams vorrangig Integration oder Segregation. Dies stimmt mit Studien überein, die das auditorische Perzept in Abhängigkeit vom Frequenzabstand untersuchen (Bregman 1999; Rose und Moore 1997; Carlyon 2004). In der Stimuluskonfiguration mit weitem Frequenzabstand, der ohne visuelle Stimulation ein segregiertes Perzept erwarten ließe, wurde jedoch auch ohne visuelle Stimulation von den Probanden zeitweise eine integrierte Wahrnehmung angezeigt. Auch bei Verwendung der

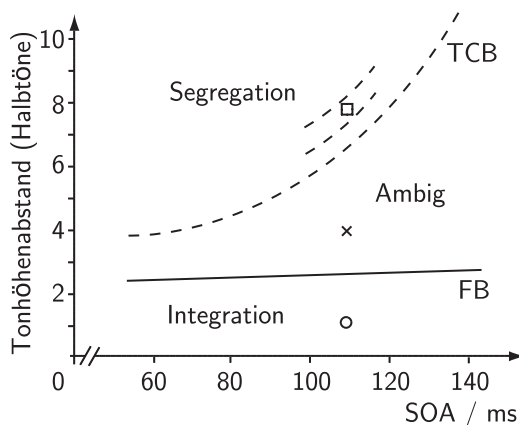


Abb. 6.1: Van-Noorden-Diagramm (nach Bregman 1999, S. 60). Eingelegt sind die Parameterkonfigurationen der vorliegenden Arbeit, die eine *per se* segregierte (Quadrat), integrierte (Kreis) oder ambige Organisation (Kreuz) erzeugen sollten. Die Ergebnisse deuten darauf hin, dass der Verlauf der TCB variabler als der Verlauf der FB zu sein scheint, da das (psychoakustisch) gemessene Perzept bei Stimuli mit großem Tonhöhenabstand bereits Anteile von integrierter Organisation enthält. Dargestellt sind deshalb auch weitere denkbare Verläufe der TCB.

Konfiguration mit engem Frequenzabstand entspricht das Perzept nicht, wie zunächst erwartet, über die gesamte Zeit einem integrierten Stream. Bei allen Probanden wurden Zeitintervalle beobachtet, in denen sie die Frequenzbereiche segregierten. Diese Befunde sind möglicherweise durch die Wahl der Stimulationsparameter bestimmt: Der genaue Verlauf der TCB ist quantitativ schwer vorherzusagen, weil er von vielen Stimulusparametern und Aufmerksamkeitseffekten abhängt sowie eine große interindividuelle Variabilität aufweist (Bregman 1999; Moore und Gockel 2002). Zudem hat van Noorden (1975) den bekannten Verlauf mit einem *ABAB*-Paradigma, also konstanten Frequenzen in jedem Stream gemessen. Daher ist es nicht klar, ob der Verlauf der TCB mit dem in einem $A_1B_2A_2B_1A_3B_3\dots$ -Paradigma übereinstimmt, bei dem in jedem Frequenzbereich verschiedene Töne präsentiert werden.

Interessanterweise hatte die visuelle Stimulation in der Stimuluskonfiguration mit weitem Frequenzabstand einen signifikanten Einfluss auf das auditorische Perzept: Durch synchrone Präsentation mit den Stimuli eines stream-übergreifenden Musters wurden die Stimuli länger als in einen Stream integriert wahrgenommen. Dies ist bemerkenswert, da hiermit auch für den *per se* eindeutigen Segregationsbereich die Modulation des auditorischen Perzepts durch synchrone visuelle Reize gezeigt wird. Dieser Befund ließ sich nicht auf die auditorische Stimulation mit engem Frequenzabstand zwischen akustischen Streams übertragen. Hier konnte keine Veränderung des Perzepts durch die zusätzliche Präsentation der visuellen Stimuli in Richtung segregierter Streams herbeigeführt werden. Betrachtet man das Frequenzverhältnis zwischen dem höchsten tieffrequenten und dem tiefsten hochfrequenten Ton in der Bedingung mit weitem Frequenzabstand, so liegt dieser mit sechs Halbtönen unter Verwendung des SOA von 110 ms schon unterhalb des Verlaufs der TCB nach van Noorden (1975). Ein Vergleich der Bedingungen ohne visuelle Stimulation zeigt eine Asymmetrie auf: Obgleich das integrierte Perzept in der Bedingung mit engem Frequenzabstand mit etwa 90 % fast permanent vorlag, lag es in der Bedingung mit weitem Frequenzabstand immerhin noch während 20 % der Messzeit vor. Es erscheint daher so, dass die Segregation in der Bedingung mit weitem Frequenzabstand unvollständig und daher – wie im mehrdeutigen Bereich – durch die integrierenden visuellen Informationen zu beeinflussen war.

Wurde im psychoakustischen Experiment das auditorische Streaming-Perzept ausgewertet, so sollte durch ein weiteres EEG-Experiment auch die präattentive Organisation der Stimuli in nicht-ambigen Stimuluskonfigurationen untersucht werden. Insbesondere die im psychoakustischen Experiment gewonnenen Evidenzen für die Empfindlichkeit der automatischen auditorischen Organisation für visuelle Stimuli sollten überprüft werden.

In der Bedingung mit engem Frequenzabstand und ohne visuelle Stimulation (*eng-auditorisch allein*) wurde keine MMN beobachtet. Somit konnten die in einem Frequenzbereich enthaltenen Unregelmäßigkeiten nicht präattentiv detektiert und damit auch die regelmäßigen Tripel nicht erkannt werden. Die auditorischen Stimuli wurden also erwartungsgemäß nicht segregiert. Durch die Präsentation eines visuellen Stimulus synchron zu jedem Ton der Tripel wurde in einer weiteren Bedingung (*eng-segregierend*) jedoch eine MMN evoziert. Somit wurde die zunächst klar integrierte Organisation durch den visuellen Einfluss hin zur Segregation verändert, was für eine Beeinflussung des auditorischen Perzepts durch visuelle Stimulation bei engem Frequenzabstand spricht. Dieser Befund deckt sich nicht mit dem des psychoakustischen Experiments. Der psychoakustische Zugang untersucht aber, wie oben beschrieben, das bewusste Perzept, welches durchaus von der präattentiven Verarbeitung abweichen kann.

Bei der Stimuluskonfiguration mit weitem Frequenzabstand konnte in der Bedingung ohne visuelle Stimulation (*weit-auditorisch allein*) die erwartete segregierte Organisation mit Hilfe der MMN, die zudem durch eine Polaritätsinversion an den Mastoiden bestätigt wurde, nachgewiesen werden. Durch die zusätzliche visuelle, mit dem stream-übergreifenden Muster synchronisierte Stimulation (*weit-integrierend*) verringerte sich bei gleicher auditorischer Stimuluskonfiguration zwar die Latenz der MMN, ihre Amplitude jedoch nicht. Das auditorische Perzept wurde hier offenbar durch die visuelle Stimulation nicht in Richtung Integration verschoben. Dieser Befund deckt sich daher nicht mit dem psychoakustischen Experiment. Möglicherweise ist die TCB für bewusste und unbewusste Wahrnehmung verschieden. Das hieße, dass die Stimuli mit weitem Frequenzabstand im EEG-Experiment stärker segregiert wahrgenommen wurden als im psychoakustischen Experiment. Dadurch wäre der Abstand zur TCB im EEG-Experiment so groß, dass eine Beeinflussbarkeit des Streaming-Perzepts nicht mehr möglich ist. Hier könnten weitere Experimente, die beide Methoden verwenden, Klarheit schaffen.

Die Kontrollbedingung in diesem EEG-Experiment war die Präsentation einer akustischen Tripel-Folge ohne zweiten Stream und visuelle Stimulation (*1 Stream-auditorisch allein*). Die dadurch evozierte MMN sollte als Vergleich für die zu erwartende MMN-Amplitude und Latenz der anderen Bedingungen dienen und wurde bei den meisten Probanden auch nachgewiesen. Im Gruppenmittelwert ist diese in ihrer Amplitude mit der MMN in den Bedingungen *eng-segregierend*, und *weit-auditorisch allein* vergleichbar. Dies spricht für den Nachweis einer von visuell evozierten Potentialen abgrenzbaren auditorischen MMN. Ihre Latenz ist etwas verkürzt, was durch die gegenüber den audiovisuellen Bedingungen verringerte Komplexität der Stimuluspräsentation erklärbar ist (Picton *et al.* 2000).

6.2 Audiovisuelle Interaktionen

6.2.1 Verarbeitungsstufe

In den durchgeführten elektrophysiologischen Experimenten wurde zur Charakterisierung des auditorischen Streamings die MMN als Index verwendet. Diese wird evoziert, wenn eine Regularität der auditorischen Reizfolge im sensorischen Gedächtnis repräsentiert wird, auf dessen Basis der Deviant-Detektionsprozess wirkt (Sussman 2005). Die Abweichungen von der Regularität wurden in den Experimenten durch die Umkehr von Tonfolgen (Trippel) realisiert, wie bereits für ähnliche Stimuluskonfigurationen beschrieben (Sussman *et al.* 1998a, 1999).

Unter Verwendung der Stimuluskonfiguration des EEG-Experiments im ambigen Bereich wurde durch die auditorischen Stimuli nur dann eine MMN evoziert, wenn die Stimuli eines Frequenzbereichs mit visuellen Reizen gekoppelt waren. Somit war die zunächst auditorische Repräsentation der Stimuli verändert. Durch die visuellen Stimuli muss dann eine Segregation der Frequenzbereiche bewirkt worden sein, denn die zufällige Frequenzverteilung der Stimuli eines Bereiches würde – bei der integrierten Wahrnehmung aller Stimuli – die im anderen Bereich enthaltene Regularität aufheben, und somit dem MMN-Mechanismus die Basis entziehen. Auf dieser frühen Stufe der neuronalen Verarbeitung, auf der die MMN arbeitet, muss sich also der Einfluss der visuellen Stimuli auf die auditorische Repräsentation bereits niedergeschlagen haben. Auch durch die Evozierung einer MMN in der *2 Streams–Segregation*-Bedingung und ihrer Nichtnachweisbarkeit bei gleichen auditorischen Stimulusparametern in der *2 Streams–Integration*-Bedingung wird gezeigt, dass die visuelle Stimulation die Formation von einem oder zwei Streams durch die Änderung der Erkennbarkeit von Regularitäten steuert. Die gefundene Modifikation der MMN zeigt daher einen Einfluss der visuellen Stimulation auf einer frühen Stufe der neuronalen Verarbeitung. Dieser Befund ist mit Studien konsistent, die ebenfalls Evidenzen für eine *early-level*-Interaktion zwischen visueller und auditorischer Modalität beschreiben (Besle *et al.* 2005; Giard und Peronnet 1999; Fort *et al.* 2002). Auch Schröger und Widmann (1998) untersuchten die audiovisuelle Integration in einem elektrophysiologischen Experiment und fanden, dass diese Interaktion auf einer frühen Verarbeitungsstufe und weit vor der Auslösung einer motorischen Antwort auf die Stimuli stattfindet.

Im MEG-Experiment dieser Arbeit wurde die Indizierung der automatischen Verarbeitung bei ambiger auditorischer Stimuluskonfiguration durch das Mismatch-Field (MMF), der biomagnetischen Komponente der MMN, vorgenommen. Die Ergebnisse sind in Über-

einstimmung mit denen des EEG-Experiments: Bei synchroner segregierender visueller Stimulation zeigen sie eine ausgeprägte MMF-Komponente. In der Bedingung *integrierend* hingegen wurde die Regularität der Stimuli eines Frequenzbereichs (Tripel) nicht erkannt, und somit kein nachweisbares MMF evoziert. Auch hierdurch zeigt sich die Eignung der MMN/MMF als Indikator für die automatische auditorische Wahrnehmung und somit die Charakterisierung von Streaming. Weil das MMF und die MMN gleichen neuronalen Ursprungs sind, sprechen die Ergebnisse des MEG-Experiments ebenfalls für eine audiovisuelle Interaktion bereits auf frühen Stufen der auditorischen Verarbeitung.

Das psychoakustische Experiment zeigt – zumindest im ambigen Bereich – eine deutliche Beeinflussung des attentiven auditorischen Perzepts durch die visuelle Stimulation. Die Probanden wurden hierbei nicht explizit nach den Perzepten *Integration* oder *Segregation* befragt, sondern einem indirekten AFC-Verfahren unterzogen. Dieses verlangte die Beschreibung des vorherrschenden auditorischen Perzepts (Drei- bzw. Zwei-Ton-Folge) unter Beachtung der visuellen Stimuli, wodurch eine zu starke Konzentration auf die auditorischen Stimuli vermieden wurde. Zudem wurde dadurch ausgeschlossen, dass allein die Detektionsaufgabe zur Veränderung des Perzepts führte. Bei ambiger Stimuluskonfiguration konnte hierdurch die in den EEG-Experimenten beschriebene Beeinflussung des auditorischen Perzepts durch visuelle Stimulation nachgewiesen werden. Die im EEG-Experiment nachweisbare Verstärkung eines segregierenden Perzepts war hier allerdings weniger stark ausgeprägt. Es zeigte sich jedoch eine mit dem EEG-Experiment vergleichbare Beeinflussung durch die integrierende visuelle Stimulation. Auch diese Ergebnisse lassen darauf schließen, dass die audiovisuelle Integration auf oder vor der Verarbeitungsstufe von Stream-Segregation abläuft. Ebenso wäre eine, wie auch immer geartete, attentive Rückkopplung vom visuellen auf das auditorische System möglich. Hierfür spricht, dass bimodale Interaktionen sowohl auf Verarbeitungsstufen mit einer neuronalen Latenz von etwa 40 ms, als auch auf mehreren späteren Gliedern der neuronalen Verarbeitungskette, sowohl in modalitätsspezifischen als auch unspezifischen Strukturen des Kortex zu finden sind (Giard und Peronnet 1999).

6.2.2 Einfluss der Aufmerksamkeit

In dem MEG- und den EEG-Experimenten wurde der Aufmerksamkeitszustand der Probanden durch eine mit den visuellen Reizen verbundene Aufgabe definiert. Deren Aufmerksamkeit sollte demnach nahezu vollständig auf die visuellen Stimuli fixiert sein. Die durchgeführten Experimente untersuchten in auditorisch passiven Bedingungen die durch

auditorische Reize evozierte MMN und das MMF. Sie weisen einen Einfluss der simultanen visuellen Reizpräsentation auf die perzeptuelle Organisation der Töne in Streams nach. Beim Bauchredner-Effekt – als ein Beispiel für audiovisuelle Interaktion – ist die auditorische Wahrnehmung unabhängig von der visuellen Aufmerksamkeit (Bertelson *et al.* 2000). In den Experimenten dieser Arbeit wurde dennoch eine fixierte Aufmerksamkeit angestrebt, um die Variabilität der Vigilanz gering zu halten, deren Veränderung Einfluss auf späte Potentiale nimmt (Woldorff *et al.* 1991). Zudem sollte der Einfluss der visuellen Stimuli auf die auditorische Modalität möglichst konstant sein.

In den Kontrollbedingungen der EEG-Experimente, in denen keine visuellen Stimuli präsentiert wurden, wurde die Aufmerksamkeit der Probanden ebenfalls auf die visuelle Modalität gelenkt. Wie bei MMN-Studien vielfach praktiziert (Pettigrew *et al.* 2004), wurde ein stummer Videofilm gezeigt. Die Aufmerksamkeit der Probanden ist damit von der auditorischen Stimulation abgelenkt und qualitativ vergleichbar mit den Bedingungen, in denen eine visuelle Aufgabe gestellt wurde. Das (passive) Ignorieren der auditorischen Stimuli hat nach Duncan *et al.* (1997) aber auch einen Einfluss auf die Deviantdetektion. Etwa 20–30 % der auditorischen Deviantreize werden demnach nicht mehr detektiert. Dies führt zu einer Abnahme der MMN-Amplitude, die im Vergleich zu anderen rein auditorischen MMN-Studien (z. B. Sussman *et al.* 1999) tatsächlich verringert ist.

Im psychoakustischen Experiment war die Aufmerksamkeitssituation im Vergleich zu den EEG- und MEG-Experimenten verändert. Hier wurde, um das bewusste auditorische Perzept evaluieren zu können, eine mit beiden Modalitäten verbundene Aufgabe gestellt. Die Probanden sollten sowohl auf die visuellen als auch auf die auditorischen Reize achten und dabei ihren vordergründigen Höreindruck beurteilen. Obwohl die Aufgabe keine diffizilen Anforderungen stellt, wird sie von den meisten Probanden auf Nachfrage dennoch als anstrengend beurteilt. Insbesondere das Aufteilen der Aufmerksamkeit zwischen den Modalitäten wurde als konzentrationsfordernd beschrieben. Dennoch haben alle Probanden das Experiment mit relativ wenig Pausenbedarf vollständig absolviert.

Die Auswertung der Bedingungen des psychoakustischen Experiments zeigt eine signifikante Verstärkung der Integration von auditorischen Stimuli verschiedener Frequenzbereiche durch eine entsprechende visuelle Stimulation, jedoch nur einen schwachen Einfluss auf die Segregation. Möglicherweise wirkt die höhere visuelle Stimulationsrate bei den segregierenden visuellen Stimuli der audiovisuellen Interaktion entgegen. Ein durch die visuellen Stimuli evozierter *top-down*-Prozess erfordert möglicherweise ein längeres SOA, um stärker zur audiovisuellen Integration beitragen zu können. Andererseits könnte der visuelle Einfluss auf auditorische Wahrnehmung auch asymmetrisch bezüglich Integration

und Segregation sein. Zudem ist bereits in der *auditorisch allein*-Bedingung ein Überhang des segregierten Perzepts vorhanden, wonach die Stimulusparameter möglicherweise etwas näher an der TCB zu liegen scheinen. Im Experiment wurden für die Segregation Frequenzmuster, aber für die Integration Intensitätsmuster mit den visuellen Stimuli synchronisiert. Auch dieser Unterschied wirkt sich möglicherweise auf die aufmerksamkeit-basierte audiovisuelle Interaktion aus, wie sie in diesem Experiment untersucht wurde. Diese Hypothese der Asymmetrie lässt sich durch die MMN-Experimente allerdings nicht beurteilen. Eine Verstärkung der Integration würde dort zwar zu einer Verringerung ihrer Amplitude führen, im ambigen Zustand kann sie aber durch die notwendige Bildung des Gruppenmittelwertes und den Zeitintervallen mit integriertem Perzept selbst ohne visuelle Stimulation nicht nachgewiesen werden. Zur Untersuchung der Asymmetrie könnten Folgeexperimente einen Beitrag leisten, bei denen die individuellen Wahrnehmungsgrenzen (TCB, FB) zunächst ermittelt werden, um dann mit individuell angepassten ambigen Stimuli zu arbeiten.

Frühere elektrophysiologische Experimente haben gezeigt, dass ein *top-down*-gesteuerter Einfluss auf die Wahrnehmung einer auditorischen Szene nur möglich ist, wenn die akustischen Parameter eine (automatische) ambige Wahrnehmung erzeugen (Sussman *et al.* 2002). In den hier durchgeführten EEG- und MEG Experimenten ist das auditorische Perzept zwar ambig, jedoch wurde dieses nicht durch gezielte Aufmerksamkeit auf die auditorischen Stimuli beeinflusst. Sie wurde vielmehr von den auditorischen Stimuli weg auf die visuellen Stimuli gelenkt. Es zeigt sich, dass in diesen selektiven Aufmerksamkeitszuständen die visuellen Stimuli – möglicherweise über den Umweg visueller Aufmerksamkeit – das auditorische Perzept beeinflussen. Dieser Befund ist konform mit den Ergebnissen von Sussman *et al.* (1998a), die eine Veränderung von auditorischer Verarbeitung durch selektive Aufmerksamkeit, allerdings ohne visuelle Stimulation, nachweisen.

6.2.3 Auditorische und visuelle MMN

Als Index für das auditorische Perzept segregierter Streams wurde die MMN verwendet. Dieses evozierte Potential wird neuronalen Generatoren im auditorischen Kortex zugeschrieben (Giard *et al.* 1990). Aber auch durch Irregularitäten in visuellen Stimulationssequenzen kann eine MMN evoziert werden, deren Generatoren im visuellen Kortex lokalisiert sind (Czigler *et al.* 2006). Die MMN beider Modalitäten zu separieren, gelingt durch Lokalisation der zu Grunde liegenden neuronalen Quellen und Beurteilung der SCD.

Schon durch einfache Analysen der Mittelwertkurven können Rückschlüsse auf die Lage der neuronalen MMN-Generatoren gezogen werden (Picton *et al.* 2000). Die in der EEG-Messung mit ambiger Stimuluskonfiguration enthaltene negative Komponente war nur in den Bedingungen mit segregierender visueller Stimulation (*2 Streams-segregierend*) und *1 Stream* signifikant. Würde diese allein durch die visuelle MMN erklärbar sein, müsste die Negativität auch in der *Visuell-allein*-Bedingung mit vergleichbarer Amplitude auftreten. Da dies nicht der Fall ist, und in der *1 Stream*-Bedingung eine auditorische MMN mit großer Wahrscheinlichkeit erwartet wurde (Sussman *et al.* 2005), spricht die Vergleichbarkeit mit der *2 Streams-segregierend*-Bedingung für die Präsenz der auditorischen MMN auch in der dort gemessenen Deviantantwort. Die Potentialverteilungen auf der Kopfhaut und die SCD wiesen für die *2 Streams-segregierend*-Bedingung neben den supratemporalen auch occipitale Foci auf, welche auf eine durch audiovisuelle Devianten evozierte MMN schließen lassen (Besle *et al.* 2005). Um diese von Veränderungen der visuell evozierten Potentiale abgrenzen zu können, hilft der Vergleich mit der *Visuell-allein*-Bedingung. Die Verteilung der occipital gemessenen visuell evozierten Antworten sind in beiden Bedingungen vergleichbar. Da jedoch nur in der *2 Streams-segregierend*-Bedingung auch auditorisch stimuliert wurde, lässt sich die visuelle MMN von den auditorisch evozierten Potentialen bei bimodaler Stimulation abgrenzen und daher die Modifizierung der auditorischen MMN durch visuelle Stimulation bestätigen.

Zusätzlich zu diesen Überlegungen wurde zunächst für das EEG-Experiment die Lage der neuronalen Generatoren durch ein Dipolmodell bestimmt, welches die gefundene MMN beschreibt (vgl. Kap. 6.2.3). In der *Visuell-allein*-Bedingung wurde ein Generatorenpaar im visuellen Kortex gefunden und bei bimodaler Stimulation in der *2 Streams-segregierend*-Bedingung durch ein weiteres Paar bilateral im auditorischen Kortex ergänzt. Dieser Befund entspricht der bekannten Lokalisierung bimodaler MMN-Komponenten und bestätigt somit deren Evozierung bei entsprechender Stimulation (Saint-Amour *et al.* 2007).

Die Genauigkeit der Quellenanalyse ist bei diesem Experiment begrenzt, da es auf dem Gruppenmittel beruht und somit individuelle Unterschiede nicht erfasst werden. Durch die unterschiedlichen Schädelformen der einzelnen Probanden ist es erklärbar, dass vereinzelt Generatoren außerhalb des Kortex lokalisiert werden. Qualitativ bestätigt die Quellenanalyse jedoch das Auftreten auditorischer und visueller MMN. Durch die Quellenmodellierung im EEG-Experiment ergaben sich somit klar voneinander abgegrenzte temporale und occipitale Generatoren. Ein Generatorenpaar wurde in temporalen Arealen symmetrisch zur Medianebene lokalisiert, ein weiteres in occipitalen Arealen. Da die auditorisch evozierte MMN hauptsächlich im auditorischen Kortex generiert wird (Alho 1995; Saint-Amour

et al. 2007), wurde diese offenbar auch im durchgeführten EEG-Experiment evoziert. Der Ursprung der visuellen MMN liegt im visuellen Kortex (Czigler *et al.* 2006; Maekawa *et al.* 2005; Tales *et al.* 1999). Das occipitale Generatorenpaar spricht daher für das Auftreten einer zusätzlichen visuellen MMN, und zwar durch die Umkehr einer aus Quadraten unterschiedlicher Größe bestehenden Stimulusfolge.

Das MEG-Experiment wurde durchgeführt, um die Genauigkeit der Quellenlokalisierung zu erhöhen. Auch hier konnte nur durch die Annahme zweier Generatorenpaare durch die Quellenlokalisierung in der *segregierend*-Bedingung das gemessene MEG-Signal erklärt werden. Für die *integrierend*-Bedingung wurde keine Quellenanalyse durchgeführt, da hier keine signifikante Differenzantwort auftrat. Die individuellen Daten weisen eine große interindividuelle Streuung auf, können aber, wie schon im EEG-Experiment, dem visuellen und auditorischen Kortex zugeordnet werden. Diese Trennung zeigt, dass in der Differenzantwort der *segregierend*-Bedingung ein MMF evoziert wurde und in der *integrierend*-Bedingung nicht. Auch die für die Gruppenmittel-Antwort berechneten vier Generatoren stimmen in ihrer Lage mit dem EEG-Experiment überein. Sie liegen nicht im räumlichen Zentrum der individuellen Generatoren, was aus den unterschiedlich stark ausgeprägten Antworten der einzelnen Probanden folgt. Die Unsicherheit zwischen der Generator-Lokalisation in individuellen Datensätzen und Gruppenmittelwerten liegt in der Größenordnung von etwa 1 cm und stimmt mit für derartige Messungen üblichen Werten überein (Whittingstall *et al.* 2004). Durch die biomagnetischen Messungen konnten somit im Differenzfeld (Deviant minus Standard) der *segregierend*-Bedingung auditorisch und visuell evozierte Antworten gefunden werden. Daraus kann abgeleitet werden, dass sowohl durch visuelle, aber auch durch auditorische Deviantstimuli ein MMF evoziert wurde. Sowohl im MEG- als auch im EEG-Experiment wurde die gleiche ambige Stimuluskonfiguration verwendet. Deshalb können die im jeweiligen Experiment gefundenen Magnetfelder und Potentiale als Komplementärbefunde gewertet werden. Sie liefern damit für die durchgeführten Experimente klare Evidenzen für die Evozierung einer MMN in beiden, der visuellen und auditorischen Modalität.

Der in den Bereichen nicht-ambiger auditorischer Stimuluskonfiguration gefundene Effekt synchroner visueller Stimulation auf das auditorische Perzept wurde ebenfalls durch Analyse der Potentialverteilungen, SCD und Quellenlokalisierung verifiziert. Der visuelle Einfluss auf die auditorische Organisation führte zwischen den Bedingungen *eng-segrierend* und *eng-auditorisch allein* zu signifikant verschiedenen MMN-Amplituden, verfehlte zwischen den Bedingungen mit weitem Frequenzabstand (*weit-integrierend* und *weit-auditorisch allein*) jedoch die Signifikanz. Bei Betrachtung der Potentialverteilung

auf der Kopfhaut wurde durch die zusätzliche visuelle Stimulation bei weitem Frequenzabstand der Negativitäts-Focus in Richtung zentraler Elektroden verschoben, blieb aber deutlich bestehen. Eine solche zentrale Negativität in Differenzpotentialen belegt eine Verstärkung der N1-Amplitude und ist ebenfalls ein neuronales Korrelat für die Detektion auditorischer Änderungen (Elangovan *et al.* 2005; Hyde 1997; Rinne *et al.* 2006). Bei Stimulation mit engem Frequenzabstand veränderte sich die Potentialverteilung sowie die SCD durch zusätzliche visuelle Stimulation derart, dass die für die MMN typischen frontalen Foci erkennbar sind. Dies bestätigt den deutlichen Effekt der visuellen Stimulation zur Erzeugung eines segregierenden auditorischen Perzepts. Die Quellenanalyse ergab Generatoren, deren Lage konform mit den beiden anderen Experimenten ist. In den Bedingungen, in denen eine MMN evoziert wurde, konnte diese je nach Stimulationsart dem visuellen oder auditorischen Kortex zugeordnet werden.

Kapitel 7

Zusammenfassung und Ausblick

Die vorliegende Arbeit untersucht die Wahrnehmung von Folgen abwechselnd hoher und tiefer Töne. Dabei spielen Mechanismen eine Rolle, die über die auditorische Sinnesmodalität hinausgehen, so auch die in dieser Arbeit untersuchte audiovisuelle Interaktion. In einem komplexen akustischen Stimulationsparadigma wurde dazu in zwei getrennten Frequenzbereichen mit jeweils drei kurzen Sinustönen unterschiedlicher Frequenz stimuliert. In einem der Frequenzbereiche folgten die Töne einem sich wiederholenden, aufsteigenden Tripel-Muster, während im anderen Frequenzbereich die Abfolge zufällig war. Frequenz und Stimulationsrate waren zunächst so gewählt, dass eine ambige Hörsituation entstand. Das heißt, die auditorischen Stimuli waren entweder als eine oder als zwei separate Tonfolgen (Streams) wahrnehmbar.

Es konnte gezeigt werden, dass mit Hilfe von synchron zu den Tripel gezeigten visuellen Stimuli die präattentive perzeptuelle Organisation der selben auditorischen Stimuli verändert werden kann. Der Nachweis erfolgte über ein ereigniskorreliertes Potential, die MMN: Wurden die visuellen Stimuli synchron zu den Stimuli der Tripel präsentiert, konnte durch Umkehr der Tripel eine MMN evoziert werden. Diese war nicht mehr nachweisbar, wenn die visuellen Stimuli synchron zu einem stream-übergreifenden Intensitätsmuster präsentiert wurden.

Diese mit elektrophysiologischen Methoden gewonnenen Ergebnisse zeigen eindrucksvoll, dass bereits auf einer präattentiven Verarbeitungsstufe die Organisation von Tönen im ambigen Bereich zu integrierten oder segregierten Streams verschoben werden kann. Mit einem psychoakustischen Experiment, welches das gleiche Stimulationsmaterial enthielt, wurden diese Ergebnisse durch die Untersuchung des Perzepts beim Zuhören mit gerich-

teter Aufmerksamkeit überprüft. Die Aufmerksamkeit der Probanden wurde hierbei – im Gegensatz zum EEG-Experiment – sowohl auf die visuellen Reize als auch gezielt auf die akustische Szene gelenkt und deren vorherrschendes auditorisches Perzept abgefragt. Auch bei dieser komplexen Aufmerksamkeitssituation wurde das zunächst ambige auditorische Perzept durch visuelle Stimuli hin zu den eindeutigen Perzepten Integration und Segregation verändert. Die Arbeit zeigt daher eine audiovisuelle Interaktion, die übereinstimmend sowohl präattentiv als auch in der subjektiven Wahrnehmung nachweisbar ist.

Wenn durch visuelle Stimuli das auditorische Perzept einer ambigen Hörsituation verändert werden kann, liegt die Frage nach einer möglichen Übertragung auf auditorische Stimulationsparadigmen mit *per se* eindeutiger Wahrnehmung nahe. Das psychoakustische Experiment enthielt deshalb auch Stimulationssequenzen mit vergrößertem und mit verkleinertem Frequenzabstand zwischen den akustischen Streams, so dass ohne visuelle Stimulation *per se* von einer automatischen Segregation oder Integration der Streams ausgegangen werden kann. Das zeigen – mit gewissen Einschränkungen – die psychoakustischen Messungen. Bei diesen Sequenzen, die zunächst ein eindeutiges Perzept erwarten lassen, zeigt sich bei zusätzlicher Verwendung synchroner visueller Stimuli eine teilweise Verschiebung des Perzepts. So wurde bei Verwendung der zur automatischen Segregation führenden Stimuluskonfiguration mit großem Frequenzabstand teilweise das Perzept eines integrierten Streams gemessen. Im Bereich mit engem Frequenzabstand blieb das gemessene Perzept trotz segregierender visueller Stimulation unverändert das eines integrierten Streams.

Dieser Befund sollte auch elektrophysiologisch bestätigt werden. Dazu wurde ein weiteres EEG-Experiment unter Verwendung eines im Vergleich zur ambigen Situation vergrößerten oder verkleinerten Frequenzbereichs-Abstand durchgeführt. Bei großem Abstand wurde dabei eine gegenüber der visuellen Stimulation relativ robuste MMN gefunden. Diese zeigt, im Gegensatz zum psychoakustischen Experiment, eine durch visuelle Stimuli unveränderte Organisation der auditorischen Stimuli als zwei segregierte Streams an. Bei Verwendung des kleinen Frequenzabstands wurde bei zusätzlicher segregierender visueller Stimulation eine MMN gefunden. Hier scheint somit durch die visuelle Stimulation die automatische auditorische Organisation modifizierbar zu sein. Offenbar sind die Prozesse der präattentiven Organisation von Tönen, wie sie mit Hilfe der MMN untersucht werden können, in diesen *per se* eindeutigen Bereichen nicht deckungsgleich mit dem subjektiven Perzept der Probanden. Der Einfluss weiterer *top-down*-Prozesse bleibt daher noch zu klären.

Um einen genaueren Zugang zu den neuronalen Generatoren der evozierten Signale wie der MMN zu erhalten, wurde ein MEG-Experiment mit ebenfalls ambiger Stimuluskonfiguration durchgeführt. Dabei konnten die Ergebnisse des EEG-Experiments bestätigt werden und in jeder Hemisphäre zwei Generatoren lokalisiert werden. Jeweils zwei sind temporal und occipital gelegen. Das zeigt, dass sowohl eine auditorische als auch eine visuelle MMN gemessen wurde.

Die durchgeführten Experimente stellen einen Beitrag zu aktuellen Fragestellungen der Verarbeitung multimodaler Reize dar. Effekte, die bisher bei hochkomplexen Reizen, wie z. B. Sprache und Mundbewegungen, untersucht wurden, sind oft recht eindrucksvoll. Die vorliegende Arbeit bedient sich dahingegen einfacher, physikalisch gut handhabbarer Reize wie Sinustönen und einfachen geometrischen Figuren. Sie kann durch diesen induktiven Zugang einen audiovisuellen Effekt nachweisen, der möglicherweise Grundlage für kognitive, psychologische Phänomene im Zusammenhang mit der auditorischen Szenenanalyse ist. Untersuchungen zeigen, dass technische Hörhilfen keine mit dem auditorischen System vergleichbare Klassifizierungs- und Quellentrennleistung erreichen, wenn die Mechanismen wie Grouping und Streaming nicht berücksichtigt werden (Dillier 2005). Auch deshalb ist die weitere Untersuchung dieser Prozesse von großem praktischen Nutzen.

Durch diese Arbeit wird das relativ unerforschte Gebiet der Wahrnehmung einer ambigen auditorischen Szene untersucht. Hier sind weiterführende Experimente notwendig, insbesondere um den Einfluss der Aufmerksamkeit und damit der aktiven Beeinflussung der Wahrnehmung (*top-down*-Prozesse) zu untersuchen. Dabei kann der Einsatz von funktionseller Kernspintomographie sinnvoll sein, um auch die höheren kognitiven Prozesse im Zusammenhang mit Streaming zu verstehen (Cacace *et al.* 2000). Aber auch auf dem Niveau von Einzelzelleableitungen an Primaten können weitere Experimente wertvolle Ergebnisse zu veränderten Perzepten geben. Der ambige Bereich bietet zudem die Möglichkeit, mit identischen Stimuli zu arbeiten, also die Kovarianz von den durch veränderte Stimuluseigenschaften aktivierten Neuronen und dem auditorischen Perzept zu vermeiden.

Eine Modellierung des Streaming-Phänomens ist, wie bereits in Kap. 2.2.3 beschrieben, selbst für die (isoliert betrachtete) auditorische Modalität noch in der Entwicklung und kann bisher nicht alle Eigenschaften simulieren. Unter Einbeziehung der visuellen Modalität und ihrer Interaktion mit dem auditorischen Streaming würden die Modelle um weitere Gleichungen für die visuelle Modalität sowie Termen für die Wechselwirkung ergänzt werden müssen. Dies ist nicht Gegenstand der vorliegenden Arbeit, könnte jedoch bei Vorliegen umfangreicherer Daten ein interessantes Thema für weiterführende Arbeiten sein.

Weiterführende Experimente sollten zudem den Einfluss der visuellen Präsentationsrate auf die audiovisuelle Interaktion systematisch untersuchen. Hier sind insbesondere die Anforderungen für die Synchronisation zwischen beiden Reiztypen und an die Gestalt der visuellen Reize nur ansatzweise aufgeklärt. Auch sind verstärkt sprachbasierte audiovisuelle Paradigmen in elektrophysiologischen Experimenten denkbar, um weitere Informationen über die neuronalen Prozesse der audiovisuellen Integration zu gewinnen. Ein prominentes Beispiel hierzu ist das Lippenablesen, wie es zur Kommunikation zwischen Menschen mit vermindertem Gehör verwendet wird. Wenn die Beeinflussung von auditorischer Perzeption durch visuelle Stimuli möglich ist, erscheint auch der umgekehrte Fall interessant. Hier ergeben sich Ansätze zur Untersuchung, wie die Wahrnehmung visueller Muster durch die Präsentation auditorischer Stimuli verändert werden kann.

Im Hinblick auf eine alternde Gesellschaft muss auch von einer zunehmenden Anzahl an hör- oder sehbehinderten Menschen ausgegangen werden (Remijn *et al.* 2004). Mit dem Alter nimmt zudem Fähigkeit zu selektiver Aufmerksamkeit ab, was mit der eingeschränkten Fähigkeit zur Inhibierung irrelevanter Stimuli zusammenhängt (Hasher *et al.* 1991). In dieser Hinsicht und im Hinblick auf die allgegenwärtige Reizüberflutung könnte die Entwicklung effektiver Warnsignale im täglichen Leben durch verstärkte Verwendung audiovisueller Signale erreicht werden, welche die selektive Wahrnehmung der relevanten Information fördern. Dabei sollte auch berücksichtigt werden, dass die Sensitivität für audiovisuelle Synchronität zwischen Musik- und Sprachwahrnehmung verschieden ist und nach Vatakis und Spence (2006) bei Sprache stärker ausgeprägt ist. Zu dieser Problemstellung liefert die vorliegende Arbeit einen Beitrag, indem sie zeigt, dass audiovisuelle Effekte für die Wahrnehmung relevant sind, und deshalb auch bei der Entwicklung von Lern- und Kommunikationshilfen deren Effizienz steigern können.

Abkürzungsverzeichnis

A/D-Rate	Abtastrate bei Analog-Digital-Wandlung
ANOVA	Analysis of Variance (Varianzanalyse)
AEP	auditorisch evoziertes Potential
AFC	Alternative Forced Choice
BOLD	Blood Oxygen Level Dependency
ECD	Equivalent Current Dipole
EEG	Elektroenzephalogramm
EOG	Eletrookulogramm
EKP	Ereigniskorreliertes Potential
EPSP	exitatorisches postsynaptisches Potential
FAEP	frühe akustisch evozierte Potentiale
FB	Fission Boundary
fMRI	Functional Magnetic Resonance Imaging (Kernspintomographie)
HEOG	horizontales Eletrookulogramm
IPSP	inhibitorisches postsynaptisches Potential
LM	linkes Mastoid
MAEP	mittellatente akustisch evozierte Potentiale
MEG	Magnetenzephalographie / -enzephalogramm
MMF	Mismatch-Field
MMN	Mismatch-Negativity

MW	Mittelwert
PCA	Principal Component Analysis (Hauptkomponentenzerlegung)
PET	Positron-Emissions-Tomographie
RM	rechtes Mastoid
SAEP	späte akustisch evozierte Potentiale
SCD	Source Current Density (Quellendichte)
SD	Standard Deviation (Standardabweichung)
SOA	Stimulus Onset Asynchrony
SPL	Sound Pressure Level (Schalldruckpegel)
SQUID	Superconducting Quantum Interference Device
TCB	Temporal Coherence Boundary
VEOG	vertikales Elektrookulogramm
VEP	visuell evoziertes Potential

Literaturverzeichnis

Alain und Woods 1997

ALAIN, C. ; WOODS, D. L.: Attention modulates auditory pattern memory as indexed by event-related brain potentials. In: *Psychophysiology* 34 (1997), Nr. 5, S. 534–546

Alho 1995

ALHO, K.: Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. In: *Ear and Hearing* 16 (1995), Nr. 1, S. 38–51

Alho et al. 1996

ALHO, K. ; TERVANIEMI, M. ; HUOTILAINEN, M. ; LAVIKAINEN, J. ; TIITINEN, H. ; ILMONIEMI, R. J. ; KNUUTILA, J. ; NÄÄTÄNEN, R.: Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. In: *Psychophysiology* 33 (1996), Nr. 4, S. 369–375

Almonte et al. 2005

ALMONTE, F. ; JIRSA, V. K. ; LARGE, E. W. ; TULLER, B.: Integration and segregation in auditory streaming. In: *Physica D* 212 (2005), Nr. 1–2, S. 137–159

Altenmüller 1996

ALTENMÜLLER, E.: Ereigniskorrelierte Potentiale. In: STÖHR, M. (Hrsg.) ; DICHGANS, J. (Hrsg.) ; BUETTNER, U.W. (Hrsg.): *Evozierte Potentiale*. 3. Aufl. Berlin : Springer, 1996, S. 557–587

Altenmüller et al. 1996

ALTENMÜLLER, E. ; RUETHER, K. ; DICHGANS, J.: Visuell evozierte Potentiale (VEP) und Elektroretinogramm. In: STÖHR, M. (Hrsg.) ; DICHGANS, J. (Hrsg.) ; BUETTNER, U.W. (Hrsg.): *Evozierte Potentiale*. 3. Aufl. Berlin : Springer, 1996, S. 289–409

Anstis und Saida 1985

ANSTIS, S. ; SAIDA, S.: Adaptation to auditory streaming of frequency-modulated tones. In: *Journal of Experimental Psychology: Human Perception and Performance* 11 (1985), Nr. 3, S. 257–271

Atienza et al. 2003

ATIENZA, M. ; CANTERO, J. L. ; GRAU, C. ; GOMEZ, C. ; DOMINGUEZ-MARIN, E. ; ESCERA, C.: Effect of temporal encoding on auditory object formation: A mismatch negativity study. In: *Cognitive Brain Research* 16 (2003), Nr. 3, S. 359–371

Baker et al. 2000

BAKER, K. L. ; WILLIAMS, S. M. ; RODERICK, I. N.: Evaluating frequency proximity in stream segregation. In: *Perception and Psychophysics* 62 (2000), Nr. 1, S. 81–88

Barraclough et al. 2005

BARRACLOUGH, N. E. ; XIAO, D. ; BAKER, C. I. ; ORAM, M. W. ; PERRET, D. I.: Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. In: *Journal of Cognitive Neuroscience* 17 (2005), Nr. 3, S. 377–391

Beauvois und Meddis 1991

BEAUVOIS, M. W. ; MEDDIS, R.: A Computer model of auditory stream segregation. In: *The Quarterly Journal of Experimental Psychology* 43A (1991), Nr. 3, S. 517–541

Beauvois und Meddis 1996

BEAUVOIS, M. W. ; MEDDIS, R.: Computer simulation of auditory stream segregation in alternating-tone sequences. In: *Journal of the Acoustical Society of America* 99 (1996), Nr. 4, S. 2270–2280

Beauvois und Meddis 1997

BEAUVOIS, M. W. ; MEDDIS, R.: Time decay of auditory stream biasing. In: *Perception and Psychophysics* 59 (1997), Nr. 1, S. 81–86

Bertelson et al. 2000

BERTELSON, P. ; VROOMEN, J. ; DE GELDER, B. ; DRIVER, J.: The ventriloquist effect does not depend on the direction of deliberate visual attention. In: *Perception and Psychophysics* 62 (2000), Nr. 2, S. 321–332

Berti und Schröger 2006

BERTI, S. ; SCHRÖGER, E.: Visual distraction: A behavioral and event-related brain potential study in humans. In: *Neuroreport* 17 (2006), Nr. 2, S. 151–155

Besle et al. 2005

BESLE, J. ; FORT, A. ; GIARD, M. H.: Is the auditory sensory memory sensitive to visual information? In: *Experimental Brain Research* 166 (2005), Nr. 3–4, S. 337–344

Bey und McAdams 2002

BEY, C. ; MCADAMS, S.: Schema-based processing in auditory scene analysis. In: *Perception and Psychophysics* 64 (2002), Nr. 5, S. 844–854

Blackman und Tukey 1959

BLACKMAN, R. B. ; TUKEY, J. W.: *The measurement of power spectra from the point of view of communications engineering*. New York : Dover Publications, 1959

Bodenmann und Schaer 2006

BODENMANN, G. ; SCHAER, M.: Gestaltpsychologie. In: *Sprache, Stimme, Gehör* 30 (2006), Nr. 1, S. 21–23

Botte et al. 1997

BOTTE, M. C. ; DRAKE, C. ; BROCHARD, R.: Perceptual attenuation of nonfocused auditory streams. In: *Perception and Psychophysics* 59 (1997), Nr. 3, S. 419–425

Bozzi und Vicario 1960

BOZZI, P. ; VICARIO, G.: Due fattori di unificazione fra note musicali: La vicinanza temporale e la vicinanza tonale. In: *Rivista di Psicologia* 54 (1960), S. 235–258

Breckow und Greinert 1994

BRECKOW, J. ; GREINERT, R.: *Biophysik*. Berlin : Walter de Gruyter, 1994

Bregman 1978

BREGMAN, A. S.: Auditory streaming is cumulative. In: *Journal of Experimental Psychology* 4 (1978), Nr. 3, S. 380–387

Bregman 1999

BREGMAN, A. S.: *Auditory scene analysis: The perceptual organisation of sounds*. Cambridge, Massachusetts : The MIT Press, 1999

Bregman et al. 2000

BREGMAN, A. S. ; AHAD, P. A. ; CRUM, P. A. C. ; O'REILLY, J.: Effects of time intervals and tone durations on auditory stream segregation. In: *Perception and Psychophysics* 62 (2000), Nr. 3, S. 626–636

Bregman und Campbell 1971

BREGMAN, A. S. ; CAMPBELL, J.: Primary auditory stream segregation and perception of order in rapid sequences of tones. In: *Journal of Experimental Psychology* 89 (1971), Nr. 2, S. 244–249

Broadbent 1958

BROADBENT, D.: *Perception and Communication*. New York : Pergamon Press, 1958

Cacace et al. 2000

CACACE, A. T. ; TASCIYAN, T. ; COUSINS, J. P.: Principles of functional magnetic resonance imaging: Application to auditory neuroscience. In: *Journal of the American Academy of Audiology* 11 (2000), Nr. 5, S. 239–272

Calvert 2001

CALVERT, G. A.: Crossmodal processing in the human brain: Insights from functional neuroimaging studies. In: *Cerebral Cortex* 11 (2001), Nr. 12, S. 1110–1123

Calvert et al. 1999

CALVERT, G. A. ; BRAMMER, M. ; BULLMORE, E. T. ; CAMPBELL, R. ; IVERSEN, S. D. ; DAVID, A. S.: Response amplification in sensory-specific cortices during crossmodal binding. In: *Neuroreport* 10 (1999), Nr. 12, S. 2619–2623

Carlyon 2004

CARLYON, R. P.: How the brain separates sounds. In: *Trends in Cognitive Sciences* 8 (2004), Nr. 10, S. 465–471

Carlyon et al. 2001

CARLYON, R. P. ; CUSACK, R. ; FOXTON, J. M. ; ROBERTSON, I. H.: Effects of attention and unilateral neglect on auditory stream segregation. In: *Journal of Experimental Psychology* 27 (2001), Nr. 1, S. 115–127

Colin et al. 2004

COLIN, C. ; RADEAU, M. ; DELTENRE, A. Soquetand P.: Generalization of the generation of an MMN by illusory McGurk percepts: Voiceless consonants. In: *Clinical Neurophysiology* 115 (2004), Nr. 9, S. 1989–2000

Colin et al. 2002a

COLIN, C. ; RADEAU, M. ; SOQUET, A. ; DACHY, B. ; DELTENRE, P.: Electrophysiology of spatial scene analysis: The mismatch negativity (MMN) is sensitive to the ventriloquism illusion. In: *Clinical Neurophysiology* 113 (2002), Nr. 4, S. 507–518

Colin et al. 2002b

COLIN, C. ; RADEAU, M. ; SOQUET, A. ; DEMOLIN, D. ; COLIN, F. ; DELTENRE, P.: Mismatch negativity evoked by the McGurk-MacDonald effect: A phonetic representation within short-term memory. In: *Clinical Neurophysiology* 113 (2002), Nr. 4, S. 495–506

Cowan 1984

COWAN, N.: On short and long auditory stores. In: *Psychological Bulletin* 96 (1984), Nr. 2, S. 341–370

Cowan 1995

COWAN, N.: Sensory memory and its role in information processing. In: KARMOS, G. (Hrsg.) ; MOLNAR, M. (Hrsg.) ; CSEPE, V. (Hrsg.): *Perspective of Event-Related Potentials Research, EEG Supplement 40*. New York : Elsevier, 1995, S. 21–31

Crank und Nicolson 1996

CRANK, J. ; NICOLSON, P.: A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. In: *Advances in Computational Mathematics* 6 (1996), Nr. 3–4, S. 207–226

Cusack 2005

CUSACK, R.: The intraparietal sulcus and perceptual organization. In: *Journal of Cognitive Neuroscience* 17 (2005), Nr. 4, S. 641–651

Cusack et al. 2004

CUSACK, R. ; DEEKS, J. ; AIKMAN, G. ; CARLYON, R. P.: Effects of location, frequency region, and time course of Selective attention on auditory scene analysis. In: *Journal of Experimental Psychology* 30 (2004), Nr. 4, S. 643–656

Cusack und Roberts 2000

CUSACK, R. ; ROBERTS, B.: Effects of differences in timbre on sequential grouping. In: *Perception and Psychophysics* 62 (2000), Nr. 5, S. 1112–1120

Czigler et al. 2006

CZIGLER, I. ; WEISZ, J. ; WINKLER, I.: ERPs and deviance detection: Visual

mismatch negativity to repeated visual stimuli. In: *Neuroscience Letters* 401 (2006), Nr. 1–2, S. 178–182

Dalebout und Fox 2001

DALEBOUT, S. D. ; FOX, L. G.: Reliability of the mismatch negativity in the responses of individual listeners. In: *Journal of the American Academy of Audiology* 12 (2001), Nr. 5, S. 245–253

Deike et al. 2004

DEIKE, S. ; GASCHLER-MARKEFSKI, B. ; BRECHMANN, A. ; SCHEICH, H.: Auditory stream segregation relying on timbre involves left auditory cortex. In: *Neuroreport* 15 (2004), Nr. 9, S. 1511–1514

Dillier 2005

DILLIER, N.: Signalverarbeitung für die rehabilitative Audiologie. In: *XVII. Winterschule für Medizinische Physik*. Pichl/Steiermark, 2005

Dodd 1977

DODD, B.: The role of vision in the perception of speech. In: *Perception* 6 (1977), Nr. 1, S. 31–40

Duncan et al. 1997

DUNCAN, J. ; MARTENS, S. ; WARD, R.: Restricted attentional capacity within but not between sensory modalities. In: *Nature* 387 (1997), Nr. 6635, S. 808–810

Eggermont und Ponton 2002

EGGERMONT, J. J. ; PONTON, C. W.: The neurophysiology of auditory perception: From single units to evoked potentials. In: *Audiology and Neuro-Otology* 7 (2002), Nr. 2, S. 71–99

Elangovan et al. 2005

ELANGOVAN, S. ; CRANFORD, J. L. ; WALKER, L. ; STUART, A.: A comparison of the mismatch negativity and a differential waveform response. In: *International Journal of Audiology* 44 (2005), Nr. 11, S. 637–646

Engel et al. 2001

ENGEL, A. K. ; FRIES, P. ; SINGER, W.: Dynamic predictions: Oscillations and synchrony in top-down processing. In: *Nature Reviews Neuroscience* 2 (2001), Nr. 10, S. 704–716

Escera et al. 2000

ESCERA, C. ; ALHO, K. ; SCHRÖGER, E. ; WINKLER, I.: Involuntary attention and distractibility as evaluated with event-related brain potentials. In: *Audiology and Neuro-Otology* 5 (2000), Nr. 3-4, S. 151–166

Field 2005

FIELD, A.: *Discovering statistics using SPSS*. London : Sage Publications, 2005

Fishman et al. 2004

FISHMAN, Y. ; AREZZO, J. C. ; STEINSCHNEIDER, M.: Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate, and tone duration. In: *Journal of the Acoustical Society of America* 116 (2004), Nr. 3, S. 1656–1670

Fort et al. 2002

FORT, A. ; DELPUECH, C. ; PERNIE, J. ; GIARD, M. H.: Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. In: *Cerebral Cortex* 12 (2002), Nr. 10, S. 1031–1038

Frassinetti et al. 2005

FRASSINETTI, F. ; BOLOGNINI, N. ; BOTTARI, D. ; BONORA, A. ; LADAVAS, E.: Audiovisual integration in patients with visual deficit. In: *Journal of Cognitive Neuroscience* 17 (2005), Nr. 9, S. 1442–1452

Geddes und Baker 1967

GEDDES, L. A. ; BAKER, L. E.: The specific resistance of biological material—a compendium of data for the biomedical engineer and physiologist. In: *Medical & Biological Engineering* 5 (1967), Nr. 3, S. 271–293

Gelfand 1998

GELFAND, S. A.: *Hearing*. Basel : Marcel Dekker, 1998

Gentilucci und Cattaneo 2005

GENTILUCCI, M. ; CATTANEO, L.: Automatic audiovisual integration in speech perception. In: *Experimental Brain Research* 167 (2005), Nr. 1, S. 66–75

Giard et al. 1995

GIARD, M. H. ; LAVIKAINEN, J. ; REINIKAINEN, K. ; PERRIN, F. ; BERTRAND, O. ; PERNIER, J. ; NÄÄTÄNEN, R.: Separate representation of stimulus frequency,

intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. In: *Journal of Cognitive Neuroscience* 7 (1995), Nr. 2, S. 133–143

Giard und Peronnet 1999

GIARD, M. H. ; PERONNET, F.: Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. In: *Journal of Cognitive Neuroscience* 11 (1999), Nr. 5, S. 473–490

Giard et al. 1990

GIARD, M. H. ; PERRIN, F. ; PERNIER, J. ; BOUCHET, P.: Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. In: *Psychophysiology* 27 (1990), Nr. 6, S. 627–640

Gockel und MicheyL 1999

GOCKEL, H. ; MICHEYL, R. P. Carlyon C.: Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes. In: *Journal of the Acoustical Society of America* 106 (1999), Nr. 6, S. 3553–3563

Green und Swets 1966

GREEN, D. M. ; SWETS, J. A.: *Signal Detection Theory and Psychophysics*. New York : Wiley, 1966

Gutschalk et al. 2005

GUTSCHALK, A. ; MICHEYL, C. ; MELCHER, J. R. ; RUPP, A. ; SCHERG, M. ; OXENHAM, A. J.: Neuromagnetic correlates of streaming in human auditory cortex. In: *The Journal of Neuroscience* 25 (2005), Nr. 22, S. 5382–5388

Guttman et al. 2005

GUTTMAN, S. E. ; GILROY, L. A. ; BLAKE, R.: Hearing what the eyes see: Auditory encoding of visual temporal sequences. In: *Psychological Science* 16 (2005), Nr. 3, S. 228–235

Hari et al. 1992

HARI, R. ; RIF, J. ; TIHONEN, J. ; SAMS, M.: Neuromagnetic mismatch fields to single and paired tones. In: *Electroencephalography and Clinical Neurophysiology* 82 (1992), Nr. 2, S. 152–154

Hartmann und Johnson 1991

HARTMANN, W. M. ; JOHNSON, D.: Stream segregation and peripheral channeling. In: *Music Perception* 9 (1991), S. 155–184

Hasher et al. 1991

HASHER, L. ; STOLTZFUS, E. R. ; ZACKS, R. T. ; RYPMA, B.: Age and inhibition. In: *Journal of experimental psychology. Learning, memory, and cognition* 17 (1991), Nr. 1, S. 163–169

Hämäläinen et al. 1993

HÄMÄLÄINEN, M. ; HARI, R. ; ILMONIEMI, R. J. ; KNUUTILA, J. ; LOUNASMAA, O. V.: Magnetoencephalography – Theory, instrumentation, and applications to noninvasive studies of the working human brain. In: *Reviews of Modern Physics* 65 (1993), Nr. 2, S. 413–497

Hoppe et al. 1982

HOPPE, W. ; LOHMANN, W. ; MARKL, H. ; ZIEGLER, H.: *Biophysik*. 2. Aufl. Berlin : Springer, 1982

Hyde 1997

HYDE, M.: The N1 response and its applications. In: *Audiology and Neuro-Otology* 2 (1997), Nr. 5, S. 281–307

Jacobsen und Schröger 2001

JACOBSEN, T. ; SCHRÖGER, E.: Is there pre-attentive memory-based comparison of pitch? In: *Psychophysiology* 38 (2001), Nr. 4, S. 723–727

Jacobsen und Schröger 2003

JACOBSEN, T. ; SCHRÖGER, E.: Measuring duration mismatch negativity. In: *Clinical Neurophysiology* 114 (2003), Nr. 6, S. 1133–1143

Jacobson 1994

JACOBSON, J. T.: *Principles & Applications in Auditory Evoked Potentials*. Boston : Allyn and Bacon, 1994

Jasper 1958

JASPER, H. H.: The ten-twenty electrode system of the international federation. In: *Electroencephalography and Clinical Neurophysiology* 10 (1958), S. 371–375

Javitt et al. 1993

JAVITT, D. C. ; DONESHKA, P. ; ZYLBERMAN, I. ; RITTER, W. ; VAUGHAN JR., H. G.: Impairment of early cortical processing in schizophrenia: An event-related potential confirmation study. In: *Biological Psychiatry* 33 (1993), Nr. 7, S. 513–519

Katayama und Polich 1999

KATAYAMA, J. ; POLICH, J.: Auditory and visual P300 topography from a 3 stimulus paradigm. In: *Clinical Neurophysiology* 110 (1999), Nr. 3, S. 463–468

King und Calvert 2001

KING, A. J. ; CALVERT, G. A.: Multisensory integration: Perceptual grouping by eye and ear. In: *Current Biology* 11 (2001), Nr. 8, S. R322–R325

König et al. 2007

KÖNIG, R. ; SIELUZYCKI, C. ; DURKA, P. J.: Tiny signals from the human brain: Acquisition and processing of biomagnetic fields in magnetoencephalography. In: *Journal of Low Temperature Physics* 146 (2007), Nr. 5–6, S. 697–718

Kropotov et al. 1995

KROPOTOV, J. D. ; NÄÄTÄNEN, R. ; SEVOSTIANOV, A.V. ; REINIKAINEN, K. ; KROPOTOVA, O.V.: Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. In: *Psychophysiology* 32 (1995), Nr. 4, S. 418–422

Kubovy und van Valkenburg 2001

KUBOVY, M. ; VAN VALKENBURG, D.: Auditory and visual objects. In: *Cognition* 80 (2001), Nr. 1–2, S. 97–126

Levänen et al. 1996

LEVÄNEN, S. ; HARI, R. ; MCEVOY, L. ; SAMS, M.: Deviant auditory stimuli activate human auditory left and right auditory cortex differently. In: *Cerebral Cortex* 6 (1996), Nr. 2, S. 288–296

Maekawa et al. 2005

MAEKAWA, T. ; GOTO, Y. ; KINUKAWA, N. ; TANIKAWA, T. ; KANBA, S. ; TOBI-MATSU, S.: Functional characterization of mismatch negativity to a visual stimulus. In: *Clinical Neurophysiology* 116 (2005), Nr. 10, S. 2392–2402

May und Tiitinen 2004

MAY, P. J. C. ; TIITINEN, H.: Auditory scene analysis and sensory memory: the

role of the auditory N100m. In: *Neurology and Clinical Neurophysiology* 2004 (2004), Nr. 19, S. 1–5

McGurk und McDonald 1976

MCGURK, H. ; McDONALD, J.: Hearing lips and seeing voices. In: *Nature* 264 (1976), Nr. 5588, S. 746–748

Meredith et al. 1987

MEREDITH, M. A. ; NEMITZ, J. W. ; STEIN, B. E.: Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. In: *Journal of Neuroscience* 7 (1987), Nr. 10, S. 3215–3229

Miller 1947

MILLER, G. A.: The masking of speech. In: *Psychological Bulletin* 44 (1947), S. 105–129

Miller und Heise 1950

MILLER, G. A. ; HEISE, G. A.: The trill threshold. In: *Journal of the Acoustical Society of America* 22 (1950), S. 637–638

Mitchell und Maslin 2007

MITCHELL, T. V. ; MASLIN, M. T.: How vision matters for individuals with hearing loss. In: *International Journal of Audiology* 46 (2007), S. 500–511

Molholm et al. 2002

MOLHOLM, S. ; RITTER, W. ; MURRAY, M. M. ; JAVITT, D. C. ; SCHROEDER, C. E. ; FOXE, J. J.: Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. In: *Cognitive Brain Research* 14 (2002), Nr. 1, S. 115–128

Moore 1997

MOORE, B. C. J.: *An introduction to the psychology of hearing*. 4. Aufl. San Diego : Academic, 1997

Moore 1998

MOORE, B. C. J.: *Cochlear hearing loss*. London : Whurr, 1998

Moore und Gockel 2002

MOORE, B. C. J. ; GOCKEL, H.: Factors influencing sequential stream segregation. In: *Acta Acustica united with Acustica* 88 (2002), Nr. 3, S. 320–332

Moore 1983

MOORE, E. J.: *Bases of Auditory Brain-Stem Evoked Responses*. New York : Grune & Stratton, 1983

Morr et al. 2002

MORR, M. L. ; SHAFER, V. L. ; KREUZER, J. A. ; KURTZBERG, D.: Maturation of mismatch negativity in typically developing infants and preschool children. In: *Ear and Hearing* 23 (2002), Nr. 2, S. 118–136

Musacchia et al. 2005

MUSACCHIA, G. ; SAMS, M. ; NICOL, T. ; KRAUS, N.: Seeing speech affects acoustic information processing in the human brainstem. In: *Experimental Brain Research* 168 (2005), Nr. 1–2, S. 1–10

Nager et al. 2003

NAGER, W. ; TEDER-SÄLEJÄRVI, W. ; KUNZE, S. ; MÜNTE, T. F.: Preattentive evaluation of multiple perceptual streams in human audition. In: *Neuroreport* 14 (2003), Nr. 6, S. 871–874

Nelken 2004

NELKEN, I.: Processing of complex stimuli and natural scenes in the auditory cortex. In: *Current Opinion in Neurobiology* 14 (2004), Nr. 4, S. 474–480

Nicholls et al. 2004

NICHOLLS, M. E. R. ; SEARLE, D. A. ; BRADSHAW, J. L.: Read my lips. In: *Research Report* 15 (2004), Nr. 2, S. 138–141

Näätänen 1992

NÄÄTÄNEN, R.: *Attention and brain function*. Hillsdale, New Jersey : Lawrence Erlbaum Assoc., 1992

Näätänen 2003

NÄÄTÄNEN, R.: Mismatch negativity: clinical research and possible applications. In: *International Journal of Psychophysiology* 48 (2003), Nr. 2, S. 179–188

Näätänen und Escera 2000

NÄÄTÄNEN, R. ; ESCERA, C.: Mismatch negativity: Clinical and other applications. In: *Audiology and Neuro-Otology* 5 (2000), Nr. 3–4, S. 105–110

Näätänen et al. 1978

NÄÄTÄNEN, R. ; GAILLARD, A. W. K. ; MÄNTYSALO, S.: Early selective-attention effect on evoked potential reinterpreted. In: *Acta Psychologica* 42 (1978), Nr. 4, S. 313–329

Näätänen et al. 1993

NÄÄTÄNEN, R. ; PAAVILAINEN, P. ; TIITINEN, H. ; JIANG, D. ; ALHO, K.: Attention and mismatch negativity. In: *Psychophysiology* 30 (1993), Nr. 5, S. 436–450

Näätänen und Picton 1987

NÄÄTÄNEN, R. ; PICTON, T.: The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. In: *Psychophysiology* 24 (1987), Nr. 4, S. 375–425

Näätänen et al. 2001

NÄÄTÄNEN, R. ; TERVANIEMI, M. ; SUSSMAN, E. S. ; PAAVILAINEN, P. ; WINKLER, I.: 'Primitive intelligence' in the auditory cortex. In: *Trends in Neurosciences* 24 (2001), Nr. 5, S. 283–288

Näätänen und Winkler 1999

NÄÄTÄNEN, R. ; WINKLER, I.: The concept of auditory stimulus representation in cognitive neuroscience. In: *Psychological Bulletin* 125 (1999), Nr. 6, S. 826–859

O'Leary und Rhodes 1984

O'LEARY, A. ; RHODES, G.: Cross-modal effects on visual and auditory object perception. In: *Perception and Psychophysics* 35 (1984), Nr. 6, S. 565–569

Opitz et al. 1999

OPITZ, B. ; MECKLINGER, A. ; VON CRAMON, D. Y. ; KRUGGEL, F.: Combining electrophysiological and hemodynamic measures of the auditory oddball. In: *Psychophysiology* 36 (1999), Nr. 1, S. 142–147

Opitz et al. 2002

OPITZ, B. ; RINNE, T. ; MECKLINGER, A. ; VON CRAMON, D. Y. ; SCHRÖGER, E.: Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. In: *Neuroimage* 15 (2002), Nr. 1, S. 167–174

Paavilainen et al. 2003

PAAVILAINEN, P. ; MIKKONEN, M. ; KILPELÄINEN, M. ; LEHTINEN, R. ; SAARELA,

M. ; TAPOLA, L.: Evidence for the different additivity of the temporal and frontal generators of mismatch negativity: A human auditory event-related potential study. In: *Neuroscience Letters* 349 (2003), Nr. 2, S. 79–82

Paavilainen et al. 1995

PAAVILAINEN, P. ; SAARINEN, J. ; TERVANIEMI, M. ; NÄÄTÄNEN, R.: Mismatch negativity to changes in abstract sound features during dichotic listening. In: *Journal of Psychophysiology* 9 (1995), Nr. 3, S. 243–249

Palmer 1999

PALMER, S. E.: *Vision science: Photons to phenomenology*. Cambridge, MA : MIT Press, 1999

Pantev et al. 1988

PANTEV, C. ; HOKE, M. ; LEHNERTZ, K. ; LÜTKENHÖNER, B. ; ANOGIANAKIS, G. ; WITTKOWSKI, W.: Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. In: *Electroencephalography and Clinical Neurophysiology* 69 (1988), Nr. 2, S. 160–170

Pantev und Lütkenhöner 2000

PANTEV, C. ; LÜTKENHÖNER, B.: Magnetoencephalographic studies of functional organization and plasticity of the human auditory cortex. In: *Journal of Clinical Neurophysiology* 17 (2000), Nr. 2, S. 130–142

Pardo und Sams 1993

PARDO, P.J. ; SAMS, M.: Human auditory cortex responses to rising versus falling glides. In: *Neuroscience Letters* 159 (1993), Nr. 1–2, S. 43–45

Paulesu et al. 1995

PAULESU, E. ; HARRISON, J. ; BARON COHEN, S. ; WATSON, J. D. G. ; GOLDSTEIN, L. ; HEATHER, J. ; FRACKOWIAK, R. S. J. ; FRITH, C. D.: The physiology of colored hearing. A PET activation study of color-word synaesthesia. In: *Brain* 118 (1995), Nr. 3, S. 661–676

Pazo-Alvarez et al. 2003

PAZO-ALVAREZ, P. ; CADAVEIRA, F. ; AMENEDO, E.: MMN in the visual modality: A review. In: *Biological Psychology* 63 (2003), Nr. 3, S. 199–330

Pettigrew et al. 2004

PETTIGREW, C. M. ; MURDOCH, B. E. ; PONTON, C. W. ; KEI, J. ; CHENERY, H. ;

ALKU, P.: Subtitled videos and Mismatch negativity (MMN) investigations of spoken word processing. In: *Journal of the American Academy of Audiology* 15 (2004), Nr. 7, S. 469–485

Picton et al. 2000

PICTON, T. W. ; ALAIN, C. ; OTTEN, L. ; RITTER, W. ; ACHIM, A.: Mismatch negativity: Different water in the same river. In: *Audiologie and Neuro-Otology* 5 (2000), Nr. 3–4, S. 111–139

Pressnitzer und Hupé 2006

PRESSNITZER, D. ; HUPÉ, J. M.: Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. In: *Current Biology* 16 (2006), Nr. 3, S. 1351–1357

Rahne et al. 2007a

RAHNE, T. ; BÖCKMANN, M. ; VON SPECHT, H. ; SUSSMAN, E. S.: Visual cues can modulate integration and segregation of objects in auditory scene analysis. In: *Brain Research* 1144 (2007), Nr. 1, S. 127–135

Rahne et al. 2007b

RAHNE, T. ; DEIKE, S. ; SELEZNEVA, E. ; BROSCHE, M. ; KÖNIG, R. ; SCHEICH, H. ; BÖCKMANN, M. ; BRECHMANN, A.: A multilevel and cross-modal approach towards neuronal mechanisms of auditory streaming. In: *Brain Research* (im Druck) (2007)

Rahne et al. 2005

RAHNE, T. ; VON SPECHT, H. ; BÖCKMANN, M.: Einschätzung der Diskriminationsleistung linear frequenzmodulierter Töne mittels Mismatch Negativity. In: *Zeitschrift für Audiologie* 44 (2005), Nr. 4, S. 186–195

Rall 1977

RALL, W.: Core conductor theory and cable properties of neurons. In: KANDEL, E. R. (Hrsg.): *Handbook of physiology, the nervous system*. Bethesda : American Physiological Society, 1977, S. 39–97

Rausch et al. 2000

RAUSCH, M. ; WIDDIG, W. ; EYSEL, U. T. ; PENNER, I. K. ; TEGENTHOFF, M.: Enhanced responsiveness of human extravisual areas to photic stimulation in patients with severely reduced vision. In: *Experimental Brain Research* 135 (2000), Nr. 1, S. 34–40

Remijn et al. 2004

REMIJN, G. B. ; ITO, H. ; NAKAJIMA, Y.: Audiovisual integration: An investigation of the 'streaming-bouncing' phenomenon. In: *Journal of Physiological Anthropology and Applied Human Science* 23 (2004), Nr. 6, S. 243–247

Rinne et al. 2006

RINNE, T. ; SARKKA, A. ; DEGERMAN, A. ; SCHRÖGER, E. ; ALHO, K.: Two separate mechanisms underlie auditory change detection and involuntary control of attention. In: *Brain Research* 1077 (2006), Nr. 1, S. 123–134

Ritter et al. 2000

RITTER, W. ; SUSSMAN, E. S. ; MOLHOLM, S.: Evidence that the mismatch negativity system works on the basis of objects. In: *Neuroreport* 11 (2000), Nr. 1, S. 61–63

Robinson 1989

ROBINSON, S. E.: Environmental noise cancellation for biomagnetic measurements. In: WILLIAMSON, S. J. (Hrsg.) ; HOKE, M. (Hrsg.) ; STROINK, G. (Hrsg.): *Advances in biomagnetism*. New York : Plenum, 1989, S. 721–724

Rogers und Bregman 1993

ROGERS, W. L. ; BREGMAN, A. S.: An experimental evaluation of three theories of stream segregation. In: *Perception and Psychophysics* 53 (1993), Nr. 2, S. 179–189

Rosburg 2003

ROSBURG, T.: Left hemispheric dipole locations of the neuromagnetic mismatch negativity to frequency, intensity and duration deviants. In: *Cognitive Brain Research* 16 (2003), Nr. 1, S. 83–90

Rose und Moore 1997

ROSE, M. M. ; MOORE, B. C. J.: Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. In: *Journal of the Acoustical Society of America* 102 (1997), Nr. 3, S. 1768–1778

Rose und Moore 2005

ROSE, M. M. ; MOORE, B. C. J.: The relationship between stream segregation and frequency discrimination in normally hearing and hearing-impaired subjects. In: *Hearing Research* 204 (2005), Nr. 1–2, S. 16–28

Rubinstein 2004

RUBINSTEIN, J. T.: An introduction to the biophysics of evoked compound action potential. In: *International Journal of Audiology* 43 (2004), Nr. Suppl. 1, S. 3–9

Saarinen et al. 1992

SAARINEN, J. ; PAAVILAINEN, P. ; SCHRÖGER, E. ; TERVANIEMI, M. ; NÄÄTÄNEN, R.: Representation of abstract attributes of auditory stimuli in the human brain. In: *Neuroreport* 3 (1992), Nr. 12, S. 1149–1151

Saint-Amour et al. 2007

SAINT-AMOUR, D. ; DE SANCTIS, P. ; MOLHOLM, S. ; RITTER, W. ; FOXE, J. J.: Seeing voices: High-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. In: *Neuropsychologia* 45 (2007), Nr. 3, S. 587–597

Sakowitz et al. 2005

SAKOWITZ, O. W. ; QUIROGA, R. Q. ; SCHURMANN, M. ; BASAR, E.: Spatio-temporal frequency characteristics of intersensory components in audiovisually evoked potentials. In: *Cognitive Brain Research* 23 (2005), Nr. 2-3, S. 316–326

Salisbury et al. 2001

SALISBURY, D. F. ; RUTHERFORD, B. ; SHENTON, M. E. ; MCCARLEY, R. W.: Button-pressing affects P300 amplitude and scalp topography. In: *Clinical Neurophysiology* 112 (2001), Nr. 9, S. 1676–1684

Sams et al. 1991

SAMS, M. ; AULANKO, R. ; HÄMÄLÄINEN, M. ; HARI, R. ; LOUNASMAA, O. V. ; LU, S. T. ; SIMOLA, J.: Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex. In: *Neuroscience Letters* 127 (1991), Nr. 1, S. 141–145

Sams und Näätänen 1991

SAMS, M. ; NÄÄTÄNEN, R.: Neuromagnetic responses of the human auditory cortex to short frequency glides. In: *Neuroscience Letters* 121 (1991), Nr. 1–2, S. 43–46

Sams et al. 1985

SAMS, M. ; PAAVILAINEN, P. ; ALHO, K. ; NÄÄTÄNEN, R.: Auditory frequency discrimination and event-related potentials. In: *Electroencephalography and Clinical Neurophysiology* 62 (1985), Nr. 6, S. 437–448

Sarvas 1987

SARVAS, J.: Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. In: *Physics in Medicine and Biology* 32 (1987), Nr. 1, S. 11–22

Scherg 1990

SCHERG, M.: Fundamentals of dipole source potential analysis. In: GRANDORI, F. (Hrsg.) ; HOKE, M. (Hrsg.) ; ROMANI, G. L. (Hrsg.): *Evoked magnetic fields* Bd. 6. Basel : Karger, 1990, S. 40–69

Scherg et al. 1989

SCHERG, M. ; VAJSAR, J. ; PICTON, T. W.: A source analysis of the late human auditory evoked potentials. In: *Journal of Cognitive Neuroscience* 1 (1989), Nr. 4, S. 336–355

Schneider und Shiffrin 1977

SCHNEIDER, W. ; SHIFFRIN, R. M.: Controlled and automatic human information processing: I. Detection, search and attention. In: *Psychological Review* 84 (1977), S. 1–66

Schröger 1997

SCHRÖGER, E.: On the detection of auditory deviations: A pre-attentive activation model. In: *Psychophysiology* 34 (1997), Nr. 3, S. 245–257

Schröger 2004

SCHRÖGER, E.: Die Mismatch-Negativierung als Werkzeug in der Hörforschung. In: *Psychophysik und Audiologie*. Leipzig : Leipziger Universitätsverlag, 2004, S. 49–64

Schröger 2005

SCHRÖGER, E.: The mismatch negativity as a tool to study auditory processing. In: *Acta Acustica united with Acustica* 91 (2005), Nr. 3, S. 490–501

Schröger et al. 1994

SCHRÖGER, E. ; PAAVILAINEN, P. ; NÄÄTÄNEN, R.: Mismatch negativity to changes in a continuous tone with regularly varying frequencies. In: *Electroencephalography and Clinical Neurophysiology* 92 (1994), Nr. 2, S. 140–147

Schröger und Widmann 1998

SCHRÖGER, E. ; WIDMANN, A.: Speeded responses to audiovisual signal changes result from bimodal integration. In: *Psychophysiology* 35 (1998), Nr. 6, S. 755–759

Shams et al. 2000

SHAMS, L. ; KAMITANI, Y. ; SHIMOJO, S.: Illusions: What you see is what you hear. In: *Nature* 408 (2000), Nr. 6814, S. 788

Shepherd 1993

SHEPHERD, G.M.: *Neurobiologie*. 2. Aufl. Berlin : Springer, 1993

Smith et al. 1982

SMITH, J. ; HAUSFELD, S. ; POWER, R. P. ; GORTA, A.: Ambiguous musical figures and auditory streaming. In: *Perception and Psychophysics* 32 (1982), Nr. 5, S. 454–464

Snyder et al. 2006

SNYDER, J. S. ; ALAIN, C. ; PICTON, T. W.: Effects of attention on neuroelectric correlates of auditory stream segregation. In: *Journal of Cognitive Neuroscience* 18 (2006), Nr. 1, S. 1–13

Soto-Faraco et al. 2004

SOTO-FARACO, S. ; NAVARRA, J. ; ALSIUS, A.: Assessing automaticity in audiovisual speech integration: Evidence from the speeded classification task. In: *Cognition* 92 (2004), Nr. B, S. 13–23

Stein 1998

STEIN, B. E.: Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. In: *Experimental Brain Research* 123 (1998), Nr. 1–2, S. 124–135

Stein und Meredith 1993

STEIN, B. E. ; MEREDITH, M. A.: *Merging of the senses*. Cambridge : MIT Press, 1993

Stevens et al. 2005

STEVENS, M. C. ; CALHOUN, V. D. ; KIEHL, K. A.: Hemispheric differences in hemodynamics elicited by auditory oddball stimuli. In: *Neuroimage* 26 (2005), Nr. 3, S. 782–792

Suchan et al. 2006

SUCHAN, B. ; LINNEWERTH, B. ; KOSTER, O. ; DAUM, I. ; SCHMID, G.: Cross-modal processing in auditory and visual working memory. In: *Neuroimage* 29 (2006), Nr. 3, S. 853–858

Sussman 2005

SUSSMAN, E. S.: Integration and segregation in auditory scene analysis. In: *Journal of the Acoustical Society of America* 117 (2005), Nr. 3, S. 1285–1298

Sussman et al. 2005

SUSSMAN, E. S. ; BREGMAN, A. S. ; WANG, W.J. ; KHAN, F.J.: Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. In: *Cognitive, Affective, and Behavioral Neuroscience* 5 (2005), Nr. 1, S. 93–110

Sussman et al. 2001

SUSSMAN, E. S. ; CEPONIENE, R. ; SHESTAKOVA, A. ; NÄÄTÄNEN, R. ; WINKLER, I.: Auditory stream segregation processes operate similarly in school-aged children and adults. In: *Hearing Research* 153 (2001), Nr. 1–2, S. 108–114

Sussman et al. 1998a

SUSSMAN, E. S. ; RITTER, W. ; VAUGHAN JR., H. G.: Attention affects the organisation of auditory input associated with the mismatch negativity system. In: *Brain Research* 789 (1998), Nr. 1, S. 130–138

Sussman et al. 1998b

SUSSMAN, E. S. ; RITTER, W. ; VAUGHAN JR., H. G.: Predictability of stimulus deviance and the mismatch negativity. In: *Neuroreport* 9 (1998), Nr. 18, S. 4167–4170

Sussman et al. 1999

SUSSMAN, E. S. ; RITTER, W. ; VAUGHAN JR., H. G.: An investigation of the auditory streaming effect using event-related brain potentials. In: *Psychophysiology* 36 (1999), Nr. 1, S. 22–34

Sussman et al. 2002

SUSSMAN, E. S. ; WINKLER, I. ; HUOTILAINEN, M. ; RITTER, W. ; NÄÄTÄNEN, R.: Top-down effects can modify the initially stimulus-driven auditory organization. In: *Cognitive Brain Research* 13 (2002), Nr. 3, S. 393–405

Sussman et al. 2007

SUSSMAN, E. S. ; WONG, R. ; HORVÁTH, J. ; WINKLER, I. ; WANG, W.: The development of the perceptual organization of sound by frequency separation in 5–11 year-old children. In: *Hearing Research* 225 (2007), Nr. 1–2, S. 117–127

Swets 1964

SWETS, J. A.: *Signal detection and recognition by human observers*. New York : Wiley, 1964

Tales et al. 1999

TALES, A. ; NEWTON, P. ; TROSCIANKO, T. ; BUTLER, S.: Mismatch negativity in the visual modality. In: *Neuroreport* 10 (1999), Nr. 16, S. 3363–3367

Tervaniemi et al. 1999

TERVANIEMI, M. ; LEHTOKOSKI, A. ; VIRTANEN, J. ; ILMONIEMI, R. J. ; NÄÄTÄNEN, R.: Test-retest reliability of mismatch negativity for duration, frequency and intensity changes. In: *Clinical Neurophysiology* 110 (1999), Nr. 8, S. 1388–1397

Tervaniemi et al. 2000

TERVANIEMI, M. ; MEDVEDEV, S. V. ; ALHO, K. ; PAKHOMOV, S. V. ; ROUDAS, M. S. ; VAN ZUIJEN, T. L. ; NÄÄTÄNEN, R.: Lateralized automatic auditory processing of phonetic versus musical information: A PET study. In: *Human Brain Mapping* 10 (2000), Nr. 2, S. 74–79

Ulanovsky et al. 2003

ULANOVSKY, N. ; LAS, L. ; NELKEN, I.: Processing of low-probability sounds by cortical neurons. In: *Nature Neuroscience* 6 (2003), Nr. 4, S. 391–398

van Noorden 1975

VAN NOORDEN, L. P. A. S.: *Temporal coherence in the perception of tone sequences*, Eindhoven University of Technology, Dissertation, 1975

van Noorden 1977

VAN NOORDEN, L. P. A. S.: Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. In: *Journal of the Acoustical Society of America* 61 (1977), Nr. 4, S. 1041–1045

Vatakis und Spence 2006

VATAKIS, A. ; SPENCE, C.: Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. In: *Neuroscience Letters* 393 (2006), Nr. 1, S. 40–44

Vliegen et al. 1999

VLIEGEN, J. ; MOORE, B. C. J. ; OXENHAM, A. J.: The role of spectral and period-

icity cues in auditory stream segregation, measured using a temporal discrimination task. In: *Journal of the Acoustical Society of America* 106 (1999), Nr. 2, S. 938–945

Vrba und Robinson 2001

VRBA, J. ; ROBINSON, S. E.: Signal processing in magnetoencephalography. In: *Methods* 25 (2001), Nr. 2, S. 249–271

Vroomen et al. 2001

VROOMEN, J. ; BERTELSON, P. ; DE GELDER, B.: The ventriloquist effect does not depend on the direction of automatic visual attention. In: *Perception and Psychophysics* 63 (2001), Nr. 4, S. 651–659

Wallace und Stein 1994

WALLACE, M. T. ; STEIN, B. E.: Cross-modal synthesis in the midbrain depends on input from cortex. In: *Journal of Neurophysiology* 71 (1994), Nr. 1, S. 429–432

Warren 1968

WARREN, R. M.: Relation of verbal transformations to other perceptual phenomena. In: *IEEE/NPL Conference on Pattern Recognition*. London : Institution of Electrical Engineers, 1968 (Conference publication No. 42), S. 1–8

Watkins et al. 2006

WATKINS, S. ; S. TANAKA, L. S. and. ; HAYNES, J. D. ; REES, G.: Sound alters activity in human V1 in association with illusory visual perception. In: *Neuroimage* 31 (2006), Nr. 3, S. 1247–1256

Whittingstall et al. 2004

WHITTINGSTALL, K. ; STROINK, G. ; DICK, B.: Dipole localization accuracy using grand-average EEG data sets. In: *Clinical Neurophysiology* 115 (2004), Nr. 9, S. 2108–2112

Wikswow Jr. 1995

WIKSWO JR., J. P.: SQUID magnetometers for biomagnetism and nondestructive testing: Important questions and initial answers. In: *IEEE Transactions on Applied Superconductivity* 6 (1995), Nr. 2, S. 47–120

Winkler et al. 2004

WINKLER, I. ; TEDER-SÄLEJÄRVI, W. ; HORVÁTH, J. ; NÄÄTÄNEN, R. ; SUSSMAN, E. S.: Human auditory cortex tracks task-irrelevant sound sources. In: *Neuroreport* 14 (2004), Nr. 16, S. 2053–2056

Woldorff et al. 1991

WOLDORFF, M. G. ; HACKLEY, S. A. ; HILLYARD, S. A.: The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. In: *Psychophysiology* 28 (1991), Nr. 1, S. 30–42

Yabe et al. 2001

YABE, H. ; WINKLER, I. ; CZIGLER, I. ; KOYAMA, S. ; KAKIGI, R. ; SUTOH, T. ; HIRUMA, T. ; KANEKO, S.: Organizing sound sequences in the human brain: The interplay of auditory streaming and temporal integration. In: *Brain Research* 897 (2001), Nr. 1–2, S. 222–227

Yumoto et al. 2005

YUMOTO, M. ; UNO, A. ; ITOH, K. ; SAITOH, O. ; KANEKO, Y. ; YATOMI, Y. ; KAGA, K.: Audiovisual phonological mismatch produces early negativity in auditory cortex. In: *Neuroreport* 16 (2005), Nr. 8, S. 803–806

Zatorre et al. 2002

ZATORRE, R. J. ; BELIN, P. ; PENHUNE, V. B.: Structure and function of auditory cortex: Music and speech. In: *Trends in Cognitive Sciences* 6 (2002), Nr. 1, S. 37–46

Zimbardo 1995

ZIMBARDO, P. G.: *Psychologie*. 6. Aufl. Augsburg : Weltbild Verlag GmbH, 1995

Zwislocki 1960

ZWISLOCKI, J. J.: Theory of temporal auditory summation. In: *Journal of the Acoustical Society of America* 32 (1960), S. 1046–1059

Lebenslauf

Personalien

Name: Dipl.-Phys. Torsten Rahne
geboren am 2. August 1979 in Lutherstadt Eisleben
Nationalität: deutsch
Familienstand: verheiratet

Wissenschaftlicher Bildungsgang

1991 – 1998 Martin-Luther-Gymnasium in Lutherstadt Eisleben
08.07.1998 Abschluss: Allgemeine Hochschulreife
1998 – 1999 Zivildienst als Krankenpfleger im Klinikum Mansfelder Land in Lutherstadt Eisleben
1999 – 2004 Studium der Physik an der Otto-von-Guericke-Universität Magdeburg, Schwerpunkte: Biophysik und medizinische Physik
22.09.2004 Abschluss als Diplom-Physiker (Diplomarbeit: „Ereigniskorrelierte Potentiale bei Stimulation mit linear frequenzmodulierten Tönen im Vergleich mit der Unterscheidungsleistung“)
2001, 2002 Wissenschaftliche Hilfskraft am Institut für Experimentelle Physik, Abteilung Festkörperphysik, der Otto-von-Guericke-Universität Magdeburg
2003 – 2004 Wissenschaftliche Hilfskraft in der Abteilung für Experimentelle Audiologie und Medizinische Physik an der Medizinischen Fakultät der Otto-von-Guericke-Universität Magdeburg
seit 2004 Promotionsstudium an der Otto-von-Guericke-Universität Magdeburg
seit 2004 Wissenschaftlicher Mitarbeiter der Abteilung für Experimentelle Audiologie und Medizinische Physik an der Medizinischen Fakultät der Otto-von-Guericke-Universität Magdeburg
2005, 2007 Forschungsaufenthalte am Albert-Einstein-College of Medicine, Department of Neuroscience in New York, USA