# Selection of Information in Auditory Virtual Reality

## Dissertation

zur Erlangung des akademischen Grades

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

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

von: Dipl.-Psych. Dörte Spring (geb. Kuhlicke)

geb. am: 10.07.1979 in: Belzig

Gutachter: Prof. Dr. Thomas F. Münte
Prof. Dr. Bernadette Jansma

eingereicht am: 29.01.2007
verteidigt am: 01.03.2007

*To my parents and parents in law*

*Especially in memory of Kathryn H. Spring,*

*For their continued support,*

*Which has made it all possible.*

*And to all my teachers, present and past,*

*For instilling the proper knowledge in me,*

*Which has enabled me to pursue this path.*

# Preamble

This dissertation is written in partial fulfillment of requirements for the degree of doctor rerum naturalium (Dr. rer. nat.) at the Otto-von-Guericke University of Magdeburg, faculty of natural sciences. The bulk of the research presented in this doctoral dissertation was conducted in the laboratory of Prof. Thomas F. Münte, institute of psychology II, department of neuropsychology, at the University of Magdeburg

When I started my work on cortical activity during selective auditory attention, already several dichotic listening tasks with rather single stimuli (e.g., tones) than whole simultaneous speeches were conducted to investigate auditory selection of attention. The well known "Cocktail-Party phenomenon" was pulled up to address the attentional selection within the auditory modality. Some of those studies already tried to explore this phenomenon in a more and more natural setting by using spatial stimulus presentation (ITD, ILD) rather than simple right-left presentations, but there was not a single experiment using individual HRTFs and continuous as well simultaneous speeches like at a real "Cocktail-Party". This fact made my research even more interesting and realistic. Nevertheless, the phenomenon of "Attention" in general remains incompletely explained regarding its origin, functionality as well as behavioral and electrophysiological effects, and therefore still needs further investigations.

During the last years of research, I had the opportunity to work with many experts in the field of EEG measurements, language, attention at the University of Magdeburg, Germany, mainly who taught me experimental and analysis techniques which became the bread and butter of my work. I also had the exceptional opportunity to work in the Cognitive laboratory at the New Mexico State University (NMSU) in Las Cruces, NM during the last months and actual writing my dissertation. Both work groups (Magdeburg, and Las Cruces) gave me the support that I needed and influenced my work decisively. The research was funded by a grant from the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation).

It gives me great pleasure to take this chance to thank some of the people without whom the completion of this dissertation would not have been possible. First, I'm grateful to Prof. Dr. Thomas F. Münte, my doctor father, who gave me the opportunity to work on this brilliant project. He also was a great source of knowledge and ideas, and a good mentor. He supported me in gaining my own experiences in this research field.

Daniel Wiswede and Jürn Möller were always available in any programming or analysis questions as well. I learned very much from them, so that I felt more and more assured with what I was doing. Thanks also to Jascha Rüsseler for his support and advises. He helped me to get a deeper understanding of Complex Event Related Potential analysis. Whenever I had some question, he was always there. We had a lot good scientific discussions. He was furthermore my teacher in scientific writing. His editorial assistance and comments were auxiliary for this doctoral dissertation. Monique Lamers and Arie van der Lugt were also very essential for the final corrected proof of this dissertation. Without their help, reading this present work would have caused a large number of MMNs and P3s.

During my stay at NMSU (Las Cruces, NM), it was Dr. James Kroger who gave me support and a lot good guideposts regarding the term “attention” in general. Not only his class that I was allowed to audit, but also his further critical discussions about my research were instructive. It certainly has changed my scientific thinking. Further on, he also helped me to improve the readability of the final version of this dissertation.

However, the work on this project of auditory selective attention within a real cocktail-party setting would have never been possible without the help of my husband Ryan Spring, who spent several nights with me in front of the computer to create stimuli and to convolve them finally with the HRTFs for individualized and spatial stimuli. He also was my moral support, my motivation and strength during the whole dissertation period. It was and still is his love that helped to concentrate and to continue my scientific work. I want to thank him warmly for his patience, understanding and constant confidence in my scientific work, especially as I spent most of the evenings and weekends in front of my computer.

During the data acquisition, I was really grateful to have Nadine Strien, Peggy Tausche, and Dana Heinze on my side. Only with their assistance, it was possible to organize and record all subjects five times in total in a relative short time. By the end, they were even more than just assistants. They became good friends who bolstered me to keep up especially in frustrating situations during the investigation. Without their help and encouragement, I would maybe still work on this series of experiments and their analyses not to mention to think about writing the actual dissertation.

Many thanks also to my parents as well as parents in law for their support und motivation. They were the reason why I never gave up to learn, to improve and to quest ahead. I would also like to thank my friends Daniel Lenz, Sabine Schneider, and Sandy Harth for encouragement, their friendship, their personal entity and recreational opportunities whenever I needed change and distraction. We were all in the same boat regarding working on our dissertations. Thus, they knew best how it feels like to be in that situation and therefore could help most with their understanding. Finally, it was Emily Chaffin who welcomed me with open arms and helped me integrating as I arrived in Las Cruces, NM. She was my best friend at work and private since then. All this made it very easy for me to continue the data analysis and the actual writing of my dissertation once I was settled there.

Last but not least, I have to thank the weatherman, who managed to provide the hottest spring and summer that the desert ever had. This helped me to stay home in the nice air conditioned apartment and to concentrate on writing instead of going out for bike trips along the southern edge of the US Rocky Mountains for further inspiration on the weekends.

Dörte Spring

Dörte Spring

Magdeburg, January 2007

# English summary

Within four similar EEG experiments, the ability to follow the content of one narrator in a realistic "Cocktail-Party" situation was addressed. The aim of these studies was to investigate selective auditory perception and the restriction of attentional focus with the underlying electrophysiology (ERPs; event-related potentials) in a setting with 2 or 4 speakers. Listeners' attention was assumed to be directed to a specific set of sounds, sound sources and not to others (cocktail party problem).

A special technique for a virtual spatial presentation of auditory stimuli was used to avoid a simple right-left presentation. Individual HRTFs (head related transfer function) from different angles were measured first. These HRTFs of a certain spatial angle were convolved with the actual stimulus material (stories, phonemes and white noise). The subject's task was always the same from experiment to experiment: to attend one of the two or four concurrent and simultaneously presented stories.

The experimental difference involved the superimposed probe stimuli (task irrelevant) which were either phonemes or white noise bursts. These probes, especially the phonemes, varied in fundamental frequency and/or spatial location in the four experiments. By doing so, the question about the nature, characteristics and precision of the attentional focus should be investigated. The working hypothesis was that directed attention to one ear would cause an attenuation of stimulus processing on the other ear reflected in a Nd/PN (negative difference/processing negativity) which was experimentally confirmed. There was an additional Pd component in two of four experiments indicating an initial selection in novel and unfamiliar situations. Furthermore, it was observed that the focus of attention is absolutely restricted to the to-be-attended auditory object only. Not only large deviants in frequency and spatial location, but even slight differences elicited a N1, PN or MMN (mismatch negativity) and P3 or RON. These were taken as a sign of attenuated stimulus processing and/or deviance detection as well as an attentional shift. Furthermore, all these effects show that there are two stages in auditory selection: a later attentional selection process (Nd/PN, P3a, RON) besides an early stimulus driven selection (Pd, N1, MMN).

The spatial and frequency effects suggest that these deviants were not included into the actual focus of attention anymore and are processed as unattended stimuli with a distracting potential. The spectral content seems to also play an important role in auditory selective attention as well. Only stimuli with exactly the same frequency spectrum as the to-be-attended auditory object (one story respectively in those four studies) fell into the attentional focus, others with less or even more frequencies in the spectrum are thoroughly excluded.

No definitive differences were found in stimulus processing within a more complex "Cocktail-Party" situation (four speakers) compared to two-speaker settings. Similar ERP components were found in both settings which suggest that the same attentional mechanism underlay the stimulus processing. Nevertheless, the four-speaker experiment compared with the two-speaker experiment revealed an apparent mirror effect for ERP results in the attended to the unattended hemispace. The only difference between both hemifields was seen in a slight shift in processing level for the unattended side, in parallel below the one for each analogous stimulus on the attended side. This result indicates that even unattended stimuli are distinguished and not treated as if they were the same stimulus. Perhaps, due to evolutionary reasons, it is highly important to be able still to differentiate between incoming information and thus, to be prepared for potentially dangerous objects.

The results also indicate that only attended stimuli are processed and elaborated deeply whereas any deviants or unattended stimuli experience an attenuated processing depending on the working memory capacity and the load of the primary task. Thus, the attentional focus is absolutely restricted to the to-be-attended object and its features only. Any changes in a feature would lead to an exclusion from the actual focus. This fact finally led to the novel "modification hypothesis" further explained in this dissertation.

Altogether, these experiments have demonstrated the validity of using HRTFs for an auditory spatial virtuality. Furthermore, it was shown that the attentional focus is highly restricted to all features of the to-be-attended stimulus. The combination of, for example, fundamental frequency, spectrum and spatial location is attended, and not just one of these features. Finally, two selective stages (early and late) decide which stimulus is included in or excluded from attentional focus.

## Deutsche Zusammenfassung

In einer Reihe von vier, aber jeweils leicht abgewandelten EEG Experimenten wurde die Fähigkeit untersucht, den Inhalt einer Erzählung innerhalb einer realistischen „Cocktial-Party“ Situation zu verfolgen. Das Ziel dieser Studien war es, die selektive auditorische Wahrnehmung, die Eigenschaften des Aufmerksamkeitsfokus und die zugrunde liegende Elektrophysiologie (EKP; ereignis-korrelierte Potentiale) in einer Zwei- oder Vier-Sprechersituation zu beleuchten. Es wurde dabei angenommen, daß die Aufmerksamkeit auf einen spezifischen Ton, Geräusch, oder Hörquelle gerichtet ist und nicht zu anderen (Cocktail-Party Problem), die ebenfalls zum selben Zeitpunkt präsent sind. Eine Selektion bzw. Filterung von Reizen wird somit vorausgesetzt.

Eine spezielle Methode zur räumlich virtuellen Darbietung auditorischer Reize anstelle einer einfachen rechts-links Stimulation wurde angewendet. Individuelle HRTFs (head related transfer function: kopfbezogene Übertragungsfunktion) verschiedenster Raumwinkel sind dafür von jedem einzelnen Probanden der Studie gemessen worden. Diese HRTFs sind individuell verschieden, da sie von der Anatomie zum Beispiel des Torsos und der äußeren Gehörganges (Ohrmuschel) entscheidend beeinflusst werden. Je nach Experiment wurden diese HRTFs eines bestimmten Raumwinkels anschließend mit dem eigentlichen Reizmaterial (Erzählungen, Phoneme und weißes Rauschen) gefaltet. Somit wurde das Reizmaterial von diesem gewählten Winkel im Raum kommend wahrgenommen. Die zu bearbeitende Aufgabe der Probanden während aller Experimente bestand immer darin, auf eine der zwei oder vier gleichzeitig dargebotenen Erzählungen zu achten. Um sicher zu gehen, daß die Versuchspersonen tatsächlich auf den vorgebenen Sprachkanal geachtet und den oder die jeweils anderen ignoriert haben, wurde im Anschluß ein Interview durchgeführt. Offene Fragen zur gehörten Erzählung sollten dabei beantwortet werden.

Die Experimente unterschieden sich nur in der Art der Probe-Reize (aufgaben-irrelevant), die jeweils den Erzählungen überlagert waren: Phoneme oder weißes Rauschen. Die Phoneme variierten darüber hinaus in der Grundfrequenz und/oder im Darbietungsort. Mit diesem Vorgehen soll die Frage nach der Natur, den Eigenschaften und der Präzision des Aufmerksamkeitsfokus geklärt werden. Wonach oder wie genaut wird selektiert? Eine allgemeine Hypothese war, daß die auf ein Ohr gerichtete Aufmerksamkeit für eine schwächere Verarbeitung für Reize der nicht beachteten Seite sorgen würde. Diese bestätigte sich durch die Beobachtung eines Aufmerksamkeitseffektes in Form einer Nd (negativen Differenz) oder PN (Prozessnegativität) im EKP. In zwei von vier Experimenten wurde zusätzlich eine Pd (positive Differenz) gefunden, die für eine erste Selektion in neuen und ungwohnten Situationen sprechen könnte. Weiterhin wurde festgestellt, daß der Aufmerksamkeitsfokus ausschließlich auf das zu beachtende auditorische Objekt (in diesem Falle eine der Erzählungen) beschränkt ist. Nicht nur deutliche Abweichungen in der Grundfrequenz oder im räumlichen Ort, sondern sogar kleinste Veränderungen der Reizeigenschaften sind in der Lage eine N1, PN oder im Falle des Oddball-Experiments eine MMN (Mismatch-Negativität) und P3 auszulösen. Diese EKP-Komponenten können als Beweis für eine geringfügigere Reizverarbeitung, Detektion einer Abweichung sowie einer Aufmerksamkeitsverlagerung angesehen werden. Diese Effekte haben außerdem gezeigt, daß es zwei Selektionsstufen gibt: einen späten aufmerksamkeitsbezogenen Selektionsprozess (Nd/PN, P3a, RON) neben einer reizgeleiteten Selektion (Pd, N1, MMN).

Dem räumlichen und Frequenzeffekt zufolge ist anzunehmen, daß Abweichungen jeglicher Art nicht mehr im Fokus der Aufmerksamkeit sind, sondern eher wie nicht beachtete Reize verarbeitet werden, die jedoch ein Ablenkungspotential durch jene Abweichungen haben. Jedoch scheint die Verarbeitung im Fokus anliegenden Bereich graduell abzunehmen. Die Abweichungseffekte, insbesondere die Frequenzeffekte, nehmen mit dem Grad der Abweichung zu. Größere Abweichungen weisen ausgeprägtere EKP-Komponenten (MMN, P3a) auf. Der spektrale Gehalte scheint ebenso eine sehr bedeutende Rolle in der auditiven

selektiven Aufmerksamkeit zu spielen. Ausschließlich Reize mit genau dem gleichen Frequenzspektrum wie das zu beachtende auditorische Objekt fallen in den Aufmerksamkeitsfokus. Andere Reize mit weniger oder sogar mehr Frequenzen im Spektrum werden uneingeschränkt vom eigentlichen Fokus ausgeschlossen.

Kein Unterschied zwischen einer komplexeren „Cocktail-Party“ Situation (vier simultane Sprechen) und einer Zwei-Sprecher Situation bezüglich der Reizverarbeitung wurde gefunden. Ähnliche EKP Komponenten wurden in beiden Szenarienarten beobachtet, was auf den gleichen der Reizverarbeitung zugrunde liegenden Aufmerksamkeitsmechanismus schließen läßt. Das vier-Sprecher Experiment im Vergleich zum zwei-Sprecher Arrangement hat einen scheinbaren Spiegeleffekt bezüglich EKP Ergebnisse des attendierten zum unattendierten Hemiraum offenbart. Der einzige Unterschied zwischen beiden akustischen Raumhälften bestand in einer parallelen Verlagerung des Verarbeitungsniveaus für die unbeachtete Seite unterhalb dessen der beachteten Hälfte und der jeweils konkurrierenden Positionen. Aufmerksamkeit allein scheint in der Lage zu sein, die Sensibilität für Reize in einem Hemifeld zu steigern; unabhängig von dem jeweils konkurrierenden Hemiraum. Dieses Resultat deutet an, daß sogar vernachlässigte Reize differenziert werden anstatt als ein und dasselbe behandelt und verarbeitet zu werden. Möglicherweise ist dies auf evolutionäre Aspekte zurückzuführen, wobei es sehr wichtig erscheint, jegliche einströmende Information separat zu behandeln, um dadurch auf potentiell gefährliche Objekte reaktionsbereit zu bleiben.

Die vorliegenden Ergebnisse zeigen generell, daß nur beachtete Reize tief und ausführlich verarbeitet werden, wohingegen abweichende und unbeachtete Reize eher nur oberflächliche verarbeitet werden in Abhängigkeit der Auslastung des Arbeitsgedächtnisses und der Beanspruchung der primären Aufgabe. Auch aus diesem Grund ist der Aufmerksamkeitsfokus wahrscheinlich auf das zu beachtende Objekt und seine Eigenschaften beschränkt. Jegliche Abwandlungen auch nur einer Eigenschaft jenes Objektes führen zum Ausschluß aus dem eigentlichen Aufmerksamkeitsfokus. Aus dieser Tatsache leitete sich schließlich die „Modifikationshypothese“ ab: Jede noch so kleine Abweichung vom eigentlich zu beachtenden auditorischen Objekt, sei es durch ein Fehlen oder einen Zusatz von Eigenschaften, führt unweigerlich zum Auschluss aus dem präzisen Aufmerksamkeitsfokus.

Zusammengefaßt haben alle Experimente gezeigt, daß der Gebrauch von HRTFs zur räumlichen Simulation sehr valide ist. Außerdem wurde demonstriert, daß der Aufmerksamkeitsfokus streng auf alle Merkmale des zu beachtenden Stimulus beschränkt ist. Es kommt auf die Kombination von Stimulusmerkmalen an und nicht nur auf eine bestimmte Eigenschaft, um einen Stimulus zu beachten. Bei der Selektion entscheiden dann zwei Prozessstufen (früh und spät), welche der Stimuli in den Aufmerksamkeitsfokus fallen und welche nicht.

## Contents

# Selection of Information in Auditory Virtual Reality

> *"Everyone knows what attention is. It is the taking possession by the mind in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others." (James, 1890)*

## Introduction/Theory

### Chapter 1

What is attention? It is generally seen as the cognitive process of selectively concentrating on one element while ignoring other elements in the environment. Unfortunately it is unclear how the brain produces the phenomena of attention; whether it is a single brain function and whether a single region of the brain is responsible for it. As such, there are numerous models explaining how attention works.

Attention is one of the most intensely studied topics within psychology and cognitive neuroscience. Some researchers may consider attention to be a very concrete process, similar to other cognitive processes (decision-making, memory, emotion, etc), because it is tied so closely to perception. To these researchers, it is a gateway to the rest of cognition. Nevertheless, attention is not completely understood yet.

Research on attention started in 1850 with William James, who was only able to study attention by introspection. About one hundred years later, Cherry, Broadbent and Treisman (Broadbent, 1958; Cherry, 1953; Treisman, 1960) used dichotic listening tasks for their attention research. Therein, two auditory messages were presented to both the left and right ears (one message to each ear). Participants were then asked to pay attention to one of the messages. By asking questions about the content of the unattended stream they saw how much information was filtered out to be able to pay attention on the other stream. This began the major debate between early-selection models and late-selection models of attention. In the 1960s and 1970s, Robert Wurtz recorded electrical signals from the brains of macaque monkeys who were trained to perform attention tasks (Goldberg & Wurtz, 1972; Wurtz & Goldberg, 1972). For the first time, it was shown that the superior colliculus was a direct neural correlate of this mental process. fMRI studies in the 1990s provided vivid images of the brain in attentive tasks. These results confirmed the earlier findings in psychophysical and primate behavioral literature. These were just the beginnings of the attention investigation. Since then, many other researchers have further studied the phenomena within both psychology and neuroscience.

Still, many questions are unanswered or not explained conclusively (e.g. that attention can be split, or that attention (preferred processing) is just “one side of the coin” whereas inhibition or attenuation represent the other side). However, there is no better psychological definition of attention than that from James (see entrance quote). Nevertheless, there are more and more neural approaches to define what attention is. Those neural based assumptions consider attention e.g. as an increase of a stimulus’ contrast or as an inhibition of irrelevant stimulus’ processing (Posner, 2004). The slow speed of progress spurred speculations about attention as a set of many separate processes without a common mechanism. This makes it even a greater challenge to disentangle the construct of attention. Perhaps someday, we can put together all these little pieces (research results) of the puzzle.

This dissertation will be mainly concerned with properties of attentional focus. Which mechanisms allow selection of attention/auditory information in complex surroundings? How is attentional focus structured? Which stimulus will be processed within the actual focus? Does a change of a single feature of a stimulus immediately effect an exclusion from the attentional focus? To answer these questions, as well as to validate a new approach of 3-D-stimulus presentation, four different electrophysiological (EEG; electroencephalogram; neurophysiologic measurement of electrical brain activity by recording from electrodes placed on the scalp) studies were conducted (chapter 2 and 3). The first experiment investigates the most important question of the method’s validity in comparison with classical findings regarding attention effects. By varying some stimulus features in the second experiment, the precision of the attentional focus is addressed. In the next study an EEG with higher spatial resolution is applied to specify the underlying source localization more in detail. Furthermore, the importance of the spectral content of auditory information for selective attention is inquired. Last but not least, the fourth experiment explores a more complex cocktail party situation, and whether the mechanisms are the same as in the simpler auditory settings. To lay the ground for these experiments, one has to know why selection is so important and which classical theories about selection processes already exist. This will be discussed in the next section.

## Filter theories

Imagine a busy and noisy day in the office with coworkers chatting, ringing phones, a running air conditioner, doors opening and closing, passing people, singing birds outside the window, and so on. This is an example of a normal every-day environment in which people are exposed to an enormous amount of simultaneous stimuli which would be impossible to process equally. Trying to process this amount of stimuli at the same time is impossible (and undesirable), because of limitations of our processing apparatus. Its restricted capacity makes a selection process necessary which reduces the incoming stream of stimuli and filters out all but the most important information in any given situation. This selective attention is a person’s ability to filter the information they are confronted with, and only pays attention to that which is useful or interesting to them.

Researchers thought about the filter process and how it could be managed. Their main ideas will be presented in the following selection of classic theories. These theories differ principally in terms of the exact place of the filter process also known as the bottleneck in attention processing.

## Broadbent's Theory

Broadbent's Filter Theory of Attention (Broadbent, 1958) forms the basis for most selective attention models, because it is the first detailed theory of focused auditory attention. His theory assumes a number of sensory channels within the sensory buffer which let simultaneous sound information pass through. Broadbent also thought that these "channels" may have some distinct neural representation in the brain. Physical characteristics (sound attributes such as pitch, loudness or spatial position) define which channel a stimulus uses. The postulation is that sound channels lead into a sensory buffer, where a particular channel is filtered based on the desired sound attributes. This early selection prevents overload of the capacity-limited mechanism beyond the filter. By this filtering, only the content of one input channel is transferred to the long-term memory store for later processing, as well as to any output mechanisms necessary to respond to the input channel. This way, overloading of capacity limited mechanisms succeeding the filter is prevented (see Figure 1).

This theory arose from the cocktail party situation in which a guest listens carefully to what someone is saying while ignoring other conversations in the room (e.g. the cocktail party problem, (Cherry, 1953)). In this scenario, the listeners must filter out the sounds of the other conversations to concentrate on the conversation they are listening to. Nevertheless, there are different conditions under which the filter switches to listen to a different channel. The switch is for example very easy and requires no conscious effort, when one’s own name is heard in the middle of a current conversation. The appearance of one’s own name determines the selection of which channel is filtered. Moray (1970) suggests that it takes approximately 0.25 seconds to switch between the channels.

The ability to switch indicates that Broadbent was incorrect in concluding that the unattended information is always rejected early in processing. He falsely assumed that the meaning of unattended information cannot be processed beyond this early stage, or that it cannot reach consciousness.

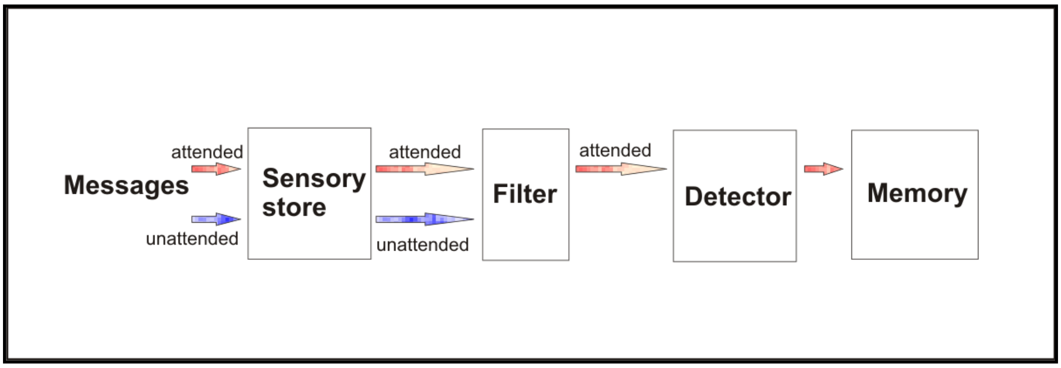


**Figure 1. A diagram of Broadbent's theoretical model of attention.**

## Treisman's Theory

Treisman (1960; 1964) noticed the problems of Broadbent's theory and proposed a hierarchy of early selective tests or filters instead of a single filter. In his model, there is one test for physical properties of words, and others for syllable patterns, grammar, and finally semantic meaning. Messages that fail various tests become attenuated to minimize interference with the selected message, but are not discarded until the final filter examining the semantic meaning of the messages. Thus, all the input selections bypass the short-term memory area of the brain. The early filters eliminate most, but not all, unattended sound channels. A subset of input channels (signals) is still allowed to enter the "dictionary units" (a series of nodes in the brain) for retrieving their semantic meaning (late selection). This network of dictionary units works as a pattern matcher. Similar-sounding stimuli finally trigger signals to the listener's output activity mechanism in the brain. If the patterns are as designated then an appropriate response to that stimulus is induced. Mismatching stimuli are not able to elicit a response because they are filtered out at this point. Hence, the real selection takes place at a later stage than Broadbent proposed. Processing of the unattended message is simply attenuated. Treisman suggested a more hierarchical selection process with processing of physical characteristics early in the hierarchy, and semantic processing at a later stage (Figure 2). According to this, the selective filter distinguishes between messages on the basis of their physical characteristics, such as location, intensity and pitch, whereas the 'dictionary' in Treisman's model allows for selection between messages on the basis of content. Although the awareness of it may be easily activated, certain information requires a very low threshold, (e.g. our name in the cocktail party example). The attenuation model proposes a decrease in the perceived loudness of an unattended message in such a way that this message is usually not loud enough to reach its threshold. Only messages with a very low threshold, such as one's own name, pass through. Exceptions are situations in which there is a general momentary decrease for all messages which allows even unattended information to pass. Nevertheless, semantic information will only be processed if there is sufficient processing capacity left. Otherwise, some later analyses are omitted for unattended stimuli.

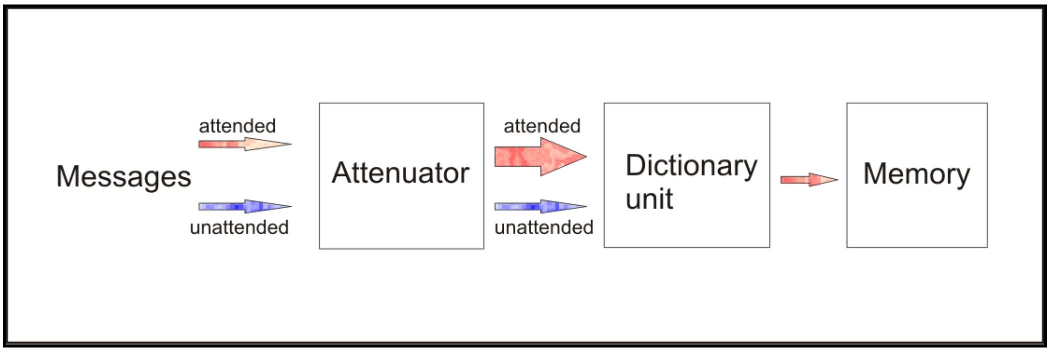


**Figure 2. A model of Treisman's attenuation theory.**

## Deutsch and Deutsch's Theory

Although Treisman (1960; 1964) places the filtering later in the processing of stimuli in comparison to Broadbent, other researchers such as Deutsch and Deutsch (Deutsch & Deutsch, 1963) postulate an even later selection. In their Response Selection Theory of Selective Attention they assume selection occurs after the pattern recognition stage (late selection) and suggest that all channels of information are recognized but are quickly forgotten unless they are pertinent in a given situation. Furthermore, Deutsch and Deutsch assume no initial filter for physical characteristics. They propose that all inputs are fully processed (including analysis of meaning). Within the dictionary network, each signal is analyzed and recognized for its importance. Only the importance and relevance of the input determines the response. Following their theory, an attention capturing sound is the sound that bears the strongest signal. With regard to the Cocktail Party Phenomenon, hearing one's own name during a conversation shows that the name when it is heard has importance.

However, is Deutsch and Deutsch's model really better than the one postulated by Treisman? Several investigations (Treisman & Geffen, 1967; Treisman & Riley, 1969) tested Treisman's theory and Deutsch and Deutsch's theory. Their experimental setup consisted of two different but simultaneous word streams. Furthermore, shadowing tasks were used which required attention to one and ignoring the other stream. The results showed that, in contrast to Deutsch and Deutsch's model, the shadowed message received greater processing than the unshadowed one. Treisman and Riley (1969) for example, told their participants to stop shadowing and to tap whenever a target occurs in either message. Thereby, many more targets were detected on the attended than the ignored message, which is inconsistent with Deutsch and Deutsch’s theory as well. Woldorff et al. (1993) in their neuropsychological studies also found support for Treisman's attenuation theory. In their experiment, selective listening to sequences of rapidly and dichotically presented tones of a certain pitch was reflected as larger magnetic brain responses for attended tones in the latency range of 20-50 ms post-stimulus as well as 80-130 ms in comparison to unattended tones. Furthermore, source localization yielded neural generators of those early attention selection processes within the primary auditory cortex and not in higher level areas or the according stages of cortical analysis as Deutsch and Deutsch assume. Thus, altogether Deutsch and Deutsch’s model seems less plausible than Treisman's. As seen, abundant evidence rather supports Treisman's model.

In general, humans, as well as most animals, respond more easily to novel objects and fast changes. This is also the reason why predators try to blend with their surroundings and move very little while sneaking up to their prey. Novel objects and fast changes are most likely carrying new information. In contrast to old objects, which are already inspected, it is profitable to analyze new information in greater detail. Fast changes should be inspected thoroughly because they, in contrast to slow changes, might affect us immediately. Thus, an early selection that may be refined in a later stage is indispensable, especially from an evolutionary point of view. It could be life-threatening, if we would process all stimuli equally until they reach the recognition stage, and if we would not distinguish earlier. Time is all that counts. This, finally, supports even more Treisman’s Attenuation Theory with a hierarchical selection pattern including both early and late selection (for a comparison see Figure 3).

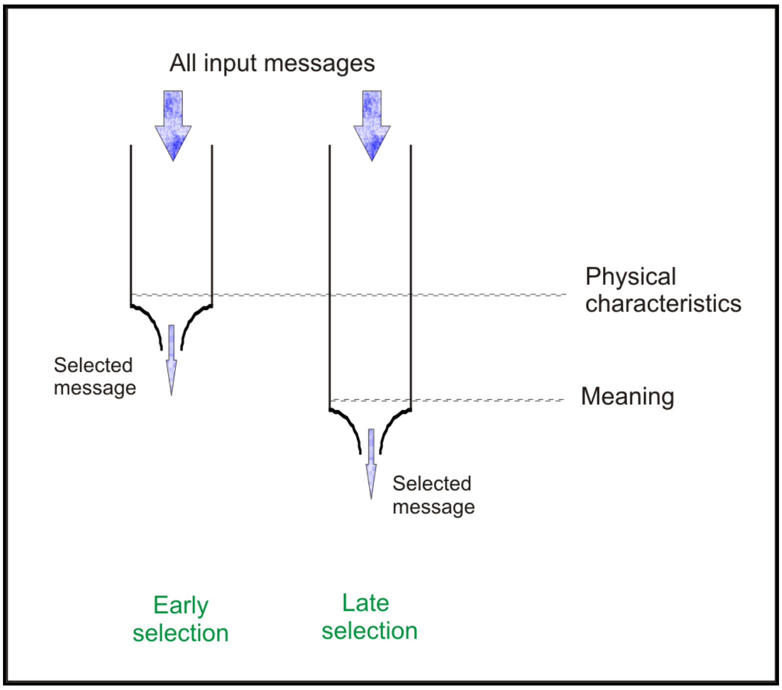


**Figure 3. A comparison between early and late selection and their bottleneck.**

There is also evidence that the point of selection depends on the task and its demands. Lavie (1995) argues for early selection in high-load tasks and late selection in low-load tasks.

This hypothesis was tested with a variation of the response competition paradigm with different loads. Either conjunctions or isolated features were processed in a simple detection or a difficult identification task. Interference of distractors was found only under low-load conditions. Usually the distractor was clearly distinct from the target. Thus, physical separation may not be a sufficient condition for selection. Overload of perception has also a main influence. Therefore, the early and late selection debate may be resolved if perceptual load of relevant information determines the selective processing of irrelevant information.

Taken together, these theories show that it is not solved yet whether selection takes place at only one stage, several stages or what the exact place or time course of these stages in the processing might be. The model that might accommodate all possibilities best is Treisman’s theory (1960; 1964) because, according to her model, unattended stimuli are attenuated before they are recognized (early selection), and not completely filtered out. At a later stage (late selection), only attended stimuli are able to reach working memory. In between both stages, attenuated, but possibly important stimuli (characterized by e.g. changes in pitch) are able to catch attention as demonstrated in some EEG studies (Woldorff et al., 1993). As was pointed out by Styles (1997), precisely discovering when and at what stage in the processing selection occurs is, however, only one small part of the issues surrounding attention. More important is to understand *why* or *how* this selection happens. Is the selection achieved by sorting by single features or is it object-based?

Several studies (both behavioral and electrophysiological) have demonstrated that preliminary analysis of auditory input and resulting partitions of it into distinct perceptual objects are the sine qua non to selectively focus attention on a certain sound source. Therefore, an object-based hypothesis was proposed in which auditory attention is allocated to perceptual objects derived from the auditory scene according to perceptual grouping principles such as grouping by proximity, similarity, continuation, closure, and symmetry. Hence, the observer's attention is allocated to an auditory object, instead of a feature per se. Pursuant to the object-based prospect, only the object defining parts and properties should be processed preferentially over other competing stimuli. There is evidence that object-based attention facilitates the response to any feature that appears inside the boundaries of the selected stimulus, even though they are not task-relevant. Hence, all the attributes of an attended object are bound together into a unitary representation. This is as if selecting a stimulus fills its boundaries with spatial attention, facilitating the information from the resulting spatial region. For more information see (Alain & Arnott, 2000; Scholl, 2001; Valdes-Sosa, Bobes, Rodriguez, & Pinilla, 1998) and also (Anllo-Vento, Schoenfeld, & Hillyard, 2004).

To gain further knowledge about how selection is carried out in the auditory modality, the use of simultaneous and competing speech streams in realistic auditory environments seems a reasonable approach. If this approach is combined with the recording of event-related brain potentials (ERPs), it is possible to determine to what extent a stimulus is processed on the neural level without the need of obtaining behavioral responses to that stimulus. Differences in ERP components between attended and unattended stimuli can be taken as evidence for an attention system that attenuates irrelevant information in favor of relevant stimulus processing as Treisman postulated.

The main aim of the present investigation is to learn more about the attentional focus, its precision and extent. Is the focus strongly restricted to the to-be-attended stimulus and its defining features? To address this question, a variety of stimulus features (frequency, spatial location, and spectral composition) were changed. Either a slight or significant variation in a stimulus' feature was applied while attention focused on a human speaker's fundamental frequency and spatial location. By doing so while EEG was recorded, changes in stimulus processing may be observed. According to Treisman's Attenuation Theory, it can be assumed that, as soon as a stimulus does not match the to-be-attended pattern in any stimulus characteristic, processing of this stimulus will be attenuated, indicating it has left the attentional focus. By altering stimuli, especially by using both slight and extreme deviant stimuli, one can also clarify the boundaries of the attentional focus much better.

### *Event-Related Brain Potentials and Attention*

One way to study selective attention in the auditory modality involves the use of event-related brain potentials (ERP) during a dichotic selective attention task. The ERP can be used to determine the time course on a millisecond-level of the neural events underlying the processing of both attended and unattended stimuli. This allows exploring which neural responses in which timeframe are actually sensitive to attention and to what degree. In general, ERPs reveal attentional modulation very clearly.

Event-related brain potentials (ERPs) are small negative or positive voltage changes in ongoing electrical brain activity evoked by a cognitive, sensory or motor event. By attaching electrodes to the scalp, the potentials can be measured. Those electrodes detect the ongoing electrical activity of the brain at their particular scalp location. EEG signals are very "noisy" (because there is always a lot going on in the brain not directly related to the stimulus event). The event-related potential is determined by averaging the brain signals time-locked to the onset of the same stimulus presented many times. Thereby the spontaneous activity is canceled out. Thus, the ERP represents the time-locked activity following a certain and definable stimulus event. Its peaks and troughs can vary in latency, amplitude, polarity, and distribution over the scalp. They reflect the activity of groups of neurons involved in the perceptual, cognitive, emotional, or motor processes in response to an event. When different groups of neurons are involved in two (or more) different processes, they can cause distinct scalp distributions or polarities. Thus, processes evoked by different conditions can be discriminated (Rugg & Coles, 1995). Some of the changes in the waveform are known to be specifically related to attention processes. These components have been shown to be affected by various types of attention deficit disorders. In the following, ERP-correlates of attentional processes will be discussed in detail. First, general attentional effects (P1, N1, and Nd) are highlighted. Then, ERP components (specifically MMN and P3) that are modulated by attention in “oddball” experiment paradigms (such as Experiment Three in this dissertation, see Chapter 4) will be discussed.

## General attention effects

### *Visual selective attention*

Until today, most research on selective attention has been conducted in the field of vision. Some findings of this research shall be mentioned here, because they illustrate the general pattern of effects.

A supramodal mechanism such as attention is, as the name already implies, a higher-level mechanism and not modality specific. Attention is therefore supposed to be equivalent for all modalities (e.g. auditory, visual, tactile). Different modalities are controlled by the same attention mechanism. This means that even task-irrelevant modalities of a to-be-attended object also gain in processing strength (Busse, Roberts, Crist, Weissmann, & Woldorff, 2005; Eimer & Van Velzen, 2002).

The earliest exogenous component modulated by visual attention represents the P1 with larger peaks for attended compared to unattended stimuli (Luck & Hillyard, 1999). Luck et al. (1990), for example, recorded ERPs as subjects attended to the left or right hemifield of a visual display while fixating a central point. The stimuli occurred unilaterally on either the attended or unattended side, or bilaterally. If stimuli were presented unilaterally in the attended hemifield or bilaterally, then an enhanced P1 for stimuli from the attended hemifield was observed. Thus, the P1 attention effect is modulated by voluntary attentional allocation between competing locations (Heinze, Luck, Mangun, & Hillyard, 1990; Luck, 2005c).

Figure 4 shows an idealized P1 modulation (increased amplitude for attended stimuli) within the ERP resulting from a selective visual experiment. Handy and Mangun (2000) showed that the P1 can also be influenced by variations in perceptual load. They investigated both low- and high-perceptual-load targets in a probabilistic spatial cuing paradigm. Thereby, “perceptual load” is defined by the amount of task-relevant information within a search array based on the total number of items and their complexity. The results showed that the magnitude of spatially selective processing (seen in a larger P1) in extrastriate visual cortex is increased with perceptual load. P1 effect reflects a mechanism that suppresses activity at unattended locations to avoid interference (Luck, 2005c). Luck also argues that P1 attention effect is more apparent the stronger the distraction. Hence, the information processing system can be biased.

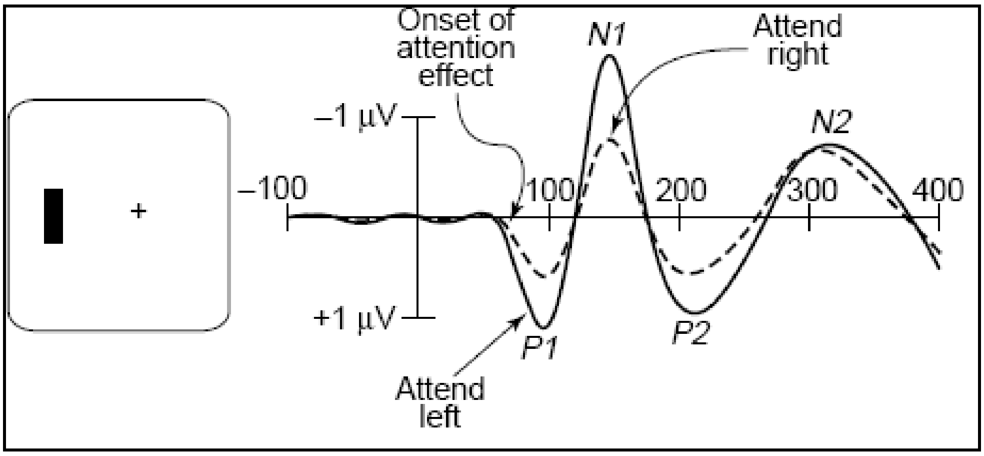


**Figure 4: Paradigm for using ERPs to study attention. Left: Stimulus display. Right: idealized results. Subjects task was to fixate a central cross while attending either to the left or right visual field. In a rapid sequence, stimuli are then presented to the left and right visual field. In the present example, the ERP elicited by a left and therefore attended stimulus evokes a larger P1 and N1 component than right/ignored stimuli.** ***Figure taken from Luck, Woodman, and Vogel (2000).***

The P1 is the earliest sensory-evoked response to be reliably modulated by attention (Eason, 1981; Luck et al., 1994; Mangun, 1995; Mangun & Hillyard, 1991; Van Voorhis & Hillyard, 1977). Hillyard et al. (1998) have interpreted their ERP results as a reflection of a sensory ‘gain control’ mechanism that simply causes larger P1 responses for attended-location stimuli compared to unattended stimuli relatively. Thus, the P1 reflects a gating (Luck et al., 2000). The higher the amplitude the wider the gate is open. The responsiveness to stimuli at attended locations is amplified, and therefore, further processing of these stimuli will be enhanced, whether the stimuli are task relevant or not.

The P1 is followed by another ERP component modulated by attention, the N1. Compared to the P1, the attentional mechanism reflected by the N1 is less clear. There is an assumption that the N1 contains more than one active attentional process (Luck & Hillyard, 1995; Luck et al., 2000; Mangun, 1995), the “N1 discrimination effect” and the “N1 reorienting effect”. As seen above, the first N1 attention effect (larger N1 amplitude for stimuli in the attended location) seems to reflect an enhanced processing of attended stimuli which may involve a discriminative process (N1 discrimination effect) applied to a restricted area in sensory space (Luck et al., 2000) ; e.g. relevant vs. irrelevant discriminative process. The second process (N1 reorienting effect) represents an enhanced negativity at the N1 latency reflecting an attentional switch from one location to another (Heinze et al., 1990). Moreover, recent research has shown that the N1 effect attenuates over time. Interestingly, only the N1 discrimination effect seems to be effected by repetition of stimulus presentation. The N1 reorienting effect, in contrast, does not change with time on task (Boksem, Meijman, & Lorist, 2005).

In general, it is assumed that the N1 reflects precise detection of an incoming stimulus (Heinze et al., 1990, study mentioned above) after it was gated (reflected in the previous P1 component). For an example of this N1 modulation see Figure 4. The N1 is regarded as an early attention related negativity, and reflects a matching process, comparing the input with an actively maintained representation of the stimulus that needs to be attended (attentional trace hypothesis).

“Attentional trace hypothesis” of selective attention was developed by Näätänen (1982; 1990; 1992). According to this hypothesis, the initial selection is accomplished by a comparison between the sensory input and an attentional trace in auditory cortex. If the same task-relevant (attended) target feature matches repeatedly and consistently incoming sensory information, it gets associated with a learned attentional response (preferred stimulus processing). Because a memory trace is formed, the task-relevant target feature is better and easily selectable. Thus, this trace constitutes a voluntary maintained representation of physical features that separate task-relevant from irrelevant (unattended) stimuli. A consciously controlled system operates this comparison process. The better the match between relevant and irrelevant stimuli, the longer the process of early selection continues.

Currently, this N1 enhancement is seen to simply typify the relative increase of neural networks’ excitability for attended versus unattended locations and shows the attentional modulation of the sensory processing with an enhanced processing of attended stimuli.

Next in time, following the N1, a selection negativity (SN; subtraction of unattended from attended stimuli) becomes obvious, comparing the ERPs to unattended visual stimuli with the ERPs to the same stimuli when they are attended, This SN, with an increased negativity in the ERPs for attended stimuli, has an onset partially overlapping the N1 time window (150-200 ms), but is also reported at a slightly later time frame with an occipital maximum with a duration of several hundreds of milliseconds. The onset latency of the SN reflects precisely the time at which an attended feature is discriminated and selectively processed (Smid, Jakob, & Heinze, 1997). The selection negativity is best observed in difference waves in which the ERPs of stimuli with unattended feature(s) is subtracted from ERPs of the same stimuli when the feature is attended to. Stimuli are not only selected on the basis of a single task-relevant feature, but also based on conjunctions of features or multiple features. Thus, different features of the same object are assumed to be processed together in parallel. Among others, this was shown by Anllo-Vento et al. (2004), who studied cortical events while presenting multi-feature objects. They first analyzed first the timing and brain sources of the SN associated with selection of the individual features. Then, the SN elicited by the selection of the conjunction of features of the to-be-attended object was inspected. Their subjects looked at sequences of bar gratings randomly mixed with either vertical or horizontal orientation and printed either in red or orange (dark grey vs. light gray; see Figure 5A). In different runs, the attention was directed to only one of the four possible conjunctions (red/vertical, red/horizontal, orange/vertical, and orange/horizontal). Thus, the presented stimuli could contain one feature (C+O- for to-be-attended color; C-O+ for to-be-attended orientation), two features (C+O+ for both to-be-attended color and orientation), or none (C-O- for unattended color and orientation) with the attended combination. The results show a clear SN (starting at about 140-180 ms, see Figure 5B) between the attended (C+O+) and completely unattended (C-O-) feature-conjunction because of the less negative going ERP wave of the latter one. Additionally, the study of Anllo-Vento et al. also showed that the color specific SN (C+O- minus C-O-) had a later onset latency (180 ms) than the orientation specific SN (C-O+ minus C-O-) at 140 ms. Interestingly, the conjunction SN (C+O+ minus C-O-) had the same onset latency as the earliest onset of the single feature SN (in this case the orientation specific SN, 140 ms). This, however, did not mean that the conjunction processing was conducted in a simple parallel manner. The single features were also integrated with one another as was reflected later on in the conjunction SN (color-orientation combination) with a later onset than the individual SNs (color and orientation). This indicates that additional neural activity is involved with the selective processing of the conjunction itself. This conjunction-specific processing can be seen in the difference wave (sum of SNs C+O- and C-O+ minus SN C+O+) which reflects a SN beginning at around 225 ms. Therefore, Anllo-Vento et al. (2004) concluded that there is a stage of independent, parallel feature selection first (140-225 ms) and thereafter a conjunction processing is started along with continued parallel processing of the color and orientation features); see Figure 5.

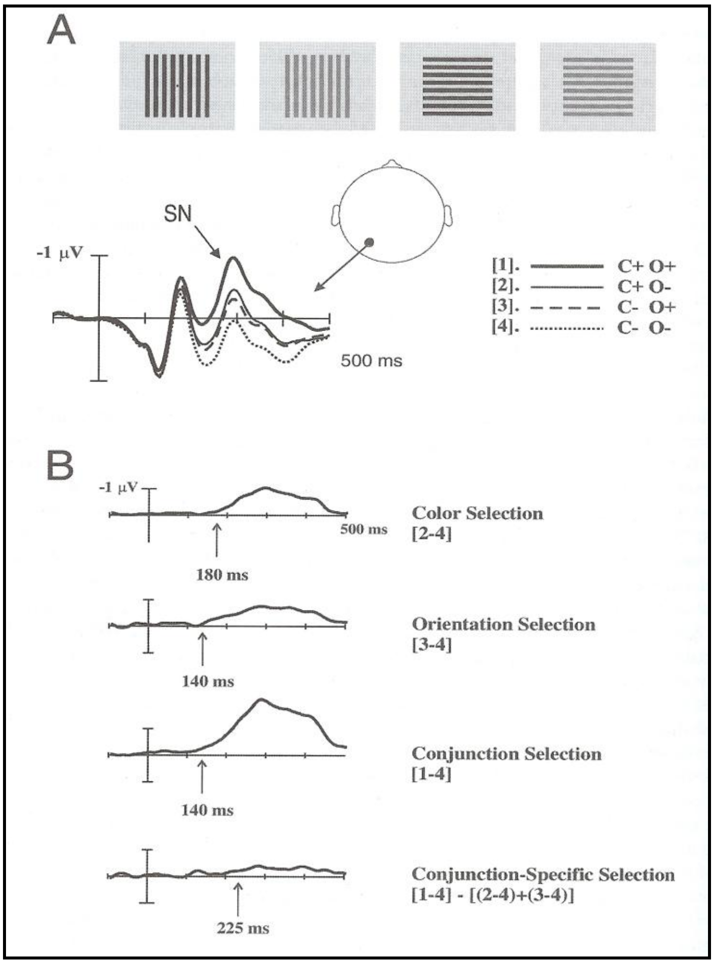


**Figure 5: (A) Diagram of the four standard stimuli presented one at a time randomly mixed. Bars in the display were either horizontal or vertical, or red (dark bars) or orange (light bars). During EEG recordings, subjects attended to a particular combination of bar color and orientation. Below the bar displays on the left, averaged ERPs over the left ventral occipital scalp are shown. The waveforms were elicited by (1) stimuli that shared both relevant features (C+O+); (2) stimuli having the relevant color (C+O-) only; (3) stimuli having the relevant orientation (C-O+) only; and (4) stimuli having neither of the relevant features (C-O-). (B) Difference waveforms display the selection negativities to selection of the relevant color, orientation, conjunction of color and orientation, and the difference between the conjunction SN and the sum of the color and orientation difference waves.** ***Figure taken from Anllo-Vento et al. (2004).***

### *Auditory selective attention*

A typical design to study auditory selective attention comprises different kinds of sounds that were presented in a rapid, randomized sequence. Therein, listeners had to attend to a particular set of sounds and to ignore unwanted sounds. Furthermore, dichotic listening tasks were used in which different sounds or messages were presented simultaneously to each ear. Stimuli from the designated ear needed to be shadowed whereas stimuli from the opposite ear were ignored. Regarding attentional modulations, ERPs of attended and unattended stimuli were compared to each other.

The P1 component is one of the earlier exogenous (stimulus driven) ERP components found in auditory ERPs. This means that the P1 changes with varying stimulus characteristics. In a study by Alain and Izenberg (2003), subjects were asked to focus their attention on tuned and mistuned stimuli presented to one ear and to ignore the tones presented to the other ear. Tuned stimuli contained only consorting harmonic components, whereas mistuned tones had one harmonic component of a complex sound changed so that it was no longer an integer multiple of the fundamental stimulus. These tuned and mistuned tones occurred either as long (standard) or short (deviant) sounds presented on both ears. Furthermore, they also manipulated the task load by varying the task instructions. The easy task load consisted of a button press response for deviant sounds (shorter (mistuned) sounds among long and tuned sounds) at the attended location only; tuning did not matter. In contrast, the hard task required in addition to the deviant detection also a categorization of tuned or mistuned sounds. Reaction times were faster when targets defined by duration only were to be detected than when by both duration and tuning were task-relevant. Further, they observed a reduced P1 wave for unattended stimuli over the hemisphere contralateral to the attended location indicating a modulation effect by auditory spatial attention.

Also similarly to the visual modality, the P1 provides evidence for an early selection effect for auditory stimuli. Thus, it reflects the early processing of auditory, as well as visual, stimuli with an onset between 60 and 100 ms (Hansen & Hillyard, 1984) and has a characteristically fronto-central scalp distribution (Coch, Sanders, & Neville, 2005). A second exogenous ERP component important for auditory stimuli processing as seen in visual processing, is the N1. At a latency of around 100 ms, this stimulus driven negativity shows its maximum with a central to fronto-central distribution within the auditory modality. Nevertheless, because the N1 is an exogenous component, thus sensory specific, the auditory cortex plays the important role for auditory stimuli compared to the occipital activity in the visual modality apparently.

Moreover, as discussed above for visual stimuli, the N1 represents a mechanism for change detection. It can be elicited by a fast change in the stimulus energy level (stimulus onset) for example. Thereby, the amplitude is determined by the physical properties of the sounds (e.g., intensity and presentation rate) whereas the MMN (mismatch negativity in an oddball design) mechanism detects regular deviations in ongoing auditory stimulation (Alho, Donauer et al., 1987).

The mismatch negativity (MMN) is a change-related brain response. It is one of the auditory ERP components, and elicited task-independently by an infrequent change in a sequence of other frequent stimuli. The MMN is evoked in response to violations of simple rules controlling the properties of auditory information. It is supposed to reflect the automatic formation of a short-term neural model of physical or abstract regularities in the auditory environment (Picton, Alain, Otten, Ritter, & Achim, 2000). A typical experimental setting to elicit an MMN provides an oddball paradigm. Therein, a standard stimulus is presented repeatedly and frequently, whereas an "oddball" stimulus appears only occasionally.

An enhanced N1 is also found by temporal and spatial attention whereas both attention effects had a similar scalp topography, suggesting common neural generators within the temporoparietal regions (auditory areas) (Lange, Krämer, & Röder, 2006). One of Hillyard et al.'s first studies (1973) has demonstrated that the amplitude of the N1 depends on attention such that a larger N1 was observed for attended compared to unattended stimuli. At that time it was assumed that this attention effect resulted from selective tonic facilitation of processing of the attended input, which is in line with Broadbent's filter theory of selective attention (Anllo-Vento et al., 2004; Broadbent, 1958).

Selective auditory attention and the effect on auditory evoked potentials are often specified by subtracting the ERPs to unattended stimuli from the ERPs to the same stimuli when they are attended (Hansen & Hillyard, 1988; Näätänen, 1982). This subtraction results in, similar to the visual SN (selection negative), a negative difference (Nd; reflecting the orientation to or further processing of an auditory input deemed relevant in preliminary sensory analyses (Näätänen & Michie, 1979)) wave, a.k.a. PN (processing negativity). Whereas the PN is a generic term for subtraction waveforms that reveal target-related processing, the Nd is a more precise measure making use of the same physical stimulus with and without attention directed to it. This subtraction procedure controls for all stimulus features except "target-ness" (the differential processing of attended and unattended channels) and possible momentary changes in the state or direction of attention. Figure 6 shows such a Nd effect by Hansen et al. (1984). Subjects attended selectively to a sequence of tones of one frequency while ignoring a sequence of different frequency and spatial origin. The ISI (interstimulus interval) varied between runs (short/fast vs. long/slow). ERPs were recorded and compared between conditions. To achieve this negative difference, the authors simply subtracted the ERP waveform of attended pips from the one of unattended tones resulting in difference wave. Because attended stimuli are processed with a more negative polarity, the result of this subtraction remains negative. Thus, the Nd reflects a stronger processing for attended stimuli. Furthermore, the same study showed that the Nd had a shorter onset and was smaller in amplitude for faster compared to slow rate stimulation. In another study, Hansen and Hillyard (1983) used multidimensional auditory stimuli in a selective-attention task. Tone pips varied orthogonally between two levels (easy and hard to discriminate) each of pitch, location, and duration. The participants were instructed to selectively listen to one stimulus class at a time. The task itself consisted of a response to longer duration target stimuli of a specific pitch or location. Whether the attention should be focused on pitch and/or location targets was indicated by pitch and/or spatial cues beforehand. Stimuli that did not match the target tone and were easily discriminable from it elicited only transitory selective effects. In contrast, tones that were hard to distinguish from the target indicated more extensive selective processing.

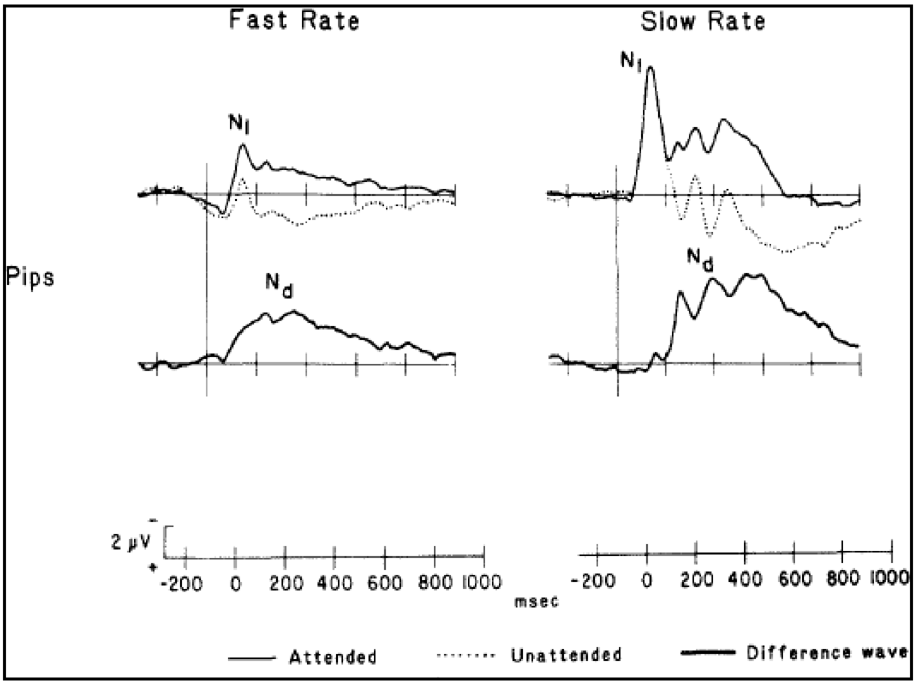


**Figure 6: Grand average ERPs recorded at Fz elicited by pips (100 ms duration) at fast and slow rate presentation. The attended (solid line) and unattended (dotted line) potentials were subtracted (attended minus unattended) to produce the difference waves plotted below each attended/unattended pair. These difference waves illustrate the resulting Nd (negative difference) explicitly. *Figure taken and modified from Hansen and Hillyard (1984).***

The Nd is generally regarded as a specific marker for an attended class of stimulus (Hansen, Dickstein, Berka, & Hillyard, 1983; Jemel, Oades, Oknina, Achenbach, & Röpcke, 2003; Näätänen, 1988; Oades, Ditmann-Balcar, & Zerbin, 1997). It is an endogenous (voluntary processing/top-down) negative component in the ERP for attended tones with a fronto-central maximum while attending to tones occurring in a rapid succession at one location and ignoring tones at another location (Alho, Donauer et al., 1987; Alho, Teder, Lavikainen, & Näätänen, 1994; Hansen et al., 1983; Hansen & Hillyard, 1980). This negativity difference consists of at least two partially overlapping components, the early component (Nde) and the late component (Ndl) respectively (Hansen & Hillyard, 1980). This effect is mostly seen in dichotic listening tasks (Hansen & Hillyard, 1988).

The Nde reflects rapid initial featural analysis of stimuli, and the selection between attended and unattended stimuli. Näätänen (Näätänen, 1992) proposed that the early Nd simply reflects a temporary feature recognition system. It determines the suitability of the stimulus for further processing. Therefore, the basic physical characteristics of each stimulus are compared to a template stored in memory. As soon as a stimulus occurs that differs from the template, the matching process terminates. Hence, the closer the match between the template and the input is, the larger the amplitude and the longer the latency of the early Nd. Several studies located the generator for the early Nd in the auditory cortex (M. Giard, Perrin, & Pernier, 1990; M. Giard, Perrin, Pernier, & Peronnet, 1988; Rif, Hari, Hämäläinen, & Sams, 1991; Woldorff et al., 1993).

Nevertheless, the Nde can overlap the N1. Some researchers also talk about the N1/Nd negativity (W. A. Teder-Sälejärvi et al., 1999). There is some evidence that the N1 is not modulated by attention itself, as usually assumed. The effects of the slow wave Nde are observed in the typical N1 time window instead (Näätänen, 1982; D. L. Woods, 1990), whereas other evidence still suggests that the N1, as well as the M100 (the MEG counterpart to the N100), is modulated by attention itself (Woldorff et al., 1993). Näätänen et al. (1978) found that the attention-related Nd, such as observed by Hillyard et al. (relatively short ISI, 1973), is dissociable in time from the N1 component. Näätänen et al. have shown that the attentional negativity (Nd) emerged around 150 ms after stimulus offset and persisted for 500 ms at least by using a longer and constant ISI (800 ms). Therefore, they suggested that the N1 effect noticed by Hillyard et al. may have been caused by an endogenous Nd overlapping the exogenous N1, rather than by an intensification of the N1 generator process. According to Näätänen, the shorter ISI could have shortened the Nd latency because of the more rapidly required processing of the stimuli.

This assumption is supported by results from other studies. Authors of these studies (Hansen & Hillyard, 1984; Parasuraman, 1980; Schwent, Hillyard, & Galambos, 1976; D. L. Woods & Alain, 1992) argued that the onset latency of the Nd wave depends on the rate of stimulus delivery - the faster the stimulus presentation rate the shorter the Nd onset latency. The early Nd and its latency (50-200 ms) is thought to be closely related to the discriminability of attended and unattended stimuli; more than to the discriminability of standard and target stimuli within the attended sequence. Thus, the more the attended and unattended stimuli differ the earlier the Nde. Parasuraman (1980), for example, presented two concurrent sequences of tone bursts with occasional targets which the subjects had to respond to. He found that the latency and amplitude of the Nde within the attended sequence varies with the distinctiveness of stimuli, but not with the distinctiveness of standard and target tones. When participants attend to a particular stream of sounds in the presence of one or more different streams of distracting stimuli and when stimulus sequences are easily discriminated from the distractors, whether they are distinguished by spatial position, frequency, or both, then the attention-related changes such as the Nd occur in ERPs (Hink, Frenton, Pfefferbaum, Tinklenberg, & Kopell, 1978; Schwent & Hillyard, 1975; Schwent, Hillyard et al., 1976).

Another characteristic of the Nde is that the amplitude and latency are usually maintained over long sessions and therefore are little affected by repeated testing (Shelley et al., 1991; D. L. Woods, 1990; D. L. Woods & Clayworth, 1987; D. L. Woods, Hillyard, & Hansen, 1984). Only one study (Donald & Little, 1981) reported a habituation of the Nde.

The late Nd is not as well understood as the Nde. Typically, it occurs in the range of 300-600 ms after stimulus onset, and is associated with selective rehearsal and maintenance of the attentional trace. Its origin is assumed to be in the mid-frontal regions (more anterior and longer in duration than the Nde). EEG, PET and fMRI studies found activation within multiple areas in the frontal and temporal cortex during auditory selective attention (Alho et al., 1999; Jäncke, Mirzazade, & Shah, 1999; Jemel et al., 2003). The functional meaning of this late Nd was described as an "attentional supervisor" (Alho, Donauer et al., 1987). It reflects keeping the stimuli in mind and enables a further processing of the stimuli (Wijers, Mulder, Gunter, & Smid, 1996). Evidence for the late Nd as being more sensitive than the Nde to the magnitude of the physical difference between the attended and unattended stimuli is provided by Hansen and Hillyard (1980). They also showed that the Ndl is larger at longer interstimulus intervals (Hansen & Hillyard, 1984). Subsequently, Näätänen assumed that it may reflect further processing after the stimulus has been identified as belonging to the to-be-attended category or the selective rehearsal of the to-be-attended stimulus. Another proposal would be that the late Nd reflects the transfer of information about the target to an executive mechanism (Näätänen, 1990). One characteristic of the late Nd is that this effect is reduced by practice. Therefore, it has been suggested that it reflects either the functioning of a central executive or the strategic updating of sensory information needed to maintain selective attention (Shelley et al., 1991; D. L. Woods, 1990). Woods et al. (1994) have shown that this component involves some form of higher-order processing, and is therefore sensitive to a conjunction of stimulus features rather than to a single feature. They found that the Nd for feature conjunction began before the analysis of individual features was finished.

So far, studies are discussed that used tones and not speech in a dichotic listening task. There are also a few studies that investigated selective spatial attention effects in the context of a spoken narrative, and mostly observed typical ERP attention effects (Hink & Hillyard, 1976; Teder, Kujala, & Näätänen, 1993; Trejo, Ryan-Jones, & Kramer, 1995; D. L. Woods et al., 1984). Irrelevant vowel probes within an attended speech passage caused more negative going ERPs than the same probes in the unattended passage in dichotic presentations (Hink & Hillyard, 1976). Woods et al. (D. L. Woods et al., 1984) presented in a more complex design various speech and tone probe stimuli embedded in dichotically presented prose passages. Speech and tonal probes were spatially and temporally coincident in the attended passage and the unattended passage. Subjects had the task to either shadow or selectively listen to one of the dichotic passages. The results for attended speech probes revealed an enhanced negativity (Nd) beginning at about 50 ms and lasting to 1000 ms (D. L. Woods et al., 1984).

Besides reflecting an activation of neural populations in auditory cortex by source localization (M. Giard et al., 1988; Näätänen, 1992), Nd waves are also influenced by the stimulus features of attended stimuli. It has been shown that Nd waves associated with pitch and location processing have different scalp distributions (Schröger, 1994; D. L. Woods, Alho, & Algazi, 1991). Attended compared to unattended tones/pitches elicited an enhanced frontal negativity whereas ERPs to selected locations showed a parietal Nd which seems to be characteristic for spatial attention. This indicates modulation of distinct auditory cortical fields.

Taken together, the processing of attended auditory stimuli is stronger and more elaborated than unattended stimuli. This selective attention effect starts very early at around 100 ms. Although the P1 and N1 are exogenous they can be modulated by attention if the stimulus characteristics are not too distracting. A separate attention effect is reflected by the Nd. This effect is endogenous and therefore driven by higher-level control mechanisms. Discriminability seems to play an important role on the size of the Nd: the more discriminable the stimuli are the larger and earlier the negative difference. Moreover, it was found that attention is not only directed to a single relevant feature of a to-be-attended object, but also to all other features/conjunctions of the same object. The scalp distribution could differ depending on stimulus features, such as location, or pitch.

The studies presented in this dissertation may shed light on some of the uncertainties discussed here, or may provide support for previous interpretations. Different pitches and locations for the stimuli are manipulated in order to investigate attentional focus in greater detail and whether changes in feature-conjunctions will lead to reduced processing of affected auditory objects. The main aim of the current studies is to examine auditory selective attention under more complex and realistic situations compared to the rather simple structured experiments mentioned above. Furthermore, the validity of the HRTF method (described in *Methologies* of this chapter) as a more realistic auditory environment is investigated.

## Specific auditory selection effects – the oddball design

### *Auditory selective attention*

### **MMN effect**

The initial recording of auditory information into our memory system is accomplished through auditory sensory (echoic) memory. It is a rapid storage of brief duration (2–3 s), and encodes physical stimulus characteristics in detail (e.g., Cowan, 1995; Darwin, Turvey, & Crowder, 1972; Neisser, 1967). To study sensory memory the ‘‘oddball’’ task is often used. In this paradigm a repeating signal (the ‘‘standard’’) is replaced periodically by an infrequent deviant (the ‘‘oddball’’). Auditory event-related potential (ERP) studies often show that oddballs reliably elicit a negative potential (relative to standards) approximately 100–300 ms after stimulus onset, even though subjects are instructed to ignore the auditory stimuli while engaged in a primary task, such as playing a video game or reading a book (Näätänen, 1990, 1992). An example of such a MMN is displayed in Figure 7.

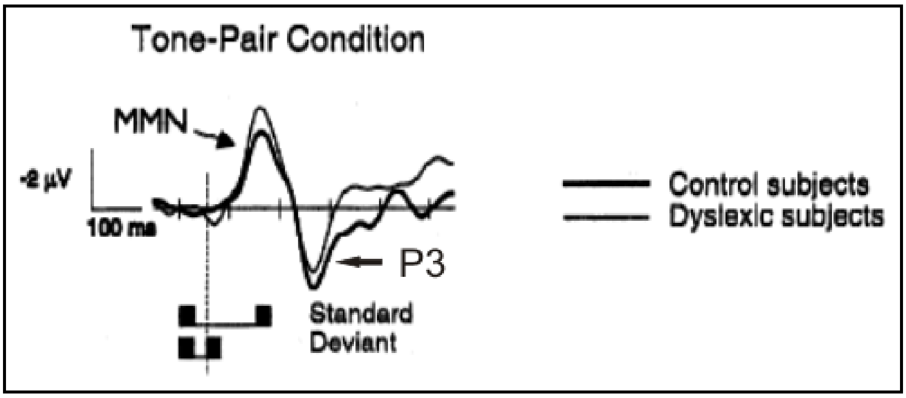


**Figure 7: Difference waves (Fz event-related potentials to standard stimuli minus those to deviant stimuli) in control (thick line) and dyslexic (thin line) subjects. While reading a book, paires of standard and deviant tones were presented. Both groups (controls and dyslexics) were able to elicit a distinct MMN. The typical MMN following P3 was observed as well.** ***Adapted from Kujala et al. (2000).***

The mismatch negativity (MMN) response is assumed to reflect the output of an automatic change-detection system. MMN increases as a function of deviance. This negativity can be elicited by all possible deviants (e.g. pitch, loudness, duration, location), only the topography varies in dependence of the deviant feature (Picton et al., 2000). Giard et al. (1995) found that the scalp topography differed as a function of the type of deviant (intensitiy, frequency, and duration). They also proposed that different neural generators underlie those different MMNs. Neural sources for frequency MMN are more anterior than the ones for the duration MMN (Frodl-Bauch, Kathmann, H.J., & Hegerl, 1997) or for deviance in pattern (Alain, Achim, & Woods, 1999). Moreover, Alho et al. (1996) suggested that processing of a deviance is rather specific to the type of standard stimulus (e.g. tone or a chord; simple or complex) than to the physical change (frequency), because they found a different MMN source after a change in tonal frequency (simple tone) compared to a similar change in a combination of musical notes (chords/complex sounds).

The anatomic basis of sensory memory and the degree to which anatomic loci are shared between sensory systems remains largely unknown. Some researchers have studied the precise location of the MMN generator within the temporal cortex by using EEG/MEG and source localization (M. H. Giard et al., 1995; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001; Tiitinen et al., 1993; Tiitinen, May, Reinikainen, & Näätänen, 1994). Other studies showed scalp potential distributions and current density distribution measures indicating an additional involvement of the parietal and frontal regions in detection of acoustic changes (Baldeweg, Klugman, Gruzelier, & Hirsch, 2002; M. Giard et al., 1990; Rinne, Alho, Ilmeniemi, Virtanen, & Näätänen, 2000; Scherg, Vajsar, & Picton, 1989).

Müller's group (2002) developed a two-stage model of auditory deviance detection. Small stimulus changes are processed in a network of posterior STG (superior temporal gyrus; first processing) and IFG (inferior frontal gyrus; second processing), whereas large stimulus changes are analyzed in more detail which additionally recruits mid-dorsolateral prefrontal cortex. Therefore, those temporal and frontal generators are thought to comprise different, but closely related functions in the pre-attentive change detection. It has been hypothesized that, in the absence of attention to an auditory stimulus, changes in the stimulus are detectable automatically in temporal cortex. For directing attention towards the deviant stimulus and preparing subsequent attentive processes, an attention reorientation mechanism is necessary and is found in frontal cortex activity (Näätänen, 1990, 1992; Näätänen & Michie, 1979). Schröger (1997) proposed a pre-attentive activation model. Therein, via several hypothetical processing stages, an obligatorily operating deviance detection system is assumed that finally leads to the conscious perception of infrequent deviant sounds among frequent standard stimuli. By encoding invariance to recent auditory stimulation into short-term representations in the auditory sensory memory and comparing each current input with these memory traces, this system acts pre-attentively. This means that no explicit intention to detect deviants or even attention is necessary, although the system may be sensitive to attentional modulations. Furthermore, Schröger assumed that the output of this feature-specific pre-attentive deviance detection system may flow into a general and integrated mismatch signal activating subsequent processes that are responsible for specific motor responses. Figure 8 shows Schröger's pre-attentive activation model graphically.

This may also be the reason why the MMN component is often followed by a P3a component (elicited by a novel stimulus in an active three-stimulus oddball-paradigm; also called the "novelty" P300 compared to the "target" P300 (P3b) in the same paradigm), indexing a switch of attention (e.g., Knight, 1996).

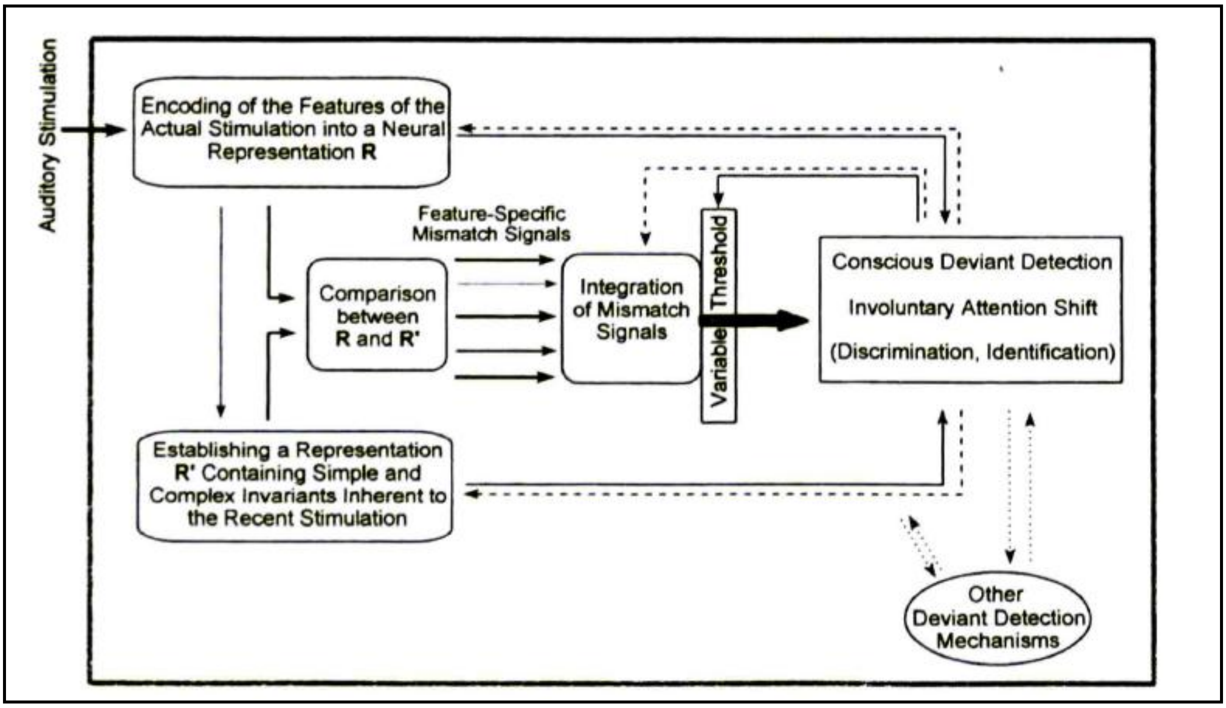


**Figure 8: This model from Schröger, 1997, explains the conscious detection of deviant sounds which represent irregularities with regard to some discrete repetitive stimulation as the result of a preattentive operating deviance detection system. Current stimulus information and invariances of recent stimulation are extracted into representations of sensory memory. These, further on, are compared on a feature-specific basis. If discrepancy is detected, appropriate and proportional mismatch signals are generated. This information then will be accumulated into an integrated mismatch signal. As soon as a variable threshold is exceeded, subsequent processes will be activated resulting in consciousness that can be seen, e.g. in a button press.** ***Adopted from Schröger (1997).***

However, the MMN has evolutionary relevance since it reflects significant (potentially dangerous) events in the environment, which helps to stay alert independently of conscious awareness, or to direct attention (e.g., detecting the alarm clock while sleeping). Thus, the MMN has the advantage that it is relatively independent of attention and can be recorded during altered states of consciousness such as coma and sleep (Näätänen et al., 2001). Nevertheless, the attentional level determines the MMN magnitude, with larger amplitude for attended stimuli than for unattended stimuli, especially for the right ear (Alain & Woods, 1994, 1997; McKenzie & Barry, 2005; Woldorff, Hackley, & Hillyard, 1991; Woldorff, Hillyard, Gallen, Hampson, & Bloom, 1998). Otherwise, the amplitude of this component reflects the quantity of change, implying that a representation of prior stimulus features underlies its generation (Näätänen, 1990).

When a sequence of identical stimuli is presented, a complete representation of physical features of the repetitive stimulus is established and stored as a memory trace (Alho et al., 1999). The efficiency of the comparison between neural traces to deviant stimuli and to the uniform stream of standard stimuli impacts the operation of the change-detection system that can be seen in the MMN. Deviance in any physical feature of the incoming stimulus, such as frequency, intensity, duration, or location, elicits a mismatch response (Näätänen, 1992).

As long as the neural representation of the standards exists, the MMN will be elicited at the moment when the deviant occurs (Winkler & Näätänen, 1995). After the memory trace for the standards has extinguished, no MMN will be elicited. Thus, an active sensory memory trace of standards is necessary for activating the change detection mechanism underlying the MMN.

Sabri and Campbell (ERP, 2001) recently discovered a monotonically increasing MMN amplitude with the rate of stimulus presentation, whereas the probability of a deviant occurring in time was held constant (e.g., one deviant on average every 9 s). Indeed, if the rate is slower than 2.4 s the MMN was absent (Alain & Woods, 1994). The reason for this appearance could be that faster rates yield stronger memory traces, perhaps because of frequent repetitions and therefore a stronger encoding within the time-limited sensory memory buffer.

Furthermore, McKenzie et al. (2005) proved that there are independent memory traces for dichotic presentations – one separate memory trace for each ear. They concluded that the ability to separate representations of the attended stimuli subsets into different memory traces is a core mechanism of sustained attention. Before this finding, the MMN in a dichotic listening paradigm was normally elicited by deviants presented to the same ear as the standard stimuli which in fact would provide evidence for an independent representation for each ear. Carlyon et al. (2001) suggested that attention affects streaming (perceived melodic line). This assumption conforms to Ritter et al.'s (2000) result showing that the attended MMN reflects attended standards and the unattended MMN reflects unattended standards. To explain these different MMNs, two different stimulus representations/memory traces regardless of their source are necessary.

## P3 component

Sutton, Braren, Zubin und John (1965) first discovered the "classic" P3 (also called P300, or P3b; the third positive wave, or the wave with a 300 ms latency). Today, it is the most investigated component, perhaps because the P3 is easy to identify (5-20µV). There is a significant increase in P3 latency from frontal to parietal electrode sites, whereas the maximum can be found over the midline electrodes and especially Pz. The P3 is basically bilaterally symmetrical (Smith et al., 1990). The latency of the P3 is said to correspond to the speed of cognitive processing and the amplitude is an indicator of the allocation of brain energy resources (Kok, 1997).

The simplest paradigm that evokes a P3 is the so called oddball-paradigm, with frequent standard and infrequent deviant/target stimuli. The task of the subject is to react to the presence of infrequent target/deviant stimulus by a given motor response (e.g. pressing a button), or just by mentally counting these stimuli. Figure 7 (MMN-figure) illustrates a P3 component that typically follows an MMN in an oddball-experiment.

Rosenfeld et al. (1992) investigated the participant's level of interest in video clips by using auditory probes and ERPs. The main task comprised counting silently the number of rare (oddball) target tones among frequent standard tones. Meanwhile, the subjects were watching exciting or boring video clips. During exciting video clips there was a parietal dominant P3 elicited by target tones that was lower in amplitude than during boring video clips.

Recent research suggests that there are two types of P3s in an oddball task (Comerchero & Polich, 1998; Comerchero & Polich, 1999; Katayama & Polich, 1998). Target tones (deviant and infrequent tones requiring a response to indicate their detection) elicit a parietal dominant P3 which peaks at about 350-450 ms after stimulus offset (target P3), when the difference between standard (frequent tones) and target tones is small, and the difference between standard and nontarget deviant tones (differ from standards; no response required) is large. In contrast, deviant tones cause a more anteriorly distributed P3 at about 300-350 ms (deviant P3).

However, the P3 can be differentiated. This differentiation emanates from at least two different, but related scalp-recorded P3 components: the P3a and P3b (Halgren, Marinkovic, & Chauvel, 1998). The P3a occurred whenever there are more than a standard and a deviant/target stimulus (also with a low probability). This component is also known as the novelty-P3, present when a new (beside the standard and target) unexpected stimulus appears which attracts attention and leads to a simple orientation. Therefore, the P3a is considered to be related to the switching of attention to a deviant event and it represents the first analysis of that stimulus (Friedman, Cycowicz, & Gaeta, 2001; Münte et al., 1995). It is supposed to reflect the cognitive processes which identify the stimulus (Squires, Squires, & Hillyard, 1975).

The amplitude of the P3a component is a function of easiness in discrimination from a standard stimulus. The more obvious a deviant is and therefore easier to discriminate from a standard, the larger the P3a (Katayama & Polich, 1998). Thus, the relative perceptual distinctiveness among stimuli affects the P3 (switch of attention) amplitude (Comerchero & Polich, 1999). Doeller et al. (2003) also found an increase of P3a amplitude as a function of deviance (the more deviant, the more distinguishable). This is the reason Comerchero and Polich suggested that P3a may be generated by target-standard discrimination rather than by stimulus novelty as first assumed (Comerchero & Polich, 1998).

This P3a, compared to the P3b (described in the next paragraph), occurs earlier, with a more frontal scalp distribution, and is a response to irrelevant stimuli outside of the focus of attention. Theoretically, the distinction between P3a and P3b emerges because the stimulus context defines the degree of attentional focus. When there is no attention directed to those deviants like in a passive oddball-experiment (no response is required), a P3a (involuntary capture of attention or orienting; passive comparator) is probable, whereas a P3b occurs in a primary discrimination task (conscious attention is needed) (Katayama & Polich, 1998).

The P3b occurs by voluntary detection and encoding of the eliciting event. This component appears more posteriorly with a peak latency 60-80 ms longer than that of the P3a. Therefore, the experimental manipulation determines whether a stimulus elicits a frontocentral (P3a) or a parietal (P3b) or both components (Friedman et al., 2001; Goldstein, Spencer, & Donchin, 2002; Spencer, Dien, & Donchin, 2001), e.g. those above mentioned target P3s in an oddball task contain more P3bs than the deviant P3s, whereas the deviant P3s include more P3as than the target P3s. Source localization revealed different underlying neural generators for both subcomponents of the P3. The frontal lobe including the anterior cingulate was activated in P3a components, whereas the temporo-parietal lobe and the STG (superior temporal gyrus) were responsible for the P3b (Polich, 2004).

Donchin and Coles (1988) see in the later P3 (or P3b) an “updating of working memory”, that revises the representation of an environment steadily. Fabiani et al. (1986) argue too that the P3b reflects the memorization processes. It is assumed that during a series of frequent stimuli the infrequent stimulus template is kept in working memory. Whenever a stimulus matches with this template, then it leads to the termination of previous neural activation due to expectations (Hruby & Marsalek, 2003). From that point of view, the P3 appears whenever an update of this representation of an environment is needed, e.g. when the inner model of an outer environment loses validity by unexpected changes. Thus, the P3 reflects changes in the environmental representations, or better to say the surprise associated with the occurrence of the less frequent stimulus (García-Larrea, Lukaszewicz, & Mauguiére, 1992). Therefore, it is not surprising that the amplitude of the P3 varies with the frequency of occurrence. The more infrequent the stimulus appears, the larger the amplitude that is independent of modality (Naumann et al., 1992). Only the shape and latency (300-500 ms) of the P3 wave differ with modality (Katayama & Polich, 1999). This indicates that the sources generating the P3 wave differ and depend on the stimulus modality (Johnson, 1989).

Furthermore, Donchin et al. (1988) as well as Polich (1996) hold the view that the P3b component mirrors the amount of voluntary and involuntary attention provided for stimulus processing. P3b is not exclusively a measure of selective attention, it is rather a measure of allocation of attentional resources (Donchin & Coles, 1988). Hence, the P3b reflects evaluative categorization of the stimulus (Kayser, Bruder, Tenke, Stewart, & Quitkin, 2000) and the awareness of the subject that an unexpected event has occurred (Leppert, Goodin, & Aminoff, 2003); the infrequent stimulus is consciously detected. For this reason the P3b is also called the endogenous potential (Comerchero & Polich, 1999).

## *Methodologies*

### Probe technique

In general, there are two major ERP techniques to estimate mental workload. The first technique uses a primary task (actual experimental subject's task). In this approach, the state of attention is assessed directly by measuring the P3 amplitude to certain discrete events embedded in the primary task (Kramer, Wickens, & Donchin, 1985; Nittono, Hamada, & Hori, 2003; Novak, Ritter, & Vaughan, 1992). The second technique represents the "probe" technique to define the neural implementation of a task/cognitive process indirectly. With the aid of the probe technique (presenting additional task irrelevant stimuli that had certain features with the actual task stimuli in common) a participant's level of attention or rather the mental workload in auditory materials can be assessed (Papanicolaou & Johnstone, 1984). This latter technique is especially used when there are stimuli in a high frequency like it is the case for continuous speech. A valid analysis of those primary task stimuli would be not given because of overlays of ERP effects from one stimulus to the following. Probe stimuli instead occur less frequent than the task relevant stimuli. Thus, with probes the patterns of regional cerebral activation are assessable in a better way by avoiding overlap effects. Furthermore, task irrelevant probes and task relevant primary stimuli share the same relevant stimulus features. Therefore, probe stimuli are assumed to be processed like task relevant stimuli as well. This fact then allows indirect conclusions from the processing of probes to processing of task relevant stimuli. Moreover, this method assesses cerebral engagement without confounding with stimulus and response-specific activity (Papanicolaou & Johnstone, 1984). Furthermore, the probe technique can be divided into two subtypes: the relevant and irrelevant probe techniques, in which subjects have to ignore or pay attention to probe stimuli.

Several studies have employed reaction times to probe stimuli as an index of the amount of attention allocated during any given task (Basil, 1994; Lang, Newhagen, & Reeves, 1996). Their assessment of the mental workload showed that the more attention was captured by visual material (TV, computer, video-game), the longer the reaction time to a probe stimulus.

In the dichotic listening task (two or four simultaneously presented prose streams; one or two stories in the left or right hearing fields) in the following experiments, probe sounds were presented randomly to both ears. Probe stimuli differed from the prose streams either by location (spatial deviants), by frequency (frequency deviants), only by the spectral composition (white noise deviants) or by none of the physical characteristics (standards). The task of the subject was to listen to one of the presented narratives to be able to answer some questions about this story after each run. Thus, the probes were task irrelevant (irrelevant probe technique) and were presented either on the attended or unattended side. In this way, information responsible for the selection of a certain speaker in this complex auditory scene can be defined and its neural manifestation can be observed. However, the probability of the deviant probes (5% or equiprobable to standard probes) varied from experiment to experiment.

## HRTFs

Usually, sounds in real life appear to come from a particular location in space. In contrast, when listening to sound recording with headphones, sounds appear to originate from inside of our head. In the past few years, there has been a significant increase in interest in the synthesis of three-dimensional spatial sound. In several areas, accurately synthesized spatial sound is of growing importance.

The Head-Related Transfer Function (HRTF) can very faithfully induce perception of spatial location. The HRTF has become central to hearing-aid research, 3-D auditory computer interfaces, and any scientific study of spatial hearing, because it completely characterizes the acoustic information available for sound localization (Blauert, 1997). Thus it represents the standard reference on the psychophysics of three-dimensional hearing.

The aim of the studies in this dissertation was to present auditory stimuli in space in a way that created same hearing impression for all subjects. This would enable a better comparison between subjects, and it would also lead to a more valid and general interpretation of the results. Therefore, the use of HRTFs as a presentation method was essential.

The main attribute of an HRTF is that it captures the position-dependent spectral changes that occur when a sound wave propagates from a sound source to the listener's ear drum. The spectral changes result from diffraction of sound waves by the torso, head, and outer ears or pinnae, all of which vary substantially from person to person. The character of the HRTF also differs enormously between people. These inter-subject variations are often quite significant. For example, serious localization errors can occur when one person hears the source through another person's HRTF. Furthermore, the character of the HRTF depends on the azimuth, elevation, and range from the listener to the source.

Figure 9 shows an example of the HRTF from three persons measured in the same laboratory. These have been adjusted to show a usable signal-to-noise ratio (SNR) of 60 dB. Any spectral difference below this amount would probably be either imperceptible, or out of the overall range of the playback system. In the figures, it may be seen that not only the spectral content changes as a function of position, but also the differences in relative level (dB) between the three subjects are depicted. These are caused by the fact that the pinnae vary in overall shape and size between individuals (Begault, 2000).

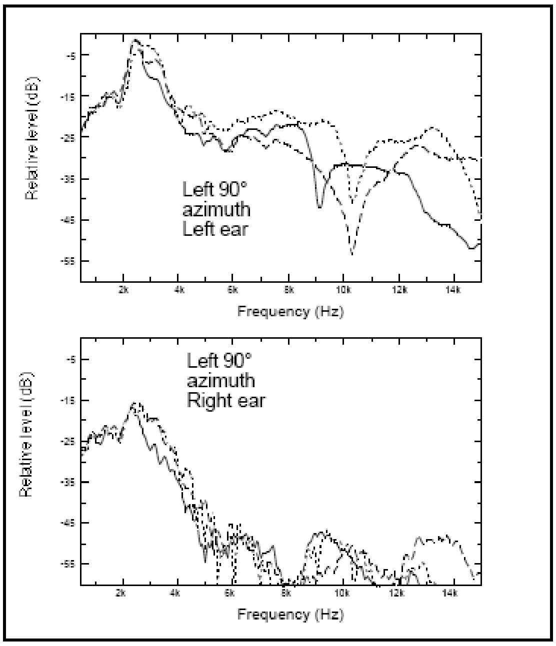


**Figure 9. HRTFs, left and right ear (top and bottom), 90 degrees azimuth, for three different people.** ***Data originally measured by Fred Wightman and Doris Kistler, University of Wisconsin—Madison. (Begault, 2000)***

There are two serious drawbacks for this purely empirical approach: (a) there is no scientific insight into the factors that control spatial hearing is provided, and (b) complex and expensive equipment for applications is required.

My research is based on the belief that the naturally gained HRTF is the most accurate simulation of spatial hearing. This method itself is only an approximation of the real perception. The use of a small microphone in the ear canal modifies the frequency response, position, reflection and refraction of sounds waves, and the air pressure. Those alterations make it impossible to accurately simulate the acoustic environment of the human ear using this method (HRTF measurements). Nevertheless, HRTF generation only attempts to measure effects of external structures and immediate results inside the ear canal. Everything that happens from middle ear to the auditory cortex is not modeled.

No physically- or mathematically-based model can estimate all single parameters in the same way. These parameters cannot be customized to individual listeners by correlation with a small number of properly chosen anthropometric measurements. Decreased localization accuracy has been reported by many psychoacoustic investigations of virtual auditory displays based on processing of non-individual HRTF's (Crispien, Fellbaum, Savidis, & Stephanidis, 1996; Fisher & Freedman, 1968; Wenzel, Wightman, Kistler, & Foster, 1988).

One approach would be to present stimuli by two separate loudspeakers in a certain spatial angle each in front of the subject. Using this approach, each subject would experience the spatial impression of the experimental stimuli. No long-lasting and difficult measurement and computation of individual HRTFs, or efforts to create individual stimuli afterwards, would be necessary. The answer is simple. With two loudspeakers in front of the subjects, the risk would be too high that the subjects would move their eyes, the head or the whole body to one of the loudspeaker from which the to-be-attended stream is presented. ERP (event-related potential) methods depend on a smoothly measured EEG, and those movements would cause large and irremovable artifacts that would weaken any effects. More importantly, in such cases, it would be much easier for the subject to pay attention to that stream, because it would be more intensified in presentation compared to the other stream that would appear more muted. Turning towards one stream in that situation means to turn away from the other one. This would change the hearing impression for both streams – attended and unattended. Effects resulting from that experimental setup could be explained by attentional modulation alone, since a dominance of the presented story compared to the ignored one may contribute.

In the following studies, the application of individual HRTFs to stimuli presented over earphones was possible. Undesirable head movements causing large EEG artifacts could still occur, but they would no longer influence the auditory effect. That the head movements are not useful to the subject would make collected EEG cleaner and ERPs more reliable.

Before creating individual stimuli, an extensive set of head-related transfer function measurements with small head microphones (Sennheiser KE4 Elektret-microphone-cartridges, connected to an amplifier (fabricated at the University of Oldenburg)) embedded in a subject’s ear canal was completed. This measurement was conducted in the Schinkel-Saal Magdeburg (a large concert hall with a high ceiling for fewer reflections). The subject was placed on a swivel chair in the middle of the room. With a fixed loudspeaker (azimuth, 0°), different sound angles were captured by changing the orientation of the subjects. Different angles were marked on the floor and the subjects were asked to orientate with the swivel chair to one of those markers. A Mackie HR 824 Studiomonitor loudspeaker mounted 3.66 meters from the subject presented the sound source (test sound: shift register-noise; 100 ms, but presented 80-

times in a row) for the acquisition of the HRTFs. Maximum length (ML) pseudo-random binary sequences (80 dB SPL) were used to obtain the impulse responses at a sampling rate of 44.1 kHz. The measurements consist of impulse responses from the left and right ear.

The auditory cues that allow a listener to localize the sound are roughly independent on distance, when a sound source is relatively distant from the head. At distances beyond 1 m, the head-related transfer functions (HRTFs) that characterize the relationship between the sound generated by the source and the sounds reaching the left and right ears of the listener vary only by a constant scale factor inversely proportional to the distance of the source. The auditory cues associated with the direction of the source, including interaural time delays (ITDs), the interaural intensity differences (IIDs), and directional filtering by the external ear, are all roughly independent of distance. [See also (Middlebrooks & Green, 1991)]. The situation is dramatically different in the region within 1 m of the listener's head. The ratio of the distance from the source to the near ear divided by the distance from the source to the far ear decreases substantially, when a lateral source is located near the head.

In contrast to the dramatic changes in the IID for nearby sources, the ITD like the HRTF remains roughly independent of distance (D. Brungart & Rabinowitz, 1996; Duda & Martens, 1997; Hershkowitz & Durlach, 1969).

A total of 13 different positions were sampled with a distance of 15 degrees between each other (from -90° (left) to 90° (right)). In sum, 858 impulse responses were recorded for 33 subjects. The measured ML-sequences were transformed by a Hadamard-transformation in a matlab-script resulting in the impulse responses. These impulse responses (IR) consisted of the head-related impulse response (HRIR/HRTF; to avoid confusion, the term HRTF will be used in the following dissertation.) and the electro-acoustic impulse response (EAIR; impulse response of the converter (AD/DA converter; RME Cardbus & Multiface/ambient system). In a reference measurement, the EAIR was captured as well, so that the HRTFs could be determined by unfolding of the IR by the EAIR. Because of a huge amount of floor reflections within the IR, before the transformation could be done, the IRs were cut separately by hand and were centered at their edges. HRTFs corresponding to the various locations were extracted from the impulse responses obtained by the aforementioned unfolding by EAIR.

As this process is time intensive and requires trained technicians and sophisticated equipment, this research was made possible through support by the Hörzentrum Oldenburg (AG Medizinische Physik; Haus des Hörens; Oldenburg).

Each signal/stimulus to be delivered over headphones during the experiments is first filtered with the HRTF. Thus, the perception is of an externalized sound located at the original recording position.

For a presentation of two or four simultaneous stories, first, each narrative is filtered with a HRTF of a certain spatial angle, and thereafter, all two or four stories are mixed together in one stereo-audio-file. This procedure allowed a two- or four-speaker stimulation via earphones instead of using two or four loudspeakers set-up in a hemi-circle around the subject.

## EEG

Electroencephalograms (EEG) are a useful method to investigate neural processes with an excellent (millisecond) temporal resolution. Therefore, EEG has significance for research on neurocognitive functions. Hans Berger, the developer of EEG, first described the alpha-blockade during cognitive processing (“Berger-effect”) as an objective correlate of mental states (Berger, 1929). Since then, methodological improvements accompanied scientific progress.

Whereas the earlier research focused first on the relation between EEG frequencies and behavior, currently there is more interest in the small endogenous event-related potential shifts (ERPs), ranging within a few microvolts in amplitude, that reflect neurocognitive processes corresponding to neural events following a sensory input. Furthermore, special paradigms enable investigations of neural population dynamics (functional integration or segregation between cortical areas during cognitive or perceptual activity).

One possibility for improving the spatial resolution of the EEG itself is to increase the number of electrodes to reduce the volume of tissue that each electrode uniquely averages. Spatial resolution depends on the distance between two electrodes at which the signals are identical apart from additional noise (electrode or amplifier noise). Adjacent channels are observed to be more similar to each other than widely separated electrodes, especially in ERP recordings with dense arrays of 64 electrodes and up (Srinivasan, 2005). According to some researchers, interelectrode distances of around 2–3 cm are required to avoid distortions of the scalp potential distribution (Gevins, Brickett, Costales, Le, & Reutter, 1990; Spitzer, Cohen, Fabrikant, & Hallett, 1989; Srinivasan, Nunez, & Tucker, 1998).

Hence, a dense-array EEG (64 electrodes instead of 32 electrodes in the residual experiments; see difference in Figure 10), was used in experiment three of this dissertation (see chapter 2). It could be shown that 31-channel recordings were clearly not sufficient for adequate source localization, but there is a significant increase of the localization precision from 31 to 63 electrodes, and from 31 to 123 electrodes (Michel et al., 2004), whereas the difference between 63 and 123 channels was minimal. Simulations and experimental studies could clearly indicate that at least 60 equally distributed electrodes are necessary for a correct sample of the scalp electric field and the following source localization procedure. This was the reason why one 64 electrode set up was used for this dissertation as well.

The reader of this dissertation may wonder at this point: Why was the denser-array EEG with 64 electrodes not used for all experiments? Our primary motivation was the economy of fewer electrodes that also led to the use of 64 and not more electrodes. A necessary part of electrode attachment is facial skin and scalp abrasion. It is both time consuming and a demanding task to attach e.g. 64 electrodes on the head when each site needs to be abraded. Therefore, it was thought that the results of one experiment could be taken to confirm or complement the findings of the other similar experiments that differ only in some probe stimuli.

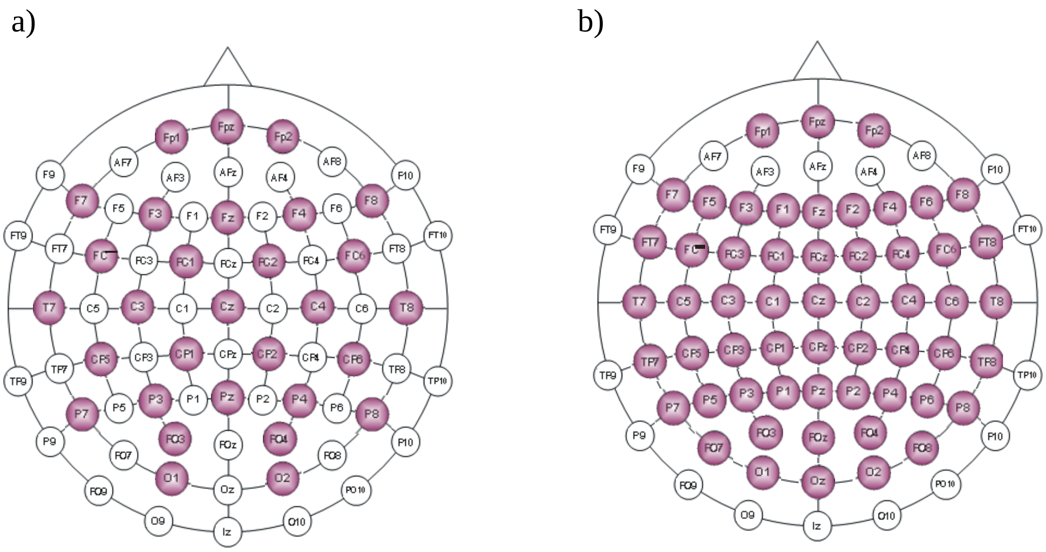


**Figure 10. Electrode position in an electro-cap with 32 (a) and 64 electrodes (b).**

## General procedure

As shown in the next figure (Figure 11), the general experimental design for all of the following experiments consisted of a cross on a computer screen (at a distance of 120 cm) for fixation while two or four simultaneous stories (one or two from either side (right/left)) were presented (see appendix A1 for a list of stories used in each experiment). Superimposed over both prose streams, there were in addition some task-irrelevant and randomly repeated probe stimuli (phonemes, a short noise) with an ISI of 250 - 750 ms. Prose streams as well as all overlaid probe stimuli were convolved with the individual HRTFs determined prior to all experiments. By doing so, all stimuli could be manipulated so as to appear from a certain spatial angle instead of being presented from the right or left only. This procedure therefore enabled a three-dimensional spatial impression for the stories and probe stimuli.

Each experiment consisted of four to six runs in which the subject had to pay attention either to the left or right side and the prose stream presented there respectively. Subjects were told which story/side they should attend to during the upcoming run. The right and left attended runs were uniformly distributed in a random order, counterbalanced in a Latin square design over subjects. In total, each experiment contained 2 (or 3) left and 2 (or 3) right attended runs.

To make sure that the subjects actually paid attention on the right side/story a short interview by oral questionnaire was accomplished. Open questions (to avoid guessing) were asked concerning some facts or descriptions mentioned in the prose stream, for an example, see appendix A2. Additionally, these off-line questionnaires are useful in gaining more insight in how subjects experienced the dichotically and spatially presented stories and whether it was quite difficult to pay attention to one of these or not. If the subjects answered less than 50% of the questions then (s)he was excluded from further analysis for this experiment.

The whole experimental session had a length of 2.25 to 3 hours (including the EEG set-up of about ½ through one hour).

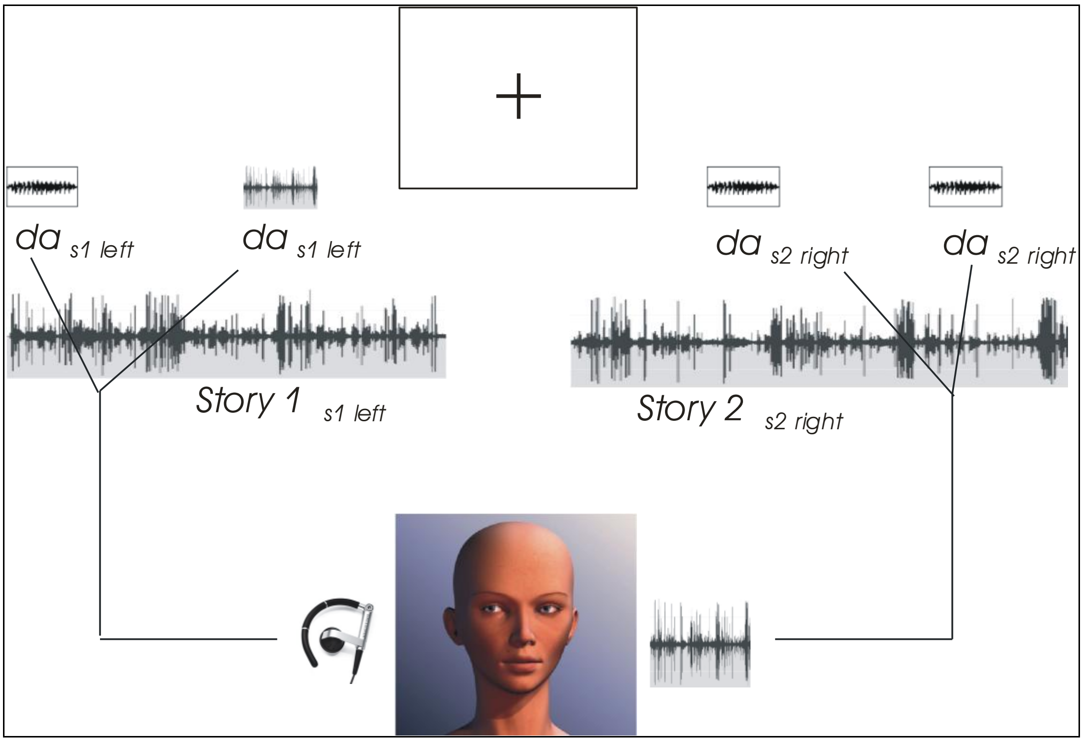


**Figure 11: General design of the experiment: subjects are seated in a comfortable chair with an earphone in each ear and instructed to attend to only one of the two or four simultaneous narratives (presented in either a woman's or a man's voice, but the same gender for both ears within one of the four experiments). The narratives had been convolved with HRTFs for each subject individually, so that they appeared as though they were spoken from a particular angle in space, whereas both narratives had a contrary spatial angle, e.g. 30° left vs. 30° right. A monitor directly in front of the participant showed a cross for fixation to reduce undesired eye movements. Simultaneously, ERPs are recorded to linguistic (/da/; differed either in frequency, spatial angle or were equal in frequency and spatial angle compared to the narrative's speaker on the same side of presentation respectively) and nonlinguistic (noise) probe stimuli (also convolved with the individual HRTFs) superimposed on each narrative.**

## Subjects

Before the actual experiments started, HRTFs of 33 young subjects in total were determined. All subjects were to be employed for all of the four experiments. Therefore, individual HRTFs could be reused for each subject respectively. This reduced the amount of to-be-measured HRTFs and the effort of stimuli preparation.

One subject drop out, after the HRTF measurement and before the actual experiments began, resulted from moving to another town. Hence, 32 subjects (26 women, 6 men; mean age 25.85, range 20-35) took part as planned on all of the four experiments.

The subjects were recruited at the University of Magdeburg. They participated in the experiments after they had given written informed consent. All subjects were healthy and without a history of hearing or neurological disorders. Vision had been corrected if necessary.

## *Structure of the dissertation*

This dissertation will address the following questions: Which mechanisms allow the selection of attention/auditory information in complex surroundings (selection of a specific speaker in a multi-speaker setting)? How is the attentional focus structured? Is it possible to understand a given auditory content by simply directing attention to a certain spatial angle and/or the speaker's frequency?

Neural correlates of pre-attentional and attentional auditory selection defined by the spatial angle and the speaker's frequency or spectral content in a multi-speaker setting should help to clarify these questions. By considering the preceding theoretical discussion and the experiments of this doctoral dissertation, the following hypothesis and outline for this dissertation can be described.

This doctoral dissertation is structured into two parts: two-speaker-settings and a four-speaker-setting. In the first part (chapter 2), three experiments are reported that investigate the selective auditory attention in a setting with two simultaneous speakers each. In concrete, experiment one concerns different fundamental frequencies of a speaker's voice. With this experiment, the question is answerable whether the selection of a speaker is carried out by spatial location only. If so, then the attention effects for both kind of probe stimuli, frequency congruent and frequency incongruent, should be the same (e.g. a clear Nd effect for spatial congruent probes when stimuli from the attended side are compared to stimuli from the unattended ear). Otherwise, if the selection depends on the speaker's frequency, then attention effects for frequency congruent probes should be larger than for frequency incongruent phonemes.

The question of how precise selection, or rather, the focus of selection is, will be addressed in the second experiment. Deviants that slightly differ in a stimulus feature from the speaker's voice characteristic (slight deviant) will be compared to extreme deviants that differ greatly from the speaker's voice in any feature. If the attentional focus is not exactly restricted to the speaker's voice (frequency and spatial location) then it may be possible that slight deviants, due to being more similar to standards (frequency and spatial congruent), are accessed by the attentional focus and therefore are able to pass the selective filter. In this case, slight deviants will be processed like standards and should show similar attention effects. Extreme deviants, on the other side, are clearly different to standard stimuli which may be the reason why they do not pass the filter of selection and therefore will be processed with less attentional resources. Thus, even on the to-be-attended side an attention effect between standards and extreme deviants are expected. Furthermore, because of the low probability of occurrence, both kinds of deviants (slight and extreme) should also show ERP components typically observed in oddball-design (MMN and P3) as a function of deviance – smaller components for slight deviants and larger effects for extreme deviants.

The third experiment investigates the spectral content of auditory information and whether it is essential in the selective process. Therefore, if selection of a speaker results from the particular spectral content of the speaker's voice, then phonemes (speaker's specific frequency spectrum) should cause a more apparent attention effect than white noise (containing all frequencies between 200 – 5000 Hz; broad frequency spectrum). Even on the to-be-attended side, there should be an attention effect (Nd; phonemes-white noise) between phonemes and the white noise, indicating that white noise is not included in the attentional focus.

The second part of this doctoral dissertation (chapter 3) tries to explore whether neural correlates of selection can be provided as well in a more complex situation (setting of four separate and simultaneous speakers). In one experiment, four speakers were positioned in a hemisphere. It was investigated how far the process of probe stimuli depended on the distance to the to-be-attended speaker. One assumption could be that priority for attentional resources will lie on the to-be-attended speaker. The second speaker on the to-be-attended side, but at another spatial location, may get less attention and therefore fewer resources for stimulus processing. Still, compared to the unattended side, it may get noticeably more attention simply due to being on the same side as the to-be-attended stream. For unattended probes on the to-be-unattended ear, there also could be gradation observable, with least attention and least available resources for to-be-unattended stimuli at a different spatial location compared to the to-be-attended prose and its probes. Therefore, the Nd effect should become weaker and weaker with distance and in dependence of the attention's ear.

In chapter four, the main results of chapter 2 and 3 are highlighted and tied together. The experimental results are discussed with respect to their implications for theories of auditory selective attention.

# Spatial location, fundamental frequency and spectral content in a two-speaker setting as critical features for auditory selective attention

## Chapter 2

### Abstract

The ability to maintain a conversation with one person at a noisy cocktail party has often been used to illustrate a general characteristic of auditory selective attention: the listeners' attention is normally directed to a specific set of sounds, to specific sound sources but not to others. The aim of the following experiments was to investigate selective auditory perception in a setting with two human speakers. The location of the different stories told by two human speakers was based on HRTFs (head related transfer functions). Individual HRTFs were applied on prose and probe stimuli to get subject-specific experimental material. ERPs were elicited by task-irrelevant stimuli (probes) during listening to one of the two stories. These probes had different relationships to the speakers' voices. The fundamental frequency, spatial location, or the spectral content was varied in experiment one through experiment three.

ERPs were recorded during auditory selective attention tasks. Listening to one of the two simultaneously presented stories yielded ERP effects related to attention and frequency (experiment one). The MMN and P3 for infrequent spatial and pitch deviants of different degrees in the second experiment provided evidence that a stable memory trace had been developed for the standard stimuli. This pattern observed for the attended and unattended voices shows the precision with which auditory streams are built.

Interestingly, white noise, even though it contains the frequency band of the speaker's voice, seems to be outside of the attentional focus. Relative to the probe phonemes derived from the speaker's voice an attentional effect is seen (experiment three). Overall, the results imply that directing attention to one ear/story leads to a processing inhibition of stimuli in the unattended stream or of stimuli from the attended spatial region that contain different features than the to-be-attended ones. Deviant probe stimuli as distracters are associated with brain activity that suggests an involuntary attentional increase before attention switches back to the to-be-attended stimuli.

### Introduction

Our auditory environment is highly complex. We therefore have to select those aspects, e.g. one of several ongoing speech signals, which are of relevance at any given time. The cocktail-party situation is a good example. Here, we voluntarily focus on one conversation while effectively blocking out concurrent streams of speech. In the visual domain the spatial location of a stimulus can provide efficient information for selection, as location is directly coded throughout this system from the retina upwards. In the auditory domain, selection by spatial location should be more difficult as the location of a sound source is computed on the basis of interaural level and time differences as well as on the basis of the filter characteristics of the outer ear. It is still not fully understood how important information is selected and irrelevant information filtered out. While there is a wealth of research showing that selective auditory attention is associated with an Nd effect in the ERP (Jemel et al., 2003;

Näätänen, 1990), the selection criteria and machanisms that act in natural auditory environments are not very well understood. What are the features determining the attentional focus and its attended stimuli? Do people select by spatial location, by fundamental frequency, or even by the spectral content of an information source? This is the general question addressed in the three experiments of the current chapter. Possible attentional cues are spatial location of presentation, the sound's fundamental frequency, and spectral content. If a stimulus is selected and passes the selective filter more resources are available for its processing. This facilitation in processing is accompanied by an increase of ERP amplitudes (negativities or positivities) (Suzuki, Nittono, & Hori, 2005; Wickens, Kramer, Vanasse, & Donchin, 1983).

There is a large body of ERP research on selective attention in the auditory modality, as described in chapter 1. One common method for these investigations was to use dichotic selective attention tasks. These settings were kept pretty simple and rather unrealistic. If stimulus location was simulated by introducing interaural level or timing differences (ILD, ITD) to create a "virtual auditory space", then interindividual hearing differences, due to different sound reflection depending on a person's upper body and pinnae anatomy, have not been taken into account. Listening to such stimulus series gives an unrealistic location of stimuli, which come from somewhere inside the head. Thus, to be able to present stimuli from a particular angle in space and to account for interindividual anatomical differences, stories and probe stimuli of the present research were convolved with individual head-related transfer functions (HRTFs).

In the following three experiments, ERPs were used to investigate possible criteria for auditory selection processes. Attention was directed to one of two (or more) voice streams. By the introduction of probe stimuli, it is possible not only to study the time course of pre-attentive and attentive processes but also to reveal the nature of the processing of attended compared to unattended stimuli.

The studies make use of the "irrelevant probes technique" (presenting task irrelevant probe stimuli on top of task relevant information; see chapter one), in all three experiments. With this method, a participant's level of attention can be assessed indirectly when the stimulus presentation such as continuous speech is too fast and disables valid ERPs for each stimulus (Papanicolaou & Johnstone, 1984). Instead of risking overlap with task relevant stimuli, the use of less frequent irrelevant probes was employed. Woods et al. (1984) have validated the usefulness of task-irrelevant probes for the investigation of attentional selection with ERPs. They presented probe stimuli embedded dichotically into prose streams. Selective listening to one prose stream elicited an enhanced negativity (starting at around 50-100 ms).

This attention-related negativity was similar to situations, in which series of stimuli similar to the probe were the target of the selective listening process (Näätänen & Alho, 2004). Thus, ERP effects to the probes in the current experiments may also be reliable indicators of attentional indicators. Support for this expectation is also provided by cross-modal studies (e.g., Nager, Estorf, & Münte, 2006) that have shown components like the Nd (negative difference) for task irrelevant modalities. They argued that spatial attention to one modality may facilitate processing to information in another modality at the same location. Therefore, Nager et al. (2006) combined visual and auditory stimuli in space. Subjects were divided into two different groups: "auditory" and "visual". The first group attended either to the left or right ear to find infrequent auditory stimuli (visual stimuli were irrelevant). The reverse was true for the second visual group: auditory stimuli were task irrelevant whereas infrequent visual stimuli should be detected. Results showed that crossmodal spatial attention can modulate the Nd to auditory stimuli when visual stimuli were attended as well as P1 and N1 components to visual stimuli when auditory stimuli were attended.

In the following experiments, phonemes and a white noise burst were employed as irrelevant stimuli superimposed on two simultaneous and dichotic prose streams. Just one of these concurrent streams had to be attended to while the other one could simply be ignored. The goal of the present study was to extend earlier efforts at evaluating how directed attention affects the ERPs elicited by irrelevant auditory probes on either the attended or the unattended prose stream in a spatial setting. The ultimate objective of this research is to investigate the structure of the focus of attention.

## *Experiment 2.1:*

### Method

This first study was conducted to investigate the validity of a new method of stimulus presentation. The use of individual HRTFs combined with the probe technique was considered to be a reliable tool to simulate spatial listening via head phones similar to a free-field situation.

Auditory selective attention has often been studied with ERPs. Most of the previous studies used simple dichotic listening tasks to present simultaneous auditory stimuli. Thus, the stimuli occurred from right and left with 180° spatial separation. In nature, however, not every sound appears from that direction. Therefore, some authors have used techniques such as interaural time differences (ITD) and interaural level differences (ILD) (Darwin & Hukin, 1999; Shinn-Cunningham & Ihlefeld, 2004) to spatially position their stimuli. Thereby, they implemented the shadowing effect of the listeners head: sounds are perceived louder on the ear on the same side as the source and more muted in the opposite ear, because the head attenuates a good amount of the sound's intensity, resulting in interaural intensity differences. Another characteristic between ears is that, because of the head's size, sounds coming from an angle (except from the azimuth) will not reach both ears at the same time. The traveling of sound waves take a bit longer to one ear resulting in a time delay between one ear and the other (ITD, interaural time difference). For example, a difference in distance of 10 cm results in a delay of 300 µs. The brain uses both types of information.

Interestingly, level differences (lateral superior olive, LSO) and timing differences (medial superior olive, MSO) are processed by different brain stem structures. Also, differences in level are more important for high frequency (above 3 KHz) sounds, while differences in timing are used for the localization of lower frequency sounds. Darwin et al. (1999) as well as Shinn-Cunningham et al. (2004) manipulated the ITD and ILD in order to position their stimuli at a certain angle in space. Another method that includes the characteristic ITD and ILD implicitly without a mathematical calculation is the "Kunstkopf" technique. Kunstkopf (dummy-head) is a stereophonic recording system developed in Germany. It utilizes an artificial head sitting on a resonator similar to a chest cavity. Recording microphones are installed inside an anatomically correct ear canal within the artificial head. Therefore, the sounds are recorded exactly at the point where the human eardrum would be located. Even the artificial pinnae are carefully designed for accurate reflection of the incoming sound. This assures the good front-back and height-depth perspective for which the technique is noted. Hence, sounds recorded by this technique are later on perceived as externalized by using lightweight headphones for optimum listening. Obviously, individual differences in the size and shape of the outer ears and the head are not taken into account by this technique. Thus, the localization precision that can be achieved with this kind of technique is somewhat suboptimal.

To date, to my knowledge, no study has used the method of individual HRTFs in humans to present auditory stimuli in space and to investigate selective auditory attention. Therefore, the goal of the following experiment was to replicate auditory attention effects by using individualized HRTFs (see chapter 1 for a description of how individual HRTFs were obtained) and to examine what the brain's response to a deviant pattern (higher fundamental frequency) compared to a to-be-attended stimulus would look like.

## Subjects

Thirty-two young subjects (26 women, mean age 25.85, range 20-35) were recruited, all students at the University of Magdeburg. They participated in the experiment after they had given written informed consent prior to participation. They received a small sum for participation. All subjects were healthy, had normal or corrected to normal vision and normal hearing.

## Stimuli

To reduce eye movements, subjects fixated on a cross (1.5 degrees visual angle) on a computer screen located in front of them during the recording. The experimental set-up comprised the presentation of two different stories at the same time, spoken by two native speakers of German of the same gender (see appendix A1).

The term "speaker" in the present investigation refers to a person's voice only from now on. Due to stimulus presentation via headphones, no loudspeaker came into play. Therefore, if loudspeakers were used in other previous studies that the present results are compared with, the term "loudspeaker" will be mentioned explicitly.

We did not use the same speaker's voice for both stories; otherwise it would have been too difficult to differentiate between the two narratives during selection. On the other side, to yield effects that are not affected by any dominance in a speaker's voice, different distinguishable human speakers with quite similar characteristics in volume and intonation were chosen. Both stories were convolved with individualized HRTFs (measured and filtered beforehand for each subject separately; see Wightman (1989a; 1989b) and were presented simultaneously and continuously via headphones at virtual locations coming from 30° degrees to the left and 30° degrees to the right. For each run, subjects were instructed which story/ear they should pay attention to. Additionally, a phoneme ("da"; 100 ms; also convolved with individualized HRTFs) of either the same frequency as the speaker's voice (F+) or about 400 Hz higher in fundamental frequency (F+400) was presented as a probe stimulus at the location of either the attended or unattended story. Phonemes were presented in randomized order with an interstimulus interval (ISI) of 250 to 750 ms (rectangular distribution).

For this experiment, 4000 probe stimuli (500 probes for each of the eight conditions; see Table 1) were presented in 4 runs (2 attend right, 2 attend left) lasting about 11 min each. The subject's task consisted of actively listening to one of the two parallel stories as instructed before a run started, and to keep the content of that story in their mind. No button press was required. After each run the subjects were asked several questions about the attended story to make sure that they really directed their attention to that story only (for an example of such a questionnaire see appendix A2). The whole session had a length of 1.25 hours (excluding the EEG set-up (about 30 minutes)).

**Table 1: Distribution of probe stimuli within the eight experimental conditions. F+ represents frequency-congruent phonemes, and F+400 conforms to frequency-incongruent probe stimuli. A+ stands for "attended" and A- for "unattended".**

| | **left** | | **Right** | |
|---|---|---|---|---|
| | ***A+*** | ***A-*** | ***A+*** | ***A-*** |
| **F+** | 500 | 500 | 500 | 500 |
| **F+400** | 500 | 500 | 500 | 500 |

## EEG-Recording and data analysis

The EEG was recorded by using an elastic cap with integrated tin electrodes (positions: Fp1, Fp2, F3, F4, F7, F8, C3, C4, P3, P4, Fpz Fz, Cz, Pz, T7, T8, Fc5, Fc6, Fc1, Fc2, Cp5, Cp6, P7, P8, P3, P4, Po1, Po2, O1, O2 of the international 10-20 system (Jasper, 1958)). The horizontal/vertical electrooculogram (EOG) was recorded using a bipolar montage between the left external canthus and a position located below the left eye. The EOG was registered to allow off-line rejection of ocular artifacts. All scalp electrodes were referenced to the left mastoid electrode. The EEG was amplified (time-constant 10 s, low pass filters 30 Hz, high pass filter .05 Hz), digitized on-line with 4 ms resolution (sampling rate of 250 Hz) and stored for further processing on hard disk. After off-line artifact rejection that excluded trials contaminated with ocular and other artifacts using individualized amplitude criteria, ERPs were obtained for epochs of 1024 ms including a 100 ms interval before the onset of the stimulus used as baseline.

The ERPs were averaged separately for attention condition (attended or unattended probes), frequency (F+/F+400), and location (left/right). After preliminary analyses had indicated no difference between effects to left and right-sided stimuli, ERPs to left and right-sided probe stimuli were collapsed to yield waveforms for electrode positions ipsi- and contralateral with regard to the location of the probes. The ERPs were generally quantified by mean amplitude measures, in some cases by peak amplitude or local peak latency measures (mentioned separately in the text), and the resulting data were subjected to repeated measures analyses of variance (ANOVAs). First, the P1, N1, positive difference, Nd, PN and P3 amplitude data were submitted to repeated measures ANOVA with factors Condition (attended, unattended, frequency-congruent, frequency-incongruent), and Electrode(s) (selected electrode(s) for spatially restricted effects), or if possible with additional factors Laterality (left or right), and Anterior-posterior (anterior or posterior). For all statistical effects (univariate F-tests) involving two or more degrees of freedom in the numerator, the Huynh-Feldt correction was applied on the data to correct for possible violations of the sphericity assumption (Huynh & Feldt, 1970).

In order to look at the scalp distribution, difference waves were computed to isolate electrophysiological correlates of attentional selection effects. These were used to create spline-interpolated isovoltage maps employing the BESA software package (Scherg & Berg, 1991).

## Results

### Subjects

The results of twenty-four subjects (20 women, 4 men; mean age 26.6, range 20-35) out of 32 were included in statistical analyses. The data of the remaining 8 subjects had to be discarded, because of too many artifacts or technical failures. If more than one third of epochs were rejected due to artifacts (exclusion criteria), a person was excluded from further analyses.

### Behavioral results

According to the questionnaire, it was concluded that all subjects were able to concentrate on the specified story. On average, 79% of the questions were answered correctly. The percentage of correct answers was above chance (more than 50% correct answers) for all subjects and varied from 54% to 98% (appendix A3a).

### Electrophysiology

The ERP figures in this section display group average ERPs (n = 24) and difference waves for either attended location, frequency-congruent probes (A+ F+), attended location, frequency-incongruent probes (A+ F+400), unattended location, frequency-congruent stimuli (A- F+), or unattended location, frequency-incongruent probe stimuli (A- F+400). ERPs from left and right stimuli were collapsed to yield ipsi- (i) and contra-lateral (c) sites.

Mean amplitudes were obtained in several time-windows separately for each subject, condition and electrode site. These were entered into overall (30 electrodes) or separate regional repeated measures analyses of variance with factors condition (2 levels) and Electrode(s), or if possible with factors Laterality (left, right), and Anterior-posterior instead of the Electrode factor (significance set to p = .05). Detailed attention and frequency effects are shown in the following paragraphs.

## *Attention effects*

### 1) Frequency-congruent stimuli

Contrary to expectation of a fronto-central Nd effect, Figure 12 shows an anterior positivity for attended probe stimuli, whereas unattended probes were negative between 300 and 400 ms. This positive difference (termed Pd from now on to distinguish this effect from the Nd; see difference waves) was largest at frontal sites, and seemed to be larger on the ipsi- compared to contralateral side. Difference waves were obtained by subtracting unattended F+ from attended F+ ERPs. This Pd was followed by a negative difference (Nd) between attended and unattended F+ stimuli in the time window of about 390-530 ms, which was most pronounced on posterior sites.

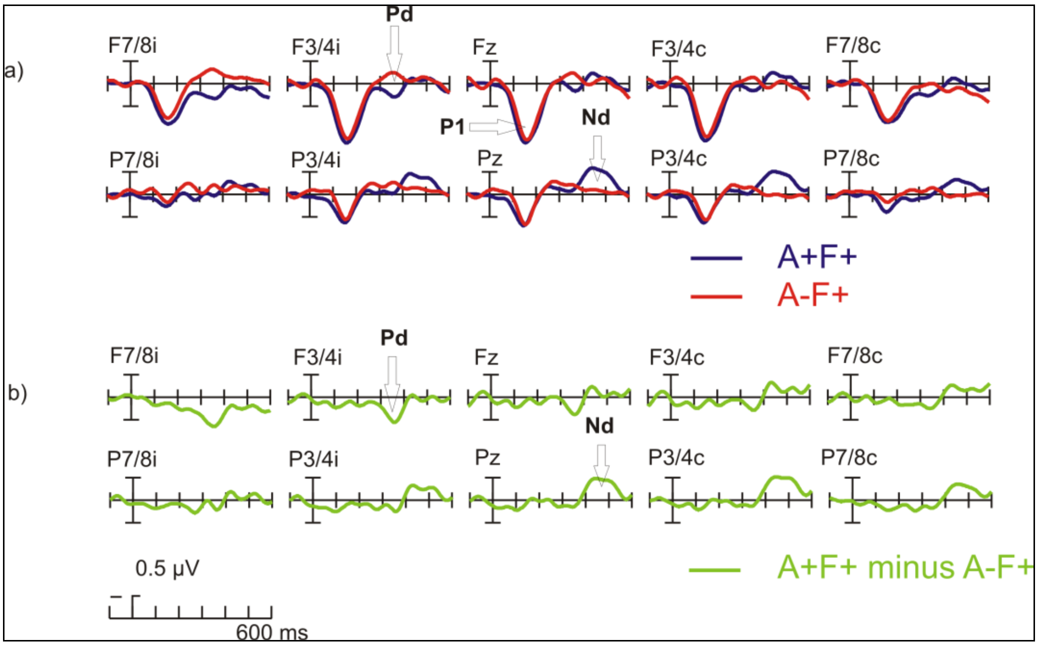


**Figure 12:** a) **Group average ERPs (blue and red curves) for frequency-congruent auditory probe stimuli (F+) on the attended (A+) and unattended (A-) side respectively for frontal electrodes, whereas ipsi- (i) and contra-lateral (c) sites were accounted for both, right and left probe stimuli combined.** b) **The green lines show difference waves (attended minus unattended stimuli) for the same electrodes displaying the frontal Pd (positive difference) and the posterior Nd (negative difference) effect.**

This first comparison between attended and unattended frequency-congruent probe stimuli (A+ F+ and A- F+) yielded main effects of Condition (Attention) in the following time windows: 95-125 ms (P1 ascent), 250-390 ms (Pd) and 390-530 ms (Nd effect) after stimulus onset.

The P1 peak itself did not differ between attended and unattended F+ phonemes but the ascent of this component seemed to rise differently. The ANOVA for the time window 95-125 ms after stimulus onset (P1 ascent) yielded a significant main effect of Condition (all head electrodes: F(1,23) = 4.79, pHF = 0.0391). Unattended F+ probes rose somewhat later than the same stimuli when they were attended, but peaked around the same time.

The Pd ANOVA especially for temporal electrodes (F7/8, T7/8, P7/8) obtained a broad significant main effect of Condition (F(1,23) = 9.16, pHF = 0.0060). A- F+ probes elicited a more negative ERP waveform compared to attended frequency-congruent phonemes. Furthermore, a two-way interaction for these temporal electrodes was found (condition x anterior-posterior interaction; F(2,46) = 3.88, pHF = 0.0284), confirming a predominantly fronto-central scalp distribution for the Pd effect. Additionally, this effect was more pronounced in the first than in the second half of the experiment. The Pd amplitude in both halves differed significantly (F(1,23) = 4.62, pHF = 0.0423; temporal sites).

With regard to the Nd, a main effect of Condition for parasagittal electrodes on the contralateral side (Fp1/2c, F3/4c, C3/4c, P3/4c, O1/2c; F(1,23) = 4.73, pHF = 0.0402) and for parietal electrodes (P7/8, P3/4, Pz: F(1,23) = 4.45, pHF = 0.0460) was revealed. Moreover, the parasagittal (F(1,23) = 4.48, pHF = 0.0452 (Fp1/2, F3/4, C3/4, P3/4, O1/2)) ANOVA reached significance in the condition x laterality interaction, confirming a predominance of this Nd over parietal and especially on the contralateral side. Thus, both conditions (A+F+ and A-F+) did show an Nd with a more negative-going waveform for attended frequency-congruent stimuli compared to the same probes when they were unattended.

Isovoltage maps of the difference waveforms (A+F+ minus A-F+) illustrate the difference in scalp distribution for the positive difference (Pd) and the later negative difference (Nd). The earlier Pd (250-390 ms) shows a fronto-central and highly ipsilateral distribution, whereas the Nd (390-530 ms) appears more posterior (parietal) and contralateral to stimulus presentation (Figure Figure 13).

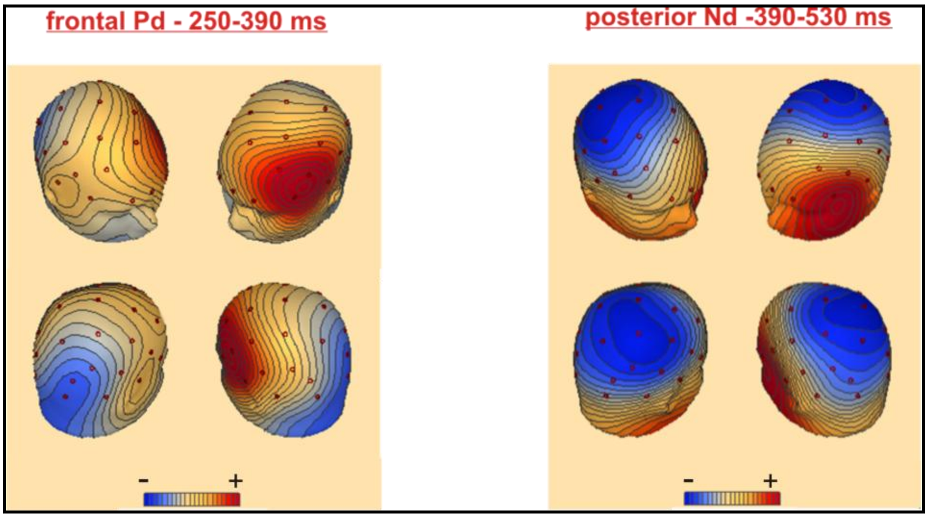


**Figure 13: Isovoltage maps of the positive difference and Nd effect: difference waveforms (A+F+ minus A-F+) were used. Isovoltage spline interpolation for the 250–390 ms and the 390-530 ms intervals were used (BESA).**

2) Frequency-incongruent stimuli

Next, analysis was performed for frequency-incongruent stimuli on the attended and unattended side respectively (Figure 14). The ERPs show differences in two main components. First, the N1 between 150 and 230 ms was different between both conditions. It was more negative for A+ F+400 than for A- F+400 probes and slightly more apparent for ipsilateral electrodes. The second time-window that showed differeces between attended and unattended F+400 probes was between 240 and 290 ms. Therein, the A+ F+400 condition showed a greater positivity compared to A- F+400 probe stimuli which was more prominent on the contralateral side.

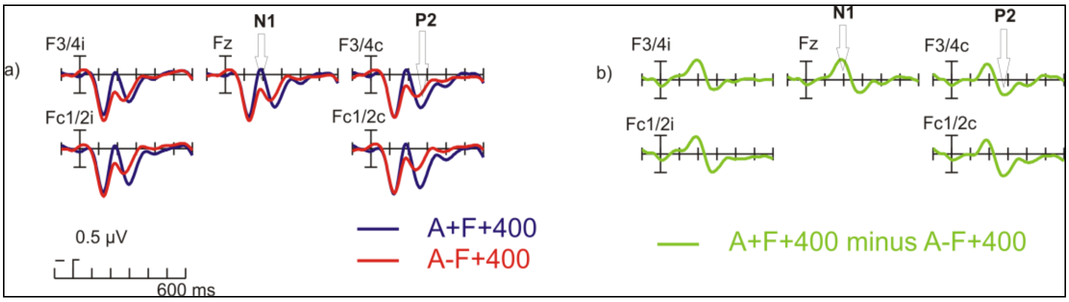


**Figure 14:** a) **Group average ERPs for frequency-incongruent auditory probe stimuli (F+400) on the attended (A+) and unattended (A-) side respectively for frontal electrodes in an ipsi- (i) and contra-lateral (c) order.** b) **Ddifference waves (attended minus unattended stimuli) display the N1 and the P2 as well.**

The earliest difference between A+ F+400 and A- F+400 probe stimuli was found for the N1 component (150-230 ms). The ANOVA showed a significant main effect of Condition with a maximum at frontal electrodes (Fp1/2, F3/4, F7/8, Fc1/2, Fc5/6: F(1,23) = 4.95, pHF = 0.0361) whereas posteriorly no statistical N1 difference was observed (pHF > 0.4) indicating a frontal distribution ipsilateral to stimulus presentation with a more negative-going ERP for the A+F+400 compared to A-F+400 probes. The parasagittal (Fp1/2, F3/4, C3/4, P3/4, O1/2) as well as temporal (F7/8, T7/8, P7/8) analysis confirmed the ipsilateral shift by a significant condition x laterality interaction (parasagittal: F(1,23) = 11.26, pHF = 0.0027; temporal: F(1,23) = 18.17, pHF = 0.0003).

Statistical analysis in the 240-290 ms time-window revealed significant results between both conditions. The maximal effect was observed at central and centro-parietal electrodes (C3/4, Cz, Cp1/2, Cp5/6: main effect of Condition: (1,23) = 8.50, pHF = 0.0078) confirming a more positive ERP for A+ F+400 than for the same stimuli when they were unattended. A lateralized effect was indicated by a condition x anterior-posterior interaction for the temporal (F7/8, T7/8, P7/8: F(2,46) = 3.44, pHF = 0.0407) analysis.

Isovoltage maps illustrate the difference in scalp distribution of the N1 and P2 components (Figure 15). The N1 has its maximum frontal to frontopolar, ipsilateral to stimulus presentation. The P2 shows a more central to centro-parietal distribution.

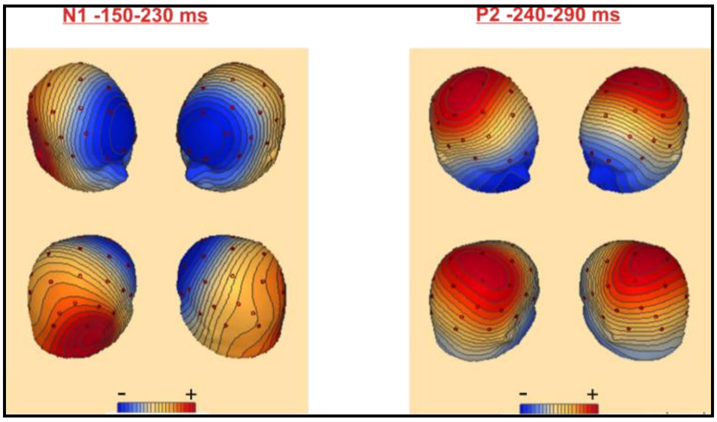


**Figure 15: Isovoltage maps of the topographical distribution of the N1 and P2: difference waveforms (A+F+400 minus A-F+400). Isovoltage spline interpolation intervals were used (BESA).**

## *Attentional effects – frequency-congruent vs. incongruent*

### 1) Stimuli on the attended side

When the two different stimuli (F+ and F+400) presented on the to-be-attended side were compared (Figure 16), an earlier P1 peak for frequency-incongruent probes compared to frequency-congruent stimuli was observed which did not differ much in amplitude. A negative shift for frequency-incongruent probes followed at around 130-210 ms (N1), more prominent over contralateral electrodes at frontal sites whereas more posteriorly (not seen in the figure) this effect was slightly larger ipsilateral to presentation. Later on (about 240-500 ms), a positivity for frequency-incongruent (F+400) compared to congruent (F+) probes followed the N1; paramount on the contralateral side as well. Both components were better seen in difference waves between frequency-congruent and incongruent stimuli (Figure 16b).

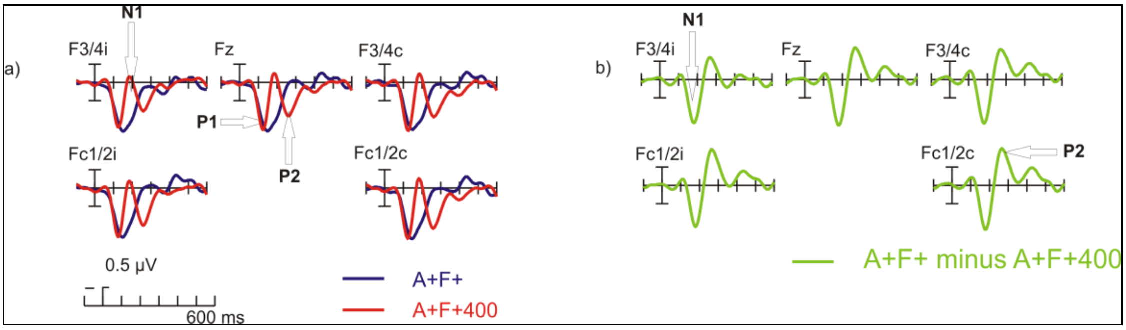


**Figure 16:** a) **Group average ERPs for frequency-congruent and incongruent auditory probe stimuli on the attended side in an ipsi- (i) and contra-lateral (c) order.** b) **Difference waves (A+F+ minus A+f+400) are shown as well displaying the N1 and the P2. Note that the N1 difference wave has a positive peak and the P2 a negative peak caused by subtracting frequency-incongruent from congruent ERP.**

Three ERP effects were observed in the comparison between attended F+ and F+400 probe stimuli: a P1, N1 (MMN-like component) and P2 effect. The first component observed for both conditions was the P1 (80-200 ms). The mean amplitude did not differ significantly. The slight difference in peak latency is the result of the N1 overlay in frequency-incongruent ERPs and not a latency effect of the P1.

In the N1 latency window (130-210 ms), the ANOVA showed a significant effect of condition (overall: F(1,23) = 8.47, pHF = 0.0079) indicating that attended frequency-incongruent stimuli (A+ F+400) elicited a negativity compared to frequency-congruent probes (A+ F+) on the to-be-attended side.

The third distinction between frequency-congruent and incongruent probes on the to-be-attended side was represented by the P2 (240-310 ms). The overall main effect of Condition (F(1,23) = 8.31, pHF = 0.0084) became significant showing a more negative-going waveform for attended F+ phonemes than for attended F+400 probes. Analysis of midline electrodes (Fpz, Fz, Cz, Pz) yielded a significant condition x anterior-posterior interaction (F(3,69) = 3.54, pHF = 0.0388) indicating a prominent central effect compared to frontopolar and posterior electrodes.

Isovoltage maps (Figure 17) show the scalp distribution for the N1 (positive, because frequency-incongruent probes were subtracted from frequency-congruent phonemes) and the later positivity (P2). The N1 shows the characteristic fronto-central distribution whereas the P3 appears more central.

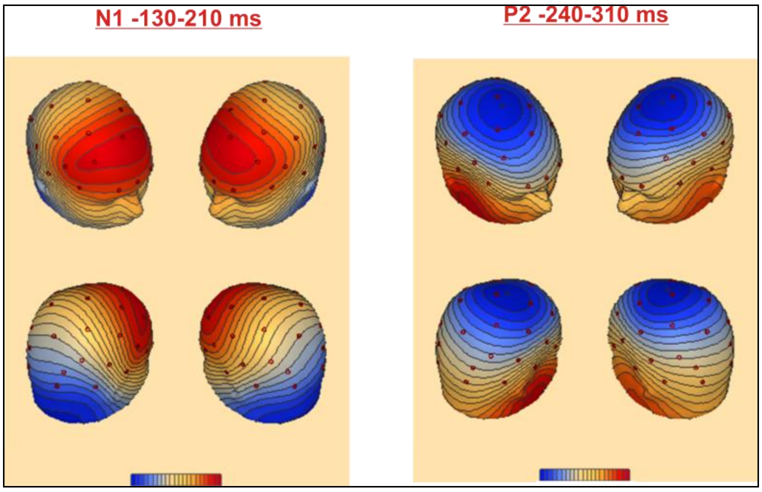


**Figure 17: Isovoltage maps of the ERP components N1 and P2: difference waveforms (A+F+ minus A+F+400) are shown within the accordant timeframes. Isovoltage spline interpolation for the 130–210 ms and the 240-310 ms intervals were used (BESA).**

2) Stimuli on the unattended side

Difference effects between frequency-congruent and incongruent probe stimuli were also observed on the attended side still remained on the unattended side (Figure 18). An early effect (80-150 ms) between both kinds of unattended stimuli (F+ and F+400) was found, indicating an earlier and a slightly larger P1 for F+400 stimuli compared to F+ probes. As seen on the attended side, a negative shift for unattended frequency-incongruent probes at around 160-200 ms (N1) was observed; more prominent over contralateral electrodes. In the time range of 220-370 ms after stimulus onset, a P2 between A- F+ and A- F+400 was revealed, that was slightly larger on the ipsilateral side. The difference waves between frequency-congruent and incongruent stimuli on the unattended ear could point up those effects (Figure 18b).

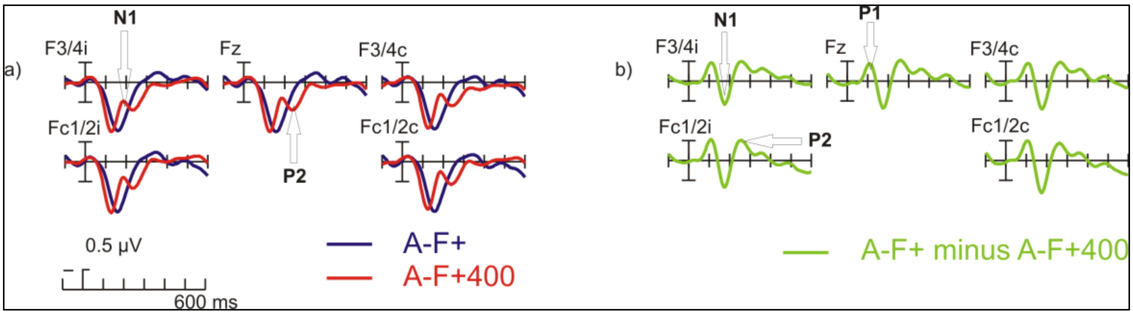


**Figure 18:** a) **Group average ERPs for frequency-congruent and incongruent auditory probe stimuli on the unattended side for frontal electrodes in an ipsi- (i) and contra-lateral (c) order.** b) **Difference waves (unattended frequency-congruent minus unattended incongruent probes) are shown as well displaying the P1 modulation, the N1 and P2. Note that the P1, and P2 difference wave has a negative and the N1 a positive peak caused by subtracting frequency-incongruent from congruent ERP.**

Statistical analyses between frequency-congruent (F+) and incongruent (F+400) probes on the to-be-ignored side revealed as well three separate differences (Table 2).

**Table 2. Significant results (main effect of condition) of pair-wise ANOVAs between frequency congruent and incongruent stimuli on the unattended side.**

| components | difference in | time window | max. effect | F-value (df) | p-value (pHF) |
|---|---|---|---|---|---|
| **P1** | peak latency | 80-200 ms | overall | 37.12 (1,23) | 0.0001 |
| **N1** | mean amplitude | 160-200 ms | Fronto-central (Fc1/2, Fc5/6) | 15.85 (1,23) | 0.0006 |
| **P2** | mean amplitude | 220-370 ms | Frontal (Fp1/2, F3/4, F7/8, Fc1/2, Fc5/6) | 5.54 (1,23) | 0.0275 |

## Discussion

The current study was conducted to investigate the validity of a new method of stimulus presentation for the investigation of auditory selective attention. The use of individual HRTFs combined with the probe technique was considered to be a reliable tool to simulate spatial listening via head phones similar to a free-field situation. Therefore, replication of attentional effects from previous auditory ERP studies was the main aim of the present study. In other words, modulations of ERP components elicited by probe stimuli in an attended situation compared to an unattended condition were expected (Papanicolaou & Johnstone, 1984; D. L. Woods et al., 1984). Attending to one of the two simultaneous stories in the current investigation was thought to increase the tonic neural activity in the sensory areas responsible for processing of the incoming stimulus. The reason for this tonic increase might be ‘bias signals’ from higher attentional control areas (see review of Hillyard et al. (1998)). However, the superimposed probe stimuli were of phasic nature. This characteristic of probe stimuli was used to analyze attention effects. By averaging the brain waves whenever those probe stimuli occurred, the tonic activity elicited by the stories was canceled out and only the specific and phasic modulation by the probes is left. As the results then showed, there were ERP differences between attended and unattended phonemes (probes).

Thus, the application of HRTFs on auditory stimuli for spatial simulation is a valid procedure. The ability of the subjects to answer questions regarding the attended story confirmed that subjects had indeed directed their attention to one speaker. Any ERP differences to probes related to the attended speaker relative to probes related to the unattended speaker can thus be taken as electrophysiological markers of attention (Teder et al., 1993). As will be pointed out below, the attention-related ERP differences show some similarities and some differences with previous ERP results obtained in spatial auditory attention tasks (e.g., Alho, Donauer et al., 1987; Schröger, 1994; Schwent, Snyder, &

Hillyard, 1976; Teder et al., 1993). It has to be kept in mind, however, that the current set-up provided a more realistic and less artificial experimental setting by a virtual auditory space than simple dichotic listening tasks (Begault, Wenzel, & Anderson, 2001; Carlille & Pralong, 1994; Pralong & Carlille, 1994). In most of these experiments, sounds presented via headphones are typically perceived inside the head (internalized), unlike real sound sources that are perceived outside of the head (externalized). With HRTFs, one can reproduce a real sound source more precisely using headphones, and auditory images are then appropriately externalized and localized (Kulkarni & Colburn, 1998).

The latency of the observed attention-related modulations conforms to previously described attention effects on auditory ERPs. These modulations provide evidence that neural resources were allocated to attended and unattended stimuli to a different extent. In other words, attended stimuli were processed preferentially and more elaborated than unattended ones. Nevertheless, after detecting their deviant feature(s) from the to-be-attended ones, they were basically processed as rather unattended stimuli even though they occurred at the to-be-attended side. Thus, the fundamental frequency can determine whether a stimulus belongs to a certain to-be-attended auditory object or a different one that may be neglected. The latter case, thereupon, would lead to exclusion from the actual attentional focus, and to a reduced stimulus processing. This procedure ensures sufficient resources for the process of relevant to-be-attended stimuli.

The following paragraphs go into detail regarding attention and frequency effects mentioned above.

## Attention effects

### *Frequency-congruent stimuli*

The main comparison in this study regarded attended and unattended frequency-congruent probe stimuli (F+) and revealed a negative difference. However, instead of the expected pure Nd, there was a preceding significant positivity deflection with a more positive A+F+ than A-F+ waveform. The reason for the occurrence of this component is unclear and rather speculative at the moment. One speculation would be that the Pd may reflect a rejection process, an active inhibition or suppression of the irrelevant stimulus (probes) for a better focus on relevant ones (story).

A broad positivity in addition to the classical Nd in auditory experiments was reported only in a single investigation yet. Beer and Röder (2005) presented in one condition of their experiment unimodal auditory stimuli with the task to detect infrequent deviants in a sequence of tones. Despite this oddball paradigm in contrast to our equiprobable stimulus presentation, they observed a broad positive difference wave (Pd: positive difference) at 120 – 160 ms which was then followed by the classical Nd. In their cross-modal condition (auditory-visual) this Pd did not reach significance.

However, the Nd's onset then would indicate the time at which a particular feature or object was discriminated and selectively processed according to its task relevance. But why did this Nd component not occur frontally as well, but instead at centro-parietal electrodes? It could very well be that the earlier positive difference overlays the classical Nd effect. Thus, only at posterior electrodes where the Pd influence weakened again, the Nd component may have been able to be evoked before the next probe stimulus was presented. The ISI was relatively short with 250-750 ms, so that the Nd did not have enough time to develop completely. Another explanation could be derived from Hansen and Hillyard. According to them (Hansen & Hillyard, 1988) the Nd effect would increase from trial to trial because of an establishing of a respective memory trace. Applied to the present study, this could mean that subjects were not able to create a proper memory trace of the presented probe stimuli because of the extreme working memory overload by simultaneously presented stories so that the Nd at frontal electrodes was not strong enough which finally caused a suppression or overlay by the Pd. By the same reason of a Pd influence, the Nd onset is later than usually seen in previous studies (about 100 ms or even earlier (D. L. Woods & Alain, 1993) instead of the current 390 ms onset time). Nevertheless, the nature of this Pd in a selective listening task with different simultaneous auditory streams like in the present study is not fully understood yet and therefore needs further detailed investigation.

Only one question remains: Why was Pd found in the present investigation compared to other auditory stimulations with continuous speech and superimposed task-irrelevant probe stimuli (for example, Teder et al., 1993)? A conclusive answer cannot be given yet. The occurrence of the Pd may due to task characteristics, as Beer and Röder (2004) suggested, because in their previous study with the same auditory stimuli no such Pd was found. Nevertheless, any interpretations are only speculative so far. From the current experiment a new conclusion following that of Beer and Röder would be that differences in task characteristics are not sufficient to evoke a Pd. Additionally, with regard to the more working memory-demanding task in the present investigation, because of the same speaker's gender and similar intonation characteristics compared to other selective listening task involving real speech streams, a logical conclusion would be that the working memory load in a certain task could be the decisive factor eliciting a positive difference. As soon as two modalities are important as in Beer and Röder's research, and the context becomes more complex or complicated (spatially close presented stories with the same speaker's gender in the current investigation) working memory is used more to its capacity limit. In general, previous auditory selective listening tasks were basically quite simple and less demanding regarding working memory, which could be the reason why no Pd for this modality was found therefore. Nevertheless, further research needs to be done in this area to verify or falsify this assumption.

*Frequency-incongruent stimuli*

When A+F+400 were compared with A-F+400, then a larger amplitude of this negativity with a more frontal distribution was found for the attended frequency-incongruent probes. This could suggest that attention may have modulated the N1 occurrence in terms of a more precise registration of attended than unattended F+400 probes. There were suggestions that the auditory N1 attention effect is associated with a relative increase of excitability of neural networks coding the attended as compared to unattended stimuli (e.g. M. Giard et al., 1988; Hillyard et al., 1973; Woldorff et al., 1993; D. Woods et al., 1994).

This pattern of an enhanced component for attended compared to unattended F+400 probes was also observed for the P2 component following the N1. The P2 was significantly larger for A+ versus A- stimuli proposing the same attentional modulation process that provides more resources or neural excitability potential for attended stimuli (M. Giard et al., 1988; Hillyard et al., 1973; Woldorff et al., 1993; D. Woods et al., 1994).

**Frequency effects**

As frequency-congruent and frequency-incongruent phonemes on both sides (A+ and A-) were analyzed, it was found that the N1 was elicited only by F+400 stimuli relative to F+ probes whether they were attended to or not. The reason why the F+400 probes caused an N1 compared to F+ stimuli could be the extreme difference in fundamental frequency (400Hz). Continuously presented stories induced an excitation of neurons repeatedly responsible for action potentials with a following refractory period. The refractory period is a critical control mechanism to avoid hyperactivity, by preventing subsequent stimuli eliciting action potentials. In an absolute refractory period, no further stimulus is able to evoke another action potential. Shortly afterwards, there is a higher excitation threshold in the relative refractory period. Additional probe stimuli with the same features as the stories, thus low depolarization below threshold, might fall sometimes disregarded into this refractory period; unable to pass the higher threshold and to initiate an action potential. Therefore, only strong and very obvious stimuli such as the frequency-incongruent probes in the present study would be able to terminate the refractory period for another action potential. This intense excitation might have been responsible for the observed N1 F+400 probes in contrast to frequency-congruent phonemes. Thus, frequency-incongruent phonemes represent a kind of deviant stimulus.

Therefore, the N1 was more like the MMN of an oddball experiment which would be evoked by deviant stimuli (Picton et al., 2000).

This N1/MMN assumption could be justified by the idea that congruent phonemes are supported by the whole prose stream because they have the same stimulus features. This fact would let the frequency-incongruent probes stand alone as in an oddball paradigm. Another argument for an MMN-like component regards the scalp distribution. Concordantly in both attentional states of the current experiment, a fronto-central scalp distribution slightly more enhanced contralateral to stimulus presentation was ascertained for the N1 component. This scalp distribution equaled the typical MMN distribution pretty well (Picton et al., 2000).

Maybe, the cause of the N1 could also have evoked the P2 component for F+400 probes compared to frequency-congruent stimuli. No P2 was elicited by F+ phonemes, whereas frequency-incongruent stimuli showed the component clearly. This may be interpretable as a sign of deviant detection as well. Normally, the P2 for simple tones has an earlier peak in the timeframe of about 180-190 ms (Shahin, Bosnyak, Trainor, & Roberts, 2003), but also has been reported at around 213 ms (Knight, Scabini, Woods, & Clayworth, 1988). The positive wave in this investigation showed a later peak (around 270 ms) and could thus also conform to an early P3a (Escera, Alho, Winkler, & Näätänen, 1998). The P3b as an alternative label for this component could be excluded because the current experimental setting contained task-irrelevant probe stimuli, and a P3b can only be elicit by active attention and a required response (Katayama & Polich, 1998; Sutton et al., 1965).

Moreover, it could be shown that a P3a cannot only be evoked by a ‘novel’ sound, but also by ‘typical’ deviant stimuli as long as they are distracting (for an overview see Polich, 2004). Comerchero and Polich (1998) suggested that a P3a may be generated by deviant-standard discrimination rather than by stimulus novelty as first assumed. Source localization yielded a generator in the frontal lobe, one decisive distinction between a P3a and P3b component (Jemel et al., 2003). Nonetheless, the observed central scalp distribution can also be observed for a P3a component instead of the typical frontal distribution especially for task-irrelevant stimuli such as occurs in the present study (e.g. Cycowicz & Friedman, 2004). At the moment, however, no certainty for determining the in this study observed component as either P2 or P3a over peak latency exists. Also, the P2 is not as extensively explored as the P3a, so that a reasonable conclusion is not possible yet. Therefore, further investigations about the distinction between the P2 and P3a are necessary, especially with a more detailed scalp topography resolution.

However, with regard to the N1 component of the current experiment which only occurred for frequency-incongruent compared to congruent stimuli, one can assume as discussed above, that the N1 might reflect a kind of difference detection (MMN). The F+400 probes varied from F+ phonemes in their fundamental frequency relative to the to-be-attended story and may not necessarily represent an attention modulation, but rather a mismatch to the general fundamental frequency while listening to one of the two stories. If this N1 deflection for F+400 stimuli reflects such a deviance detection, then the positivity afterwards is most probably a P3a rather than a P2 component, and is a kind of memory update (Polich, 2004). Then, the P3 would also indicate temporarily extra brain energy resources that were allocated to process those distracting F+400 task irrelevant stimuli. On the other hand, F+ phonemes (task irrelevant as well) did not require additional resources to be processed.

The P3a latency is associated with a measurement of stimulus classification speed independent of response processes (Kutas, McCarthy, & Donchin, 1977; McCarthy & Donchin, 1981; Pfefferbaum, Christensen, Ford, & Kopell, 1986); thus, the earlier the P3a onset the faster the stimulus classification. The earlier onset of the present P3a may be explicable by the fast stimulus presentation that required a fast stimulus processing. Furthermore, a large stimulus deviance facilitates stimulus classification could also maybe have caused this early component. Moreover, it was found that F+400 probes on the attended side elicit a larger P3a than on the unattended side. This also supports the theory that stimuli on the attended side are processed preferentially (Alain & Izenberg, 2003; Sonnadara, Alain, & Trainor, 2006).

Nevertheless, as seen in the difference waves between attended and unattended stimuli for F+ and F+400 probes, a negative difference occurred only for frequency-congruent stimuli. Thus, an obvious attentional modulation was present for F+ stimuli only. This leads to the argumentation that frequency-incongruent phonemes must have been outside the attentional focus; even when they were presented at the same location as the to-be-attended story. Otherwise they would have been as well processed differently on both sides in general. This suggested that frequency-incongruent phonemes on either side are less attended and therefore processed as though they were unattended. Furthermore, F+400 phonemes compared to F+ stimuli received less mental resources for their processing because of an exclusion from further processing (Alho, Tottola, Reinikainen, Sams, & Näätänen, 1987).

Thus, the assumed attentional advantage for F+400 probes seen in the early ERP components, especially the N1/MMN, was caused by physical differences between stimuli only, and not by attention itself. This confirmed again that those early components are more stimulus-driven rather than modulated by top-down attention. Later ERP components like the Nd are controlled by top-down attention primarily. They provide the most robust attentional effect.

Attended frequency-incongruent phonemes compared to the A-F+400 condition may have gained a little advantage for the first exogenous ERP components only because they were presented at the same location as the to-be-attended prose (spatial location advantage). Heinze et al. (1990) as well as Boksem et al. (2005) showed that, no matter how much the stimuli differ from each other at a given spatial location, all of them will enhance the early exogenous ERP components, indicating that spatial attention may select stimuli on the basis of location first before all (other) relevant features are processed. Even though the latter studies investigated visual phenomena, the findings may still be valid for auditory stimuli as well with regard to the assumption that attention and especially spatial attention is controlled by a sensory independent higher-level network system (Kastner, 2004). The first auditory support for this assumption came from Sonadara (2006) and Alain (2003).

Because of all these findings of modulated ERPs between conditions, one can also make a further additional conclusion. This study like the one from Woods et al. (1984) demonstrated that attentional modulations can be found for task-irrelevant probes, and not only for relevant stimuli as they were used in classical auditory selective listening tasks (e.g., Alho, Donauer et al., 1987; Schröger, 1994; Schwent, Snyder et al., 1976; Teder et al., 1993).

### Summary

Taken together, the validity of the method of combining HRTFs with the probe technique for a dichotic listening task was demonstrated. Furthermore, an attentional effect between attended and unattended probe stimuli (stronger processing of attended than unattended phonemes) was found in line with the literature. Moreover, stimuli that vary in fundamental frequency compared to the to-be-attended speaker’s voice showed a different as well as a reduced processing on both to-be-attended and unattended side indicating that they are processed as unattended stimuli. This means, even apart from being presented on the attended side, deviant probe stimuli were not in the attentional focus anymore to be processed strongly and therefore they were provided with less mental resources for their processing. Hence, attentional focus is determined not only by location but also by the fundamental frequency of a stimulus. The question of how sharply restricted the attentional focus on incoming stimuli is and whether a stimulus that is slightly different to attended stimuli is still selected as attended stimuli or not will be addressed in the following experiments.

## *Experiment 2.2:*

### Method

After demonstration of the HRTF technique's validity, another experiment was conducted to investigate the attentional focus more in detail. In the first study of this chapter, it could be shown that probes at the attended side, but with an extreme different fundamental frequency compared to the to-be-attended speaker's voice, were processed as unattended stimuli. This indicates at least that extreme frequency "deviants" are not in the attentional focus anymore. The current experiment rather introduces slight deviants. Are they also outside of the attentional focus and are they processed as unattended stimuli apart from being mixed into the to-be-attended stream? Or may slight deviants be able to pass the attentional filter? The latter case would argue for a slightly more, but not completely (evidenced by experiment one – extreme deviants are excluded) open attentional focus for slight variations in stimulus features.

This question will be addressed in the present second study by using two types of deviants: one with a small and one with a large frequency difference. Furthermore, the ability to pass the attentional filter and to be included in the attentional focus will be investigated for slight and extreme spatial deviants as well. The oddball paradigm was used to see whether or not deviants are noticed as being different from the to-be-attended stream. Moreover, it was shown that the MMN (mismatch negativity) varies with attention (Woldorff et al., 1991). In Woldorff et al.'s study (1991) the MMN for the unattended channel was significantly reduced compared to the one elicited by the attended auditory stream. On this basis, it may be feasible to investigate attentional effects besides deviant detection processing.

#### Subjects

The same 32 subjects as in experiment one participated.

#### Stimuli

To reduce eye movement, subjects fixated on a cross on a computer screen located in front of them during the recording. The experimental set-up and the subject's task were the same as for experiment 1. The variation from the first experiment consisted of the presentation probability and the degree as well as the kind of deviance from the to-be-attended speaker's voice. A phoneme ("da"; 100 ms) of the same fundamental frequency as the speaker's voice of the story appeared as a standard probe stimulus at the same location as the story in 80% of all probe stimuli. This implementation was the case for both stories - whether they were attended or unattended, resulting in attended and unattended standard probes at the same time. Further, the same phoneme could appear with a different fundamental frequency (F+) at the same location (L+) as the story: either as a slight frequency deviant (L+ F+60; a 60Hz higher pitch than the speaker's voice; probability = 5%) or as a more extreme frequency deviant (L+ F+400; a 400Hz higher pitch; probability = 5%). On the other hand, the spatial location of some probe stimuli was also varied in relation to the location of presented standard phonemes at 15° in space: either slight (L+15 F+; 15° deviants (additionally 15° away from the azimuth; at 30° finally); probability = 5%) deviants) or extreme (L+30 F+; 30° deviants (at 45°); probability = 5%) by keeping the phonemes' fundamental frequency (frequency of the speaker's voice) constant. However, just one attribute was changed at a time – either the frequency or the spatial location, but not both together.

These selected variations or combinations of stimulus patterns were chosen to reduce the experimental design matrix to clear arranged and analyzable components according to the question of interest (either frequency or spatial effects regarding attentional load). Phonemes were presented in randomized order with the same interstimulus interval (ISI) of 250 to 750 ms (rectangular distribution) as in experiment one, ensuring comparability of the results. For this experiment, 6000 probe stimuli (4800 standard probes (2400 attended, 2400 unattended) and 300 probes in each of the four deviant conditions (150 attended, 150 unattended) resulting in 1200 deviant probes (see Table 3). The stories and probe stimuli were presented in 6 runs (3 attended right/unattended left, 3 attended left/unattended right; 10 min each). The subject's task consisted of actively listening to one of the two parallel stories as instructed before a run started, and to keep the content of that story in their mind. No button press was required. After each run subjects were asked several questions about the to-be-attended story to make sure that they really directed their attention to that single story only. The whole session had a length of 1.75 h (excluding the EEG set-up).

**Table 3: Number of probe stimuli within the twenty experimental conditions. L+ represents the same location as the speaker's voice of the story, and F+ conforms to the same frequency of the speaker's voice, respectively. Deviant variations are indicated by a certain value following either L+ (for location: 15 represents 15°, or 30 stands for 30°) or F+ (for frequency deviants: 60 represents 60Hz higher, or 400 stands for 400Hz higher). A+ stands for "attended" and A- for "unattended".**

| | **left** | | **right** | |
|---|---|---|---|---|
| | ***A+*** | ***A-*** | ***A+*** | ***A-*** |
| **L+F+** (standard) | 1200 | 1200 | 1200 | 1200 |
| **L+F+60** (low frequency deviant) | 75 | 75 | 75 | 75 |
| **L+F+400** (highly frequency deviant) | 75 | 75 | 75 | 75 |
| **L+15F+** (low spatial deviant) | 75 | 75 | 75 | 75 |
| **L+30F+** (higly spatial deviant) | 75 | 75 | 75 | 75 |

## EEG-Recording and data analysis

EEG-recording and data analysis were performed as in experiment 1. Only the differences in analysis methods are described. The ERPs were averaged separately for attention condition (attended/unattended probes), experimental condition (standard or deviant (spatial (L+15, L+30), frequency (F+60, F+400))) and location (left or right). After preliminary analyses indicated no difference between effects to left and right-sided stimuli, ERPs to left and right-sided probe stimuli were collapsed to yield waveforms for electrode positions ipsi- and contralateral with regard to the location of the probes.

ERPs were quantified for the following components in the following time-windows: P1 (80-170 ms), Nd (320-450 ms), MMN (220-340 ms), P3 (310-360 ms), and RON (reorienting negativity; 400-600 ms). The RON is seen as a sign of a switch back towards a primary task (Schröger, Giard, & Wolff, 2000), or in this case towards the to-be-attended stimulus, after attention was diverted by a distractor. The repeated measures ANOVAs contained the factors Condition (standard, slight or extreme deviant), and Electrode(s) (selected electrode(s) for spatially restricted effects); or if possible the factors Lateratility (left or right), and Anterior-posterior (anterior or posterior) instead of the factor Electrode(s).

## Results

### Subjects

The results of fifteen subjects (13 women, 2 men, mean age 27.0, range 20-33) out of 32 were statistically analyzed. The data of the remaining 17 subjects had to be discarded, because of too many artifacts or technical failures. The same exclusion criteria as in the first study were used to exclude a person from further analyses.

### Behavioral results

According to the questionnaire, it could be concluded that all subjects were able to concentrate on the specified story, seen in 65% correct answers in average. The percentage of correct answers was above chance for all subjects and varied from 51% to 84% (appendix A3b).

### Electrophysiology

The following ERP figures display group average ERPs (n = 15) and difference waves for either attended location, standard probes (A+L+F+), attended location, slight (A+L+F+60) or extreme frequency deviants (A+L+F+400), attended location, slight (A+L+15F+) or extreme spatial deviants (A+L+30F+), unattended location, standard probes (A-L+F+), unattended location, slight (A-L+F+60) or extreme frequency deviants (A-L+F+400), unattended location, slight (A-L+15F+) or extreme spatial deviants (A-L+30F+). ERPs from left and right stimuli were collapsed to yield ipsi- (i) and contra-lateral (c) sites.

A complete omnibus ANOVA including attentional, frequency, and spatial dimensions was not possible due to the fractional factorial design (see previous paragraph about stimuli). Therefore, frequency and spatial effects regarding attentional load were computed separately in the following paragraphs by repeated measures ANOVAs with factors Condition (attended vs. unattended, standard vs. slight and extreme deviant), and Electrode(s) (all or spatially selected head electrodes) or Laterality (left, right) and Anterior-posterior (anterior, posterior) if possible (set to p = .05).

## *Attentional effects – attended vs. unattended*

1) Standard probes – spatial and frequency-congruent stimuli

The first comparison concerned the standard stimuli (L+F+) on the attended and unattended side. In contrast to experiment 1, a clear fronto-central Nd effect (320-450 ms), without an overlaying Pd component, was observed with a more negative-going ERP wave for attended standards compared to A-L+F+ stimuli, Figure 19. This Nd effect showed an emphasis contralateral to stimulus presentation. Difference waves were obtained by subtracting unattended L+F+ stimuli from the attended L+F+ probes (standards).

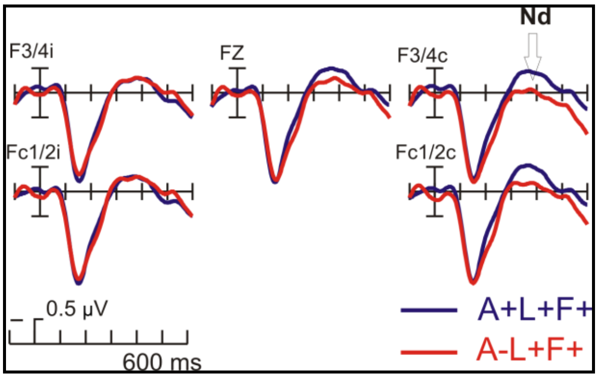


**Figure 19: Group average ERPs for auditory standard probe stimuli (L+ F+) when they were attended (A+) and the same stimuli when they were unattended (A-).**

ANOVA of the root mean square revealed a statistical negative displacement (Nd; 320-450 ms) between frequency and spatial congruent probe stimuli in both attention conditions (overall (all electrodes) main effect of Condition: F(1,14) = 5.89, pHF = .0293). A slight maximum of this Nd effect was observed at electrode Fc1/2c (F(1,14) = 9.04, pHF = 0.0094). In general, unattended probes have shown a less negative wave than attended stimuli.

The isovoltage map shows a fronto-central Nd scalp distribution with an emphasis contralateral to stimulus presentation (Figure 20).

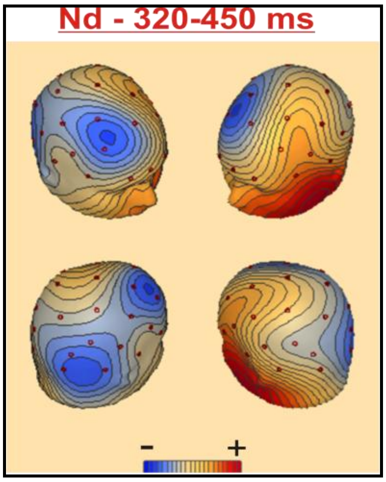


**Figure 20: Isovoltage map for visualization of the topographical distribution of the Nd effect: difference waveforms (A+L+F+ minus A-L+F+) are shown within the accordant timeframes. Isovoltage spline interpolation for the 320-450 ms interval was used (BESA).**

## *Deviance effects – standard vs. slight vs. extreme deviant*

1) <u>Frequency effects – slight and extreme frequency deviants vs. standard probe</u>

Figure 21a) and b) show the next comparison between A+ standard probes and both types of A+ frequency deviants (slight: L+F+60 and extreme: L+F+400) at central electrodes. There was a large MMN (175-225 ms) for extreme frequency deviants (L+F+400) whereas the MMN for slight frequency deviants was not as strong compared to standard probe stimuli (L+F+). However, both deviant MMNs seemed to vary not only in peak amplitude but also in peak latency. Thereafter, both deviants showed a P3a component (270-330 ms) which was more dominant for the A+L+F+400.

When the same but unattended stimuli were compared to one another the same but partially weaker pattern as for attended probes occurred (Figure 21c) and d)). The only component that was stronger at unattended compared to attended sites was the P3a for L+F+60 probes. Again, a salient MMN (170-215 ms) is observed for A-L+F+400 stimuli only, although A-L+F+60 probes showed a negative shift as well in that time window. This component was followed by the P3a (310-360 ms) for both deviant probes (A-L+F+60 and A-L+F+400).

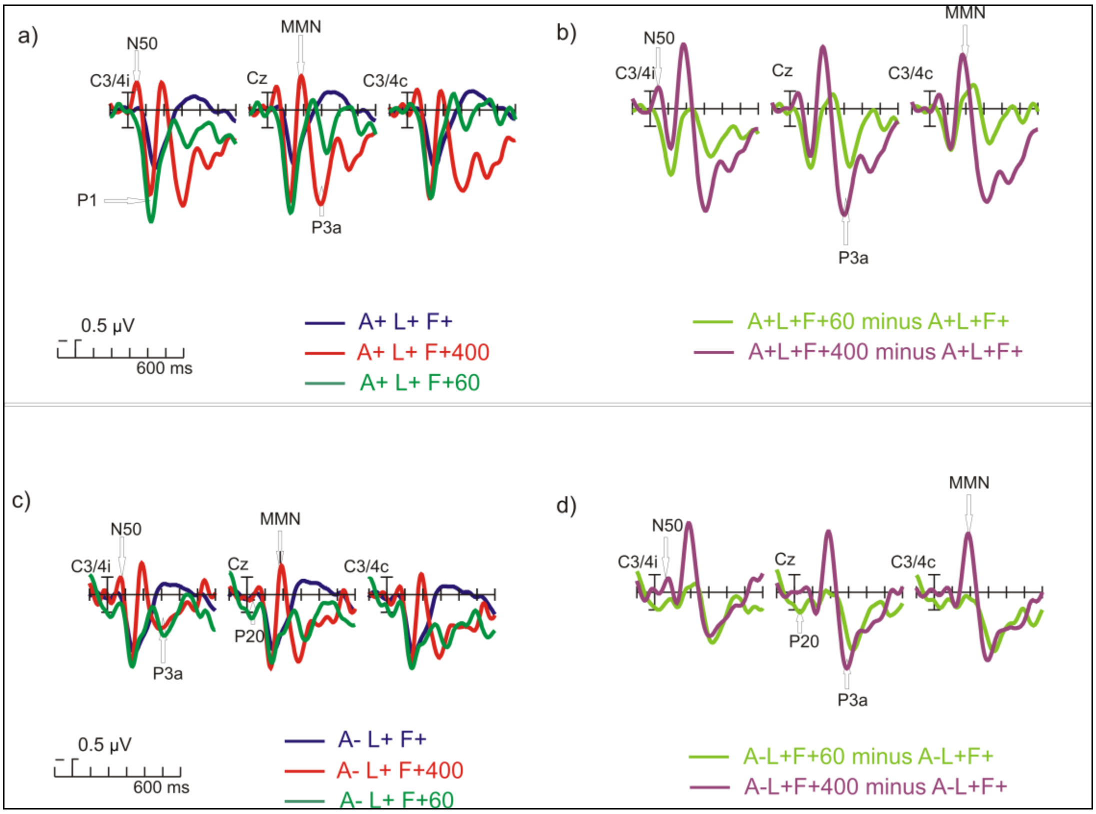


**Figure 21: Group average ERPs and difference waves for attended and unattended auditory probe stimuli on the frontal electrodes; ipsi- (i) and contra-lateral (c) sites were accounted for both, right and left probe stimuli combined. F+L+ shows that the probes had the same frequency (F) and spatial location (L) as the speakers' voice of the story on each ear respectively.** a) **Attended rare deviants (A+L+F+60 (slight deviant; 60 Hz higher frequency; 2.5%) and A+L+F+400 (extreme deviant; 400 Hz higher frequency; 2.5%)) in comparison with attended standard probes (A+L+F+; 40%).** b) **Difference waves between those attended deviants and the standards are shown.** c) **Unattended deviants (A-L+F+60 (slight deviant; 60 Hz higher frequency; 2.5%) and A-L+F+400 (extreme deviant; 400 Hz higher frequency; 2.5%)) in comparison with unattended standard probes (A-L+F+; 40%). Difference waves are illustrated in** (d). **A+ stands for attended and A- for unattended stimuli.**

The next figure (Figure 22) makes the differences between frequency deviant stimuli and standards clearer. The same, but attenuated, difference-pattern for the comparison between unattended extreme frequency deviants and standard probes as for the attended side was observed (Figure 22a), whereas the P3a difference effect between L+F+60 and L+F+ was not modulated from attended to unattended side (Figure 22b).

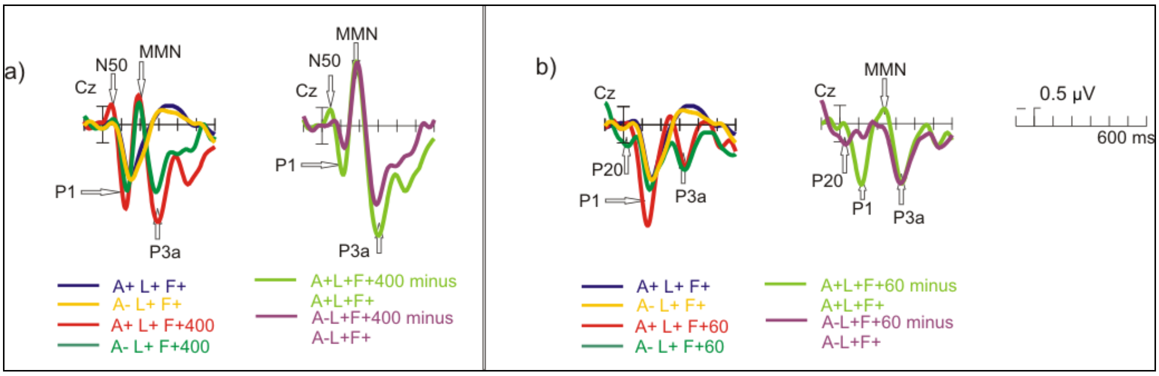


**Figure 22: Group average ERPs and their difference waves for** a) **extreme frequency deviants (A+/A- L+ F+400) compared to standard probes (A+/A- L+ F+) and** b) **slight frequency deviants (A+/A- L+ F+60) compared to standard probes (A+/A- L+ F+) at the electrode Fz, whereas right and left probe stimuli are combined.**

### a. Attended side

ANOVAs including all three attended conditions regarding spatial effects resulted in significant differences between standards, slight and extreme spatial stimuli. The main effect of condition was seen for all 30 electrodes between 175-225 ms (F(2,28) = 3.74, pHF = 0.0385), and between 250 and 350 ms (F(2,28) = 4.00, pHF = 0.0324).

In the following, pair-wise ANOVAs for attended slight or extreme frequency deviants vs. standards were computed which revealed a statistically significant MMN (175-225 ms) for extreme frequency deviants (A+L+F+400) only, and a statistically significant P3a (320-330 ms/250-350 ms) for both slight and extreme frequency deviants. Only the MMN component differed between both degrees of deviants.

#### ***MMN***

For the pair-wise comparison between standard and slight frequency deviants, none of the ANOVAs reached significance, suggesting a similar ERP at that time window for both conditions.

In contrast, the comparison between standard and extreme frequency deviants for the MMN showed a significant main effect of Condition (overall: F(1,14) = 11.12, pHF = 0.0049) indicating a negative deflection compared to standards.

Differences between A+L+F+60 and A+L+F+400 were found as well (overall main effect of Condition: F(1,14) = 4.89, pHF = 0.0441). A larger MMN for extreme frequency deviants was especially pronounced at anterior and ipsilateral electrodes. This observation was supported by a midline (Fpz, Fz, Cz, Pz) condition x anterior-posterior interaction (F(3,42) = 3.60, pHF = 0.0431) and a parasagittal (Fp1/2, F3/4, C3/4, P3/4, O1/2) condition x laterality interaction (F(1,14) = 5.37, pHF = 0.0361).

The isovoltage map for the extreme frequency deviants (L+ F+400) underscores a fronto-central distribution with a slight ipsilateral shift for the MMN (175-225 ms) component (Figure 23).

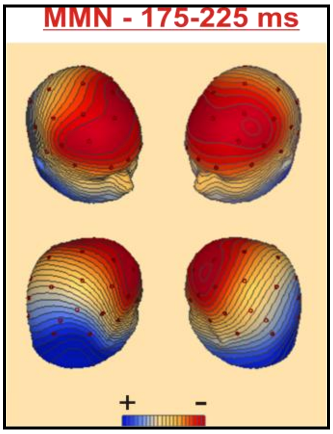


**Figure 23: Topographical maps for the MMN component found for the extreme frequency deviants by using isovoltage spline interpolation for the 175–225 ms interval. Map was generated for difference waves between attended standard probe stimuli (A+L+F+) and attended extreme frequency deviants (A+L+F+400). Note that relative scaling was used.**

### *P3a*

The P3a was significant for extreme frequency deviants (A+L+F+400) compared to standards (A+L+F+): overall main effect of Condition: F(1,14) = 9.45, pHF = 0.0082. The midline (Fpz, Fz, Cz, Pz: F(3,42) = 7.57, pHF = 0.0024) ANOVA showed a significant condition x anterior-posterior interaction, indicating larger central amplitudes.

The pair-wise comparison between standard and slight frequency deviants on the other side could also point out a significant P3a at central to centro-parietal electrodes ((C3/4, Cz, Cp1/2, Cp5/6) main effect of Condition: F(1,14) = 5.12, pHF = 0.0401). Furthermore, the condition x anterior-posterior interaction confirmed the emphasis of this effect on the central sites (midline - Fpz, Fz, Cz, Pz: F(3,42) = 5.71, pHF = 0.0083).

Isovoltage maps for both slight (A+L+F+60) and extreme frequency deviants (A+L+ F+400) on the attended side confirm the statistically ascertained central midline distribution for the early P3 (P3a) component (320-330 ms/250-350 ms) (Figure 24).

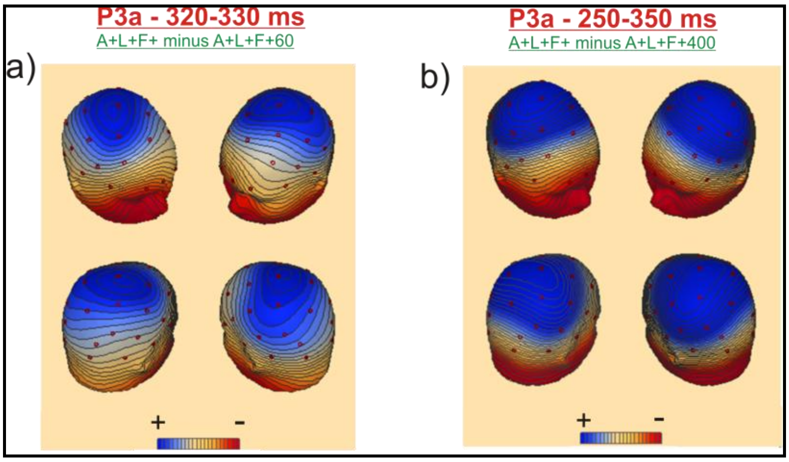


**Figure 24: Isovoltage map for the topographical distribution of the P3a: difference waveforms (**a) **A+L+F+ minus A+L+F+60;** b) **A+L+F+ minus A+L+F+400) are shown. Isovoltage spline interpolation for the 320-330 ms** (a) **and 250-350 ms** (b) **intervals were used (BESA). Note that relative scaling was used.**

### b. Unattended side

Omnibus ANOVAs with all three conditions (standard, slight and extreme deviant) together in one factor of Condition and with all 30 electrodes revealed significant differences in the MMN time frame (170-215 ms; mean amplitude: F(2,28) = 3.59, pHF = 0.0420) and also for the P3a component (240-360 ms; peak amplitude: F(2,28) = 7.74, pHF = 0.0024).

Pair-wise ANOVAs for the comparison between A-L+F+ and both frequency deviants (A-L+F+60 and A-L+F+400) revealed the same statistical pattern of significant components as for the attended side (Table 4): a significant MMN (170-215 ms) for extreme frequency deviants (A-L+F+400) only, and a statistically significant P3a (310-360 ms/240-285 ms) for both slight and extreme frequency deviants. Only the MMN component resulted in significant ANOVAs between both kinds of deviants.

**Table 4. Significant results (main effect of condition) of pair-wise ANOVAs between frequency deviants and standard stimuli on the unattended side.**

| components | conditions | max. effect | F-value (df) | p-value (pHF) |
|---|---|---|---|---|
| **MMN** | A-L+F+ vs. A-L+F+60 | -- | -- | -- |
| | A-L+F+ vs. A-L+F+400 | overall | 7.39 (1,14) | 0.0166 |
| | A-L+F+60 vs. A-L+F+400 | overall | 4.83 (1,14) | 0.0453 |
| **P3a** | A-L+F+ vs. A-L+F+60 | overall | 5.53 (1,14) | 0.0338 |
| | A-L+F+ vs. A-L+F+400 | overall | 4.61 (1,14) | 0.0498 |
| | A-L+F+60 vs. A-L+F+400 | -- | -- | -- |

### c. Attended vs. unattended side – frequency deviants

Regarding pair-wise comparisons between attended and unattended frequency deviants, neither the MMN nor the P3a differed significantly for both types of frequency deviants (L+F+60 or L+400F+), though one might expect it by visual inspection.

## 2) Spatial effects – slight and extreme spatial deviants vs. standard probe

With regard to spatial effects, Figure 25 displays an MMN and RON effect. Between 220 and 340 ms, spatial deviants elicited more negative amplitudes than standard probes, indicating an MMN for spatial deviants which does not seem to differ in mean amplitude (see difference waves) but in peak latency with an earlier peak for slight spatial deviants. Later on, between 400 and 600 ms, a re-orienting negativity (RON) was observed on central sites. On the unattended side, the same pattern occurred when both kinds of spatial deviants (A-L+15F+ and A-L+30F+) were compared to unattended standard probes (A-L+F+). The MMN (190-240 ms) seemed to be slightly larger for extreme than for slight spatial deviants especially at electrodes contralateral to stimulus presentation. As seen for attended spatial deviants, between 300 and 500 ms, a RON occurred at central sites again.

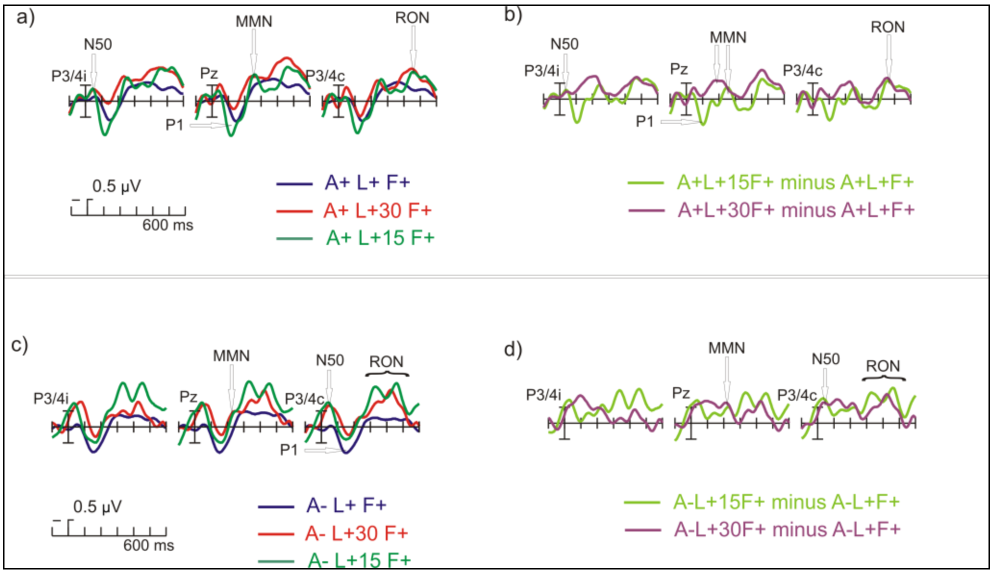


**Figure 25: Group average ERPs for attended and unattended auditory probe stimuli with right and left probe stimuli combined for ipsi- (i) and contra-lateral (c) sites.** a) **Attended rare deviants (A+ L+15 F+ (slight deviant; 15° further away from azimuth additionally to standard location) and A+ L+30 F+**

**(extreme deviant; 30° further away from azimuth additionally to standard location)) in comparison with attended standard probes (A+L+F+).** b) **Difference waves between those attended deviants and the standards are shown.** c) **The unattended deviants (A+ L+15 F+ (slight deviant) and A+ L+30 F+ (slight deviant)) in comparison with unattended standard probes (A-L+F+). The differences are seen clearer in the difference waves** (d) **respectively. A+ stands for attended and A- for unattended stimuli.**

In the following figure (Figure 26), differences between spatial deviant stimuli (L+15F+ and L+30F+) and standards (L+F+) on both attended and unattended side are shown more clearly. The same, but (slightly) boosted MMN pattern for the comparison between unattended extreme spatial deviants and standard probes (Figure 26a) and for the comparison between A-L+15F+ and A-L+F+ stimuli (Figure 26b) was observed compared to the attended side.

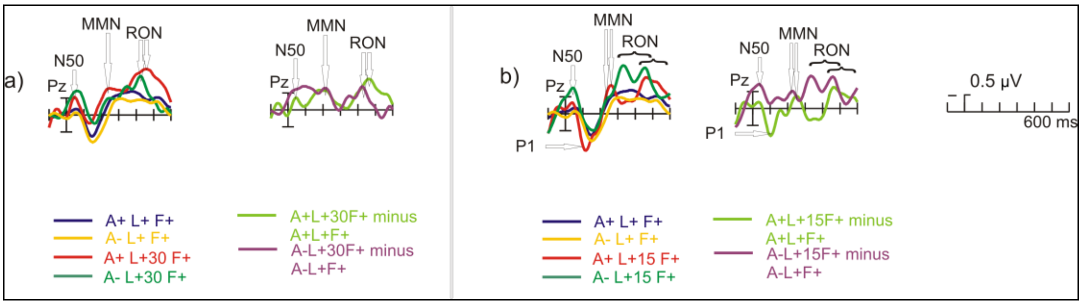


**Figure 26: Group average ERPs and their difference waves for** a) **extreme spatial deviants (A+/A- L+30 F+) compared to standard probes (A+/A- L+ F+) and** b) **slight spatial deviants (A+/A- L+15 F+) compared to standard probes (A+/A- L+ F+) at the electrode Fz, whereas right and left probe stimuli are combined.**

## a. Attended side

Omnibus ANOVAs for all 30 electrodes with all three conditions (standard, slight and extreme deviant) together in the factor of Condition revealed significant differences for the following ERP components. In the MMN time frame (220-340 ms; F(2,28) = 5.35, pHF = 0.0107) and also for the RON (400-600 ms; F(2,28) = 7.95, pHF = 0.0018) significances were calculated. In the following, these results were examined in pair-wise comparisons (peak amplitude measurements) that revealed a statistically significant MMN (220-340 ms/190-240 ms), and RON for A+L+15F+ (slight) and A+L+30F+ (extreme) deviants (400-600 ms).

### ***MMN***

The pair-wise comparison between standard and slight spatial deviants (220-340 ms) revealed an overall main effect of Condition (F(1,14) = 11.47, pHF = 0.0044). This confirms the visually observed MMN for A+L+15F+ probes relative to standards.

Significant differences were also statistically confirmed for standard probes compared to extreme spatial deviants (overall main effect of Condition: F(1,14) = 9.49, pHF = 0.0081). No significances in either peak amplitude or peak latency could be found between slight (A+L+15F+) and extreme (A+L+30F+) spatial deviants.

### ***RON***

Another component occurring for both spatial deviants in the pair-wise comparison to standard stimuli was a late negative deflection (RON) between 400-600 ms. This component has been interpreted as neural processing reflecting reorienting of attention after distraction (Schröger et al., 2000).

The first comparison addressed slight spatial deviants contrasting standard stimuli. Thereby, the overall main effect of Condition reached significance (F(1,14) = 11.69, pHF = 0.0042). This result suggests a more negative-going ERP waveform for A+L+15F+ probes than for standard stimuli.

A similar pattern of significances was shown for the comparison between standard and extreme spatial deviants. The main effect of Condition became significant (overall: F(1,14) = 11.29, pHF = 0.0047) reflecting a more negative ERP for A+L+30F+ deviants compared to the ERP for A+L+F+ probes.

No significances in amplitude could be found between slight and extreme spatial deviants, indicating the same strength of RON effect for both conditions.

## b. Unattended side

Omnibus ANOVAs with all three conditions (standard, slight and extreme deviant) together in the factor of Condition revealed significant differences for the MMN as well as for the RON. The peak amplitude differed significantly in the following time frames: 160-250 ms (MMN; overall: F(2,28) = 4.26, pHF = 0.0242), and 400-600 ms (RON; overall: F(2,28) = 9.14, pHF = 0.0009). Further on, these results were examined in pair-wise comparisons which also revealed a statistically significant MMN (190-240 ms/160-250 ms) and a significant RON (300-500 ms) for both spatial deviants relative to standard probe stimuli (Table 5). Thereby, between both kinds of spatial deviants, neither the MMN nor the RON reached significant differences.

**Table 5. Significant results (main effect of condition) of pair-wise ANOVAs between spatial deviants and standard stimuli on the unattended side.**

| components | conditions | time window | max. effect | F-value (df) | p-value (pHF) |
|---|---|---|---|---|---|
| **MMN** | A-L+F+ vs. A-L+15F+ | 190-240 ms | overall | 6.54 (1,14) | 0.0228 |
| | A-L+F+ vs. A-L+30F+ | 160-250 ms | overall | 5.49 (1,14) | 0.0344 |
| **RON** | A-L+F+ vs. A-L+F+ | 300-500 ms | overall | 11.34 (1,14) | 0.0046 |
| | A-L+F+ vs. A-L+F+ | 300-500 ms | overall | 11.74 (1,14) | 0.0041 |

## c. Attended vs. unattended side – spatial deviants

With regard to pair-wise comparisons between attended and unattended spatial deviants, neither the MMN nor the RON differed significantly for both types of spatial deviants (slight: L+15F+ or extreme: L+30F+).

## Discussion

The aim of this second experiment was to investigate the nature of the attentional focus. More specifically, it was analyzed which stimuli features are included within the attentional focus and to which degree variations in feature values lead to an exclusion from the focus of attention while attending to a certain prose stream determined by the speaker's voice characteristics (spatial location, and fundamental frequency).

This issue was investigated by using an oddball-paradigm with spatial and frequency deviants that varied slightly (slight deviants) or largely (extreme deviants) from standard probe stimuli (frequency and spatial congruent to speaker's voice characteristics). With that stimulus set up, characteristic components such as the MMN and P3a were expected as a sign for deviance detection and attentional shift. Therefore, if those components can be observed, one can argue that changes in stimulus feature were noticed (MMN) by the subject's brain which also may cause a succeeding attentional orienting to that deviant stimulus (P3a) if this detected difference was located outside of the actual focus at other stimulus specific features. This shift of attention is an evidence for refocusing. This means that the attentional focus moved from the to-be-attended prose stream to the deviating stimulus. Otherwise, if this stimulus was included in the focus of attention, no such attentional shift would have been necessary. Instead, the infrequent presentation of those deviant stimuli would have been noticed (MMN) without refocusing (P3a). Thus, changes in stimulus features alone such as spatial location or the fundamental frequency may determine whether the stimulus will be included in the focus of attention.

According to these hypotheses, statistical analyses of mean amplitude, peak amplitude or peak latency were calculated. First of all, and as a reproduction of the attentional effect reported widely in the literature (Jemel et al., 2003; Woldorff et al., 1993), the present experiment also revealed the Nd-effect for attended vs. unattended stimuli for the standard probes (frequency and spatial congruent). This effect provided evidence that attention is important to modulate stimulus processing. It demonstrates differently allocated resources (more brain resources for attended stimuli) and the level of processing activity (higher for attended than unattended stimuli) as a result of sensory gain (Woldorff et al., 1993).

As the first experiment already demonstrated, the HRTF method combined with the probe technique is a valid technique to investigate auditory attention. This present HRTF study has replicated attentional effects mentioned in previous investigations using simple binaural presentation or interaural time difference (ITD) and interaural level difference (ILD) manipulations.

Moreover, regarding the effects from both kinds of deviant stimuli – fundamental frequency or spatial location – in their slight or extreme deviant manner, statements about the characteristics of the attentional focus were derived. In an oddball-experiment, typically an MMN for deviant stimuli compared to standards is observed (Picton et al., 2000, for a review). Frequency effects for the attended side were reported first. The frequency deviant MMN reached significance for extreme frequency deviants only, whereas the slight frequency deviant's negative peak was not strong enough to differ from the standard's ERP waveform in that MMN time window. This result confirmed the reported MMN amplitude as a function of deviance (Berti, Roeber, & Schröger, 2004): a smaller MMN for slight frequency deviants, and a larger MMN for extreme frequency deviants. While the slight deviants appeared to show an MMN on visual inspection, the difference between slight deviants and standards failed to reach significance.

The MMN represents an automatic change detection depending on an auditory memory trace (repetition of a standard probe) (M. Giard et al., 1990; Näätänen, 1992) and also a preparation for an attentional switch (Schröger, 1996). Thus, occasional deviants are not included into the memory trace. Hence, whenever such a deviant appears then an attentional switch from standard probes to the deviant stimulus comes along with that. This would also explain why a P3a (see next paragraph) as an indicator for the actual switch follows the MMN within the ERP waveforms. However, no significant smaller MMNs, though visually apparent, were revealed for both slight and extreme frequency deviants on the unattended side relative to the attended location. The slight advantage for stimuli on the attended side might have been occurred because of spatial benefit. Alain and Izenberg (2003) found a facilitation of early sensory processing for items presented to a location where attention is focused – no matter how much stimuli differed. Thus, this argument of a larger MMN because of attention is no longer crucial, especially because the MMN is an exogenous component. The inspection of later endogenous components (truly modulated by attention as a top-down process) such as the P3a are helpful indicators of a stimulus' relation to the attentional focus.

Both frequency deviants evoked a significant P3a relative to standard stimuli indicating an involuntary switch of attention toward sound changes (Schröger & Wolff, 1998). Why is the currently observed component a P3a and not a P3b. The distinction between P3a and P3b emerges because the stimulus context defines the degree of attentional focus. If there is no attention directed to those deviants, a P3a (involuntary capture of attention or orienting) is probable, whereas a P3b occurs in a primary discrimination task (conscious attention is needed) (Katayama & Polich, 1998) that requires a task relevant target stimulus to respond to (Sutton et al., 1965). In the current case, the probe stimuli were all task-irrelevant and thus not be able to elicit a P3b. Moreover, it could be shown that a P3a cannot only be evoked by a 'novel' sound, but also by 'typical' deviant stimuli as long as they are distracting (for an overview see Polich, 2004). Nevertheless, a rather central instead of a classical frontal scalp distribution was observed. Other studies (e.g. Cycowicz & Friedman, 2004) also reported this central and less frontal effect for the P3a in their results for especially task-irrelevant stimuli like in the present case which first may be mistakable for the P3b distribution. Howsoever, a P3b has still a slightly more posterior appearance at centro-parietal electrodes (Comerchero & Polich, 1999).

As the P3a between attended and unattended deviant stimuli was compared, no significant difference for either frequency deviant was found. No later difference was observed either. Thus, the lack of a negative difference (Nd effect; sign of attentional modulation) in both cases (see Figure 27a) was taken as evidence for their exclusion from the attentional focus. Apart from that, as mentioned initially, the P3a is an indicator for an attentional switch. The conclusion would be that whenever an attentional switch is necessary then the stimulus must be outside of the attentional focus. Otherwise, no switch would be required because attention is assumed to be centered on all stimulus features within the attentional focus.

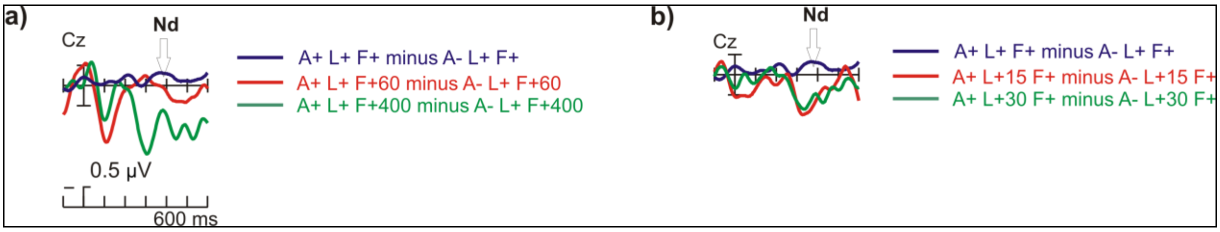


**Figure 27. Difference waves for** a) **spatial and** b) **frequency deviants in comparison to standard probes. An Nd effect occurred for standard phonemes only. Neither of the deviants' difference waves revealed such an attention effect.**

The difference in P3a amplitude between both frequency deviants on either side may have arisen from the degree of deviance from standards. Thus, there is a highly restricted focus for frequency, but with a gradual rather than an abrupt fading. This would explain the gradually increasing attentional switching potential with increasing degree of deviance. Doeller et al. (2003) also found an increase of P3a amplitude as a function of deviance. Thus, the smaller P3a for slight frequency deviants can be accounted for by the minimal variance and therefore harder discrimination from the standard probes. Extreme frequency deviants on the other hand elicited a larger P3a, because they were easier, and therefore more obvious to discriminate from standard probes (Katayama & Polich, 1998). Hence, the relative perceptual distinctiveness among stimuli affects the P3a (switch of attention) amplitude (Comerchero & Polich, 1999).

Spatial deviants on the other hand elicited, whether they were slight or extreme spatial deviants, a clear MMN. The amplitude and peak latency did not differ significantly between both degrees of spatial deviants. In contrast, some researchers (Nager, Teder-Sälejärvi, Kunze, & Münte, 2003; Paavilainen, Karlsson, Reinikainen, & Näätänen, 1989) found that MMN amplitude increased with increasing spatial separation between standard and deviant stimuli, whereas other studies like in this current study suggested that MMN elicited by changes in location is “all or nothing” (Colin, Radeau, Soquet, Dachy, & Deltenre, 2002; Shestakova, Ceponiene, Huotilainen, & Yaguchi, 2002). Furthermore, there was no significant reduced MMN component for the unattended compared to the attended deviants. This result let preliminary assume that the attentional focus is not only restricted to the to-be-attended speaker’s voice (equal to spatial deviants) only, but also, and even more, to its location. Therefore, even little changes in the spatial orientation had an enormous effect regarding integration in the attentional focus. Because of being presented at another spatial location on the same attended side as the to-be-attended prose even partial attention benefits in stimulus processing were not available anymore. Thus, attentional advantage was provided only for stimuli at the same location (Alain & Izenberg, 2003) as pointed out above, and therefore spatial deviants on the attended side were out of focus and processed like their counterparts on the unattended side.

Furthermore, there was no P3a for spatial deviants which was also seen in Doeller et al.’s (2003). This let assume that deviant stimuli within the attended stream like the frequency deviants of the current study (apart from being attended and focused directly) have a higher probability to cause P3a components than stimuli within the unattended but concurrent stream on the same side. Hence, spatial deviants are seen as belonging to another but not to the to-be-attended prose stream because of their different spatial location. Thus, they are not able to elicit a significant P3a as an indicator of attentional switch.

Apart from this P3a lack for spatial deviants, they evoked an additional late negativity instead; reorienting negativity (RON; discovered by Schröger et al. (1998)). The RON is thought to mirror the reorientation back to the task-relevant information after an attentional switch towards a deviant and distracting stimulus (Berti et al., 2004; Escera, Alho, Schröger, & Winkler, 2000; Rüsseler, Kowalczuk, Johannes, Wieringa, & Münte, 2002). The occurrence of a RON for spatial deviants in the present study provided further evidence for the assumptions made before, that even small deviances in space are not included in the attentional focus anymore. It shows that there must have been a switch of attention toward those spatial deviants; otherwise no reorienting would have occurred. Furthermore, this assumption was confirmed by comparing ERP between the attended and unattended side. There was no significant RON difference for both spatial deviants. Thus, no attentional modulation occurred (see Figure 27b).

### Summary

In summary, it could be found that the actual attentional focus is quite restricted to stimulus features of the attended object. It is not enough to be presented at the same to-be-attended side or in a quite similar pattern as the actual attended object. Even small, and not only extreme deviances in any stimulus's feature (for instance, spatial location, and fundamental frequency) cause an exclusion from the attentional focus and lead therefore to a different processing than for standard stimuli sharing the same physical features with the to-be-attended story's voice. Thus, the precision of the attentional focus was clearly limited to the to-be-attended object; in the present case the speaker's voice of the to-be-listened-to story. This attended object only is processed preferentially whereas all other stimuli experience a lower priority in stimulus processing. All different deviant components (MMN, P3a, and RON) showed thereby an agreeing pattern without any attentional modulation. Thus, any kinds of deviants were outside of the focus of attention even though they were presented on the same side as the to-be-attended story. Nevertheless, the strongest effect was provided for spatial deviants.

## *Experiment 2.3:*

### Method

The preceding experiments had shown that the attention focus is quite restricted to certain input features. Spatially, no deviant was able to pass the filter determined by the attentional focus and to get more mental resources for processing. With regard to the fundamental frequency of to-be-attended stimuli, a slightly wider attentional focus was found which let pass at least slight frequency deviants as well besides the frequency-congruent stimuli. Thereafter, another question arose which led to this third experiment. How important is the spectral information of a speaker's voice for selective attention? Is the attentional focus restricted to the frequency spectrum of a speaker's voice only or is it relatively independent of a stimulus' frequency range? Again, with a passive probe technique including phonemes (by the speaker's voice) and a noise probe superimposed on two spatially separate presented prose streams this question should be addressed. This time, an EEG with a higher spatial resolution (64 instead of 32 electrodes as in the other experiments) was used. By applying the same experimental design for all of the experiments with changes in the kind of probe stimuli only, one can expect the same basic mechanisms for the required attentional processes in each set-up.

#### **Subjects**

The same 32 subjects as in the previous two experiments were recruited.

## Stimuli

The procedure was the same as in the preceding experiment.

Only the superimposed probe stimuli varied. A phoneme ("da") of the same frequency as the speaker's voice respectively or a white noise (44 kHz; 5ms onset-offset; band pass filtered 200 – 5000 Hz) was presented (100 ms each; also convolved with individualized HRTFs) at the location of either the attended or unattended story.

White noise contains a broad frequency spectrum, while speech/phonemes are relatively band-limited at high frequencies.

For this experiment, 4000 probe stimuli in total (500 probes for each of the eight conditions: left attended white noise; left unattended white noise; right attended white noise; right unattended white noise; left attended phonemes; left unattended phonemes; right attended phonemes; right unattended phonemes; see Table 6) were presented in 4 runs (2 attended right, 2 attended left) lasting about 11 min each. The subject's task consisted of actively listening to one of the two parallel stories as instructed before a run started, and to keep the content of that story in their mind. No button press was required. After each run the subjects were asked several questions about the attended story to make sure that they really directed their attention to that story only. The whole session had a length of 1.25 h (excluding the EEG set-up).

**Table 6: Distribution of probe stimuli within eight experimental conditions. Phon represents phonemes, and noise conforms to white noise probe stimuli. A+ stands for "attended" and A- for "unattended".**

| | **left** | | **right** | |
|---|---|---|---|---|
| | ***A+*** | ***A-*** | ***A+*** | ***A-*** |
| **phon** | 500 | 500 | 500 | 500 |
| **noise** | 500 | 500 | 500 | 500 |

## EEG-Recording and data analysis

The EEG was recorded by using an Easy-Cap with 64 plug-in silver-silver-chloride electrodes (positions: Fp1, Fp2, F1, F2, F3, F4, F5, F6, F7, F8, Ft7, Ft8, C1, C2, C3, C4, C5, C6, P1, P2, P3, P4, P5, P6, P7, P8, Fpz, Fz, Cz, Pz, T7, T8, Tp7, Tp8, Fc1, Fc2, Fc3, Fc4, Fc5, Fc6, Cp1, Cp2, Cp3, Cp4, Cp5, Cp6, P7, P8, P3, P4, Po3, Po4, Po7, Po8, O1, O2 of the international 10-20 system (Jasper, 1958)). All other parameters of EEG-recording were identical to those used in experiments 1 and 2.

The ERPs were averaged separately for attention condition (attended or unattended probes), experimental condition (phoneme or white noise) and location (left or right). After preliminary analyses had indicated no difference between effects to left and right-sided stimuli, ERPs to left and right-sided probe stimuli were collapsed to yield waveforms for electrode positions ipsi- and contralateral with regard to the location of the probes. ERPs were quantified by mean amplitude measures and the resulting data were subjected to repeated measures analyses of variance (ANOVAs). First, the Nd, P1, MMN and P3 amplitude data were submitted to an overall two-way ANOVA with factors of condition, and electrodes (all 56 head sites). Then, some closer ANOVAs were calculated: three two-way midline, lateral, and anterior-posterior ANOVA with the factors of condition and sites, and three three-way ANOVAs (frontal, parasagittal, and temporal ANOVAs) with the factors of condition, laterality, and anterior-posterior each.

## Results

### Subjects

The results of seventeen subjects (14 women, 6men; mean age 27.7, range 20-35) out of 32 were statistically analyzed. The data of the remaining 15 subjects had to be discarded, because of too many artivacts or technical failures. If more than one third of epochs were rejected due to artifacts (exclusion criteria), a person was excluded from further analyses.

### Behavioral results

According to the questionnaire, it could be concluded that all subjects were able to concentrate on the specified story, seen in 71% correct answers in average. The percentage of correct answers was above chance for all subjects and varied from 50% to 92% (appendix A3c).

### Electrophysiology

The following ERP figures display group average ERPs (n = 17) and difference waves for either attended location, phonemes (A+ phon), attended location, white noise probes (A+ noise), unattended location, phonemes (A- phon), or unattended location, white noise probe stimuli (A- noise). ERPs from left and right stimuli were collapsed to yield ipsi- (i) and contra-lateral (c) sites.

Mean amplitudes were obtained in several time-windows separately for each subject, condition and electrode site. Significant results (p < .05) of the calculated omnibus and separate regional or single electrode ANOVAs are reported in the following paragraphs broken down into attentional and deviant effects.

The separate electrode groups for the regional ANOVAs in this current experiment with 64 electrodes in total are as follows: **lateral** (Fp1/2, F1/2, F3/4, F5/6, F7/8, Fc1/2, Fc3/4, Fc5/6, Fc7/8, C1/2, C3/4, C5/6, C7/8, Cp1/2, Cp3/4, Cp5/6, Cp7/8, P1/2, P3/4, P5/6, P7/8, Po3/4, Po7/8, O1/2), **anterior-posterior** (Fz, F3/4, F7/8, Fcz, Fc 3/4, Fc7/8, Cz, C3/4, C7/8, Cpz, Cp3/4, Cp7/8, Pz, P3/4, P7/8, Poz, Po3/4, Po7/8), **frontal** (F1/2, F3/4, F5/6, Fc1/2, Fc3/4, Fc5/6), **midline** (Fpz, Fz, Fcz, Cz, Cpz, Pz, Poz, Oz), **parasagittal** (Fp1/2, F1/2, F3/4, Fc1/2, Fc3/4, C1/2, C3/4, Cp1/2, Cp3/4, P1/2, P3/4, Po3/4, O1/2), and **temporal** (F5/6, F7/8, Fc5/6, Fc7/8, C5/6, C7/8, Cp5/6, Cp7/8, P5/6, P7/8, Po7/8).

## *Attentional effects – attended vs. unattended*

### 1) Phoneme probe stimuli

Figure 28 displays group average ERPs to phonemes on the attended and unattended side respectively. A clear P1 component (100-200 ms) could be observed which did not differ largely between both sides (attended vs. unattended). Furthermore, unattended phonemes were less negative between 300 ms to 500 ms after stimulus onset, indicating the existence of the Nd for this experimental condition compared to attended phonemes (see difference waves).

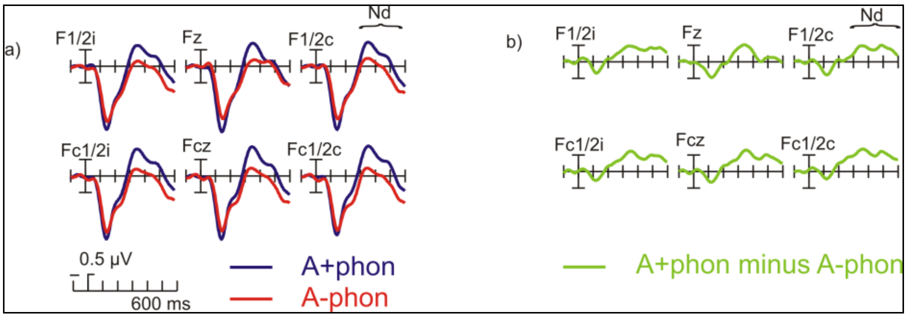


**Figure 28:** (a) **Group average ERPs for auditory probe stimuli (phon = phonemes) on the attended (A+) and unattended (A-) side respectively, in which the ipsi- (i) and contra-lateral (c) sites were accounted.** (b) **Difference waves (attended phonemes minus unattended phonemes) are shown.**

Within this first comparison, attended vs. unattended phonemes, the Nd effect with a more negative-going wave for attended than for unattended stimuli was shown. The effect was predominant at frontal/central electrodes reflected in statistical significance (frontal main effect of Condition: F(1,16) = 11.35; pHF = .003).

The isovoltage map of the difference waveforms (A+phon minus A-phon) underscores the fronto-central scalp distribution of the negative difference (Nd) (Figure 29).

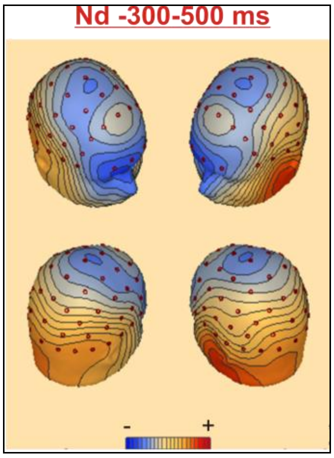


**Figure 29: Isovoltage maps for visualization of the topographical distribution of the Nd effect: difference waveforms (A+phon minus A-phon) are shown within the accordant timeframes (red areas are positive; blue areas are negative). Isovoltage spline interpolation for the 300–500 ms interval was used (BESA).**

2) White noise probe stimuli

In the comparison between attended white noise stimuli and the same stimuli as they were unattended, then a more dominant N300 component for unattended noise was observed on frontal electrodes (Figure 30). Difference waves in the same figure showed this effect – a more negative-going waveform for A-noise compared to A+noise stimuli – in the 300-400 ms time window more precisely. On the other side, no Nd attention effect could be found.

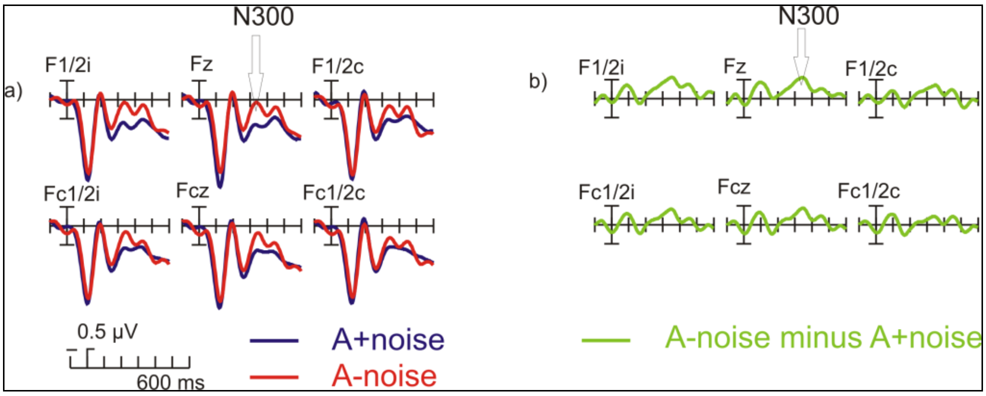


**Figure 30:** (a) **Group average ERPs respectively for auditory probe stimuli (noise = white noise) on the attended and unattended side respectively, in which the ipsi- (i) and contra-lateral (c) sites were accounted.** (b) **Difference waves (unattended white noise minus attended white noise) are shown as well.**

In this comparison, only one difference effect was observed. The N300 component seemed to be larger for unattended than for attended white noise stimuli. The ANOVA yielded a significant main effect of Condition (overall: F(1,16) = 4.62; pHF = .043).

## *Deviant effects – phonemes vs. white noise*

1) Attended phonemes vs. white noise probe stimuli

Differences between both types of probe stimuli phonemes and white noise on the attended are seen in Figure 31. Those ERPs displayed a strong MMN-like N1 component (150-230 ms) for noise relative to phonemes. Note that both types of probes were equiprobable. Later on, a prominent positivity could be seen between 230 ms to 330 ms (P3a-like) after stimulus onset.

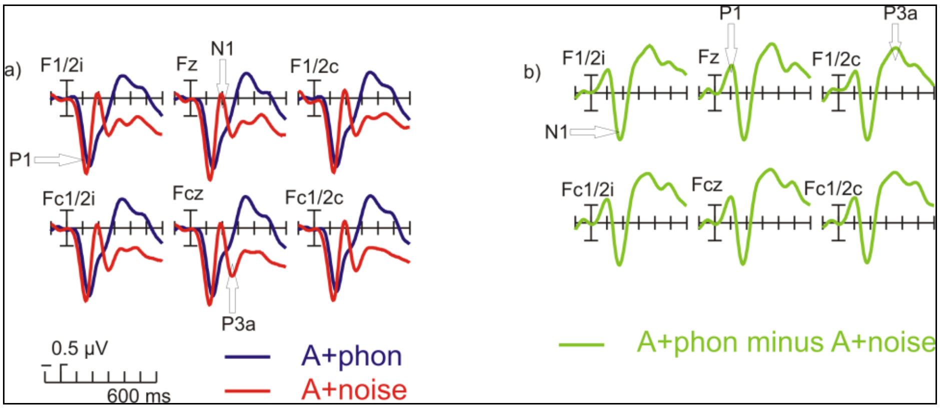


**Figure 31:** (a) **Group average ERPs for auditory probe stimuli (phon = phonemes; noise = white noise) on the attended side, in an ipsi- (i) and contra-lateral (c) view.** (b) **Difference waves (attended phonemes minus attended white noise) are shown as well.**

By comparing attended phonemes (A+phon) with white noise (A+noise) probe stimuli, the N1 effect (150-230 ms) as well as the P3a component (230-330 ms), and the PN effect (230-600 ms) were verified statistically.

With regard to the MMN-like N1 component in the time window of 150-230 ms, statistical conformation could be found. White noise stimuli caused a “mismatch”/N1 compared to phonemes that differed in magnitude over the scalp. The overall (56 electrodes) main effect of Condition became significant (F(1,16) = 5.01, pHF = 0.036). Nevertheless, the condition x anterior-posterior interactions (midline ANOVA: Fpz, Fz, Cz, Pz (F(3,48) = 6.78; pHF = .004) provided a frontal N1/MMN-like distribution

The isovoltage map of the difference waveforms (A+phon minus A+noise) showed the characteristic fronto-central scalp distribution of the N1 component (Figure 32a).

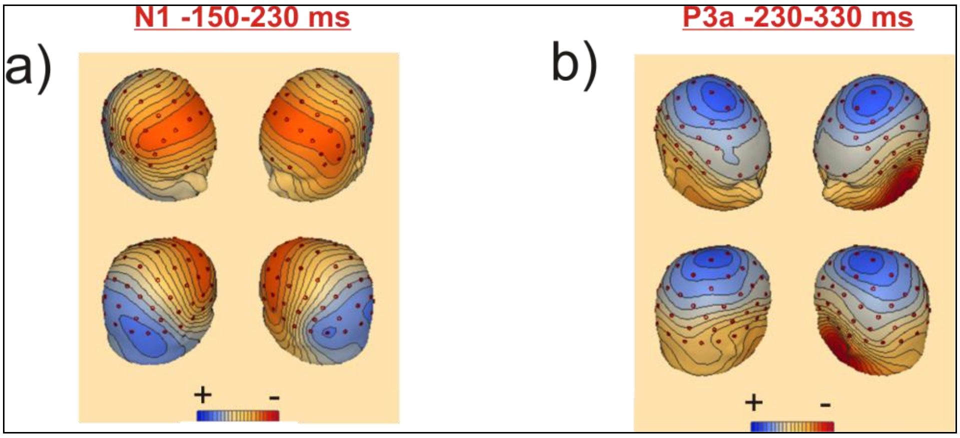


**Figure 32. Isovoltage maps for visualization of the topographical distribution of the** (a) **N1, and** (b) **P3a-like component: difference waveforms (A+phon minus A+noise) are shown within the accordant timeframes. Isovoltage spline interpolation for the respective interval was used (BESA).**

The ANOVA for noise stimuli compared to phonemes at the time window between 230 and 330 ms yielded a P3a-like component for white noise (anterior-posterior main effect of Condition F(1,16) = 4.49, pHF = 0.046). This effect was not predominant at all sites, but mostly seen at central electrodes compared to anterior and posterior electrodes confirmed by the anterior-posterior interaction (anterior-posterior ANOVA: F(2,32) = 10.69, pHF = 0.001).

Isovoltage maps of the difference waveforms (A+phon minus A+noise) underscored the characteristic central scalp distribution for both P3a-like component (Figure 32b).

2) Unattended phonemes vs. white noise probe stimuli

When the same probe stimuli (phonemes and white noise) on the unattended side were compared to one another, the same pattern as for attended stimuli was observed, Figure 33. An MMN-like N1 component (150-250 ms) for noise probes but not for phonemes which were presented with the same probability. Later on, a prominent positivity could be seen between 230 ms to 350 ms (P3a-like) after stimulus onset.

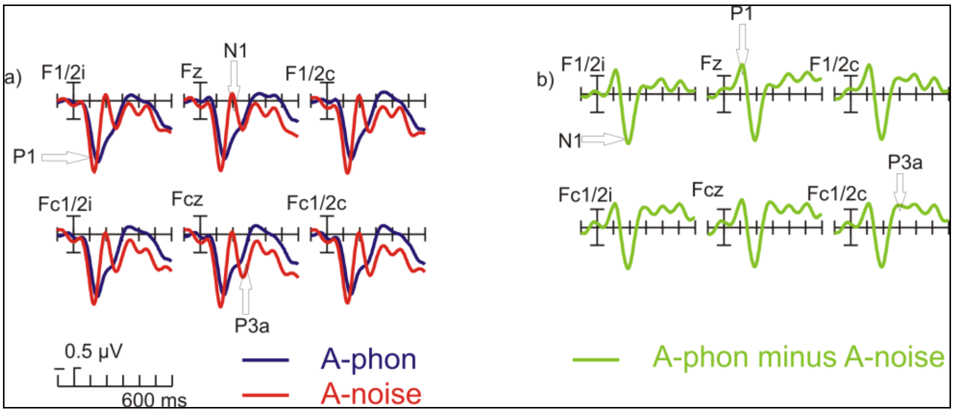


**Figure 33:** (a) **Group average ERPs for auditory probe stimuli (phon = phonemes; noise = white noise) on the unattended side for the fronto-central electrodes, in which the ipsi- (i) and contra-lateral (c) sites were accounted.** (b) **Difference waves (unattended phonemes minus attended white noise) are shown as well.**

As unattended phonemes (A-phon) were compared with white noise (A-noise) probe stimuli, the same statistical pattern as for equivalent stimuli on the attended side appeared. The N1 effect (150-250 ms) as well as the P3a-like component (230-350 ms) reached significance (Table 7).

**Table 7. Significant results (main effect of condition) of pair-wise ANOVAs between white noise and phonemes on the unattended side.**

| components | max. effect | F-value (df) | p-value (pHF) |
|---|---|---|---|
| **N1/MMN-like** | frontal | 15.70 (1,16) | 0.001 |
| **P3a-like** | overall | 4.38 (1,16) | 0.048 |

Isovoltage maps of the difference waveforms (A-phon minus A-noise) evidenced the characteristic frontal scalp distribution of the N1 component (Figure 34a), and showed a central P3a scalp distribution (Figure 34b).

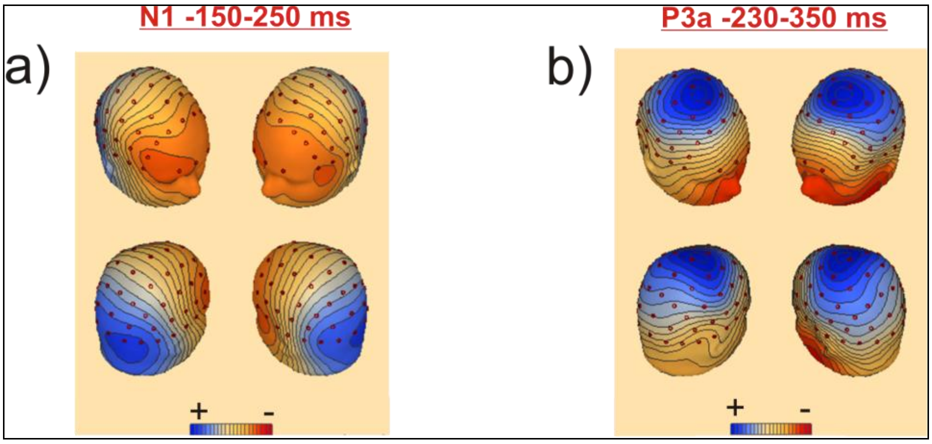


**Figure 34. Isovoltage maps for visualization of the topographical distribution of the** (a) **N1, and** (b) **P3a-like component: difference waveforms (A-phon minus A-noise) are shown within the accordant timeframes. Isovoltage spline interpolation for the respective interval was used (BESA).**

## Discussion

This third study aimed to investigate attention related processes in the brain with a higher (spatial) resolution EEG by using 64 channels instead of the standard 32-channel-set-up. With this method, findings of preceding studies (experiment 2.1 and 2.2) as well as the study in chapter 3 should be supported. Furthermore, at the same time, the question about the sensitivity of the attentional focus for the stimulus' frequency spectrum should be answered.

### ***Attentional effects***

Before we will answer the question about the role of a stimulus' spectral content, let us take a look at a general attention effect first. Again, in contrast to the first study but in agreement with the second experiment of this chapter, a clear and broad attentional effect in form of the widely discussed Nd wave (negative difference; (Araki et al., 2005; Münte, Kohlmetz, Nager, & Altenmüller, 2001; Näätänen & Picton, 1987; Woldorff & Hillyard, 1991)) was observed between A+phon and A-phon probe stimuli (Figure 35). This result confirms that attention facilitates the processing of to-be-attended stimuli by allocating more mental resources to those stimuli (Woldorff et al., 1993). This difference in processing level of a stimulus depending on the attentional status is displayed in a more negative ERP for attended than unattended stimuli. Thus, the current study with two different spatial streams revealed as well the need for selective attention to one of the two spatial channels. While paying attention to one ear, it was necessary to ignore the input from the other ear at the same time. Hence, the processing of spatial information was indispensable. The fact that this Nd wave was more apparent on fronto-central electrodes, let assume that the frontal cortex was highly involved in such attentional selection processes (Jemel et al., 2003).

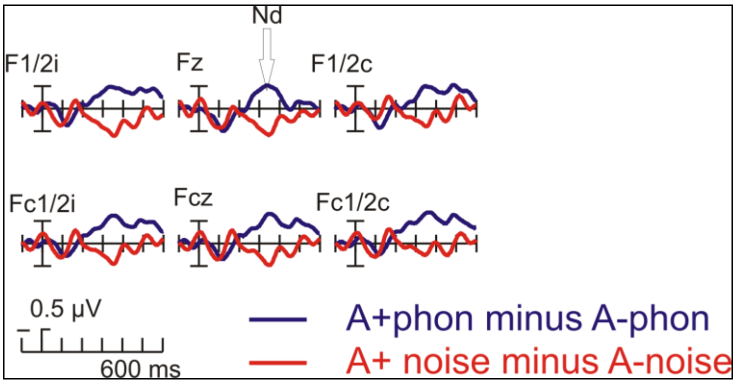


**Figure 35. Difference waves for phonemes (phon) and white noise (noise) in comparison to one another to unveil Nd effects if existent. Only phonemes showed such an attention effect.**

In the comparison of white noise probes (A+ vs. A-) no such an Nd wave as for phonemes was found (Figure 35) indicating a similar processing and the same amount of resources for both attentional conditions (Woldorff et al., 1993). The lack of an Nd effect in this case can be taken as evidence for a selection by a stimulus' spectral content. White noise includes a broad frequency spectrum while speech/phonemes are relatively band-limited at high frequencies. These additional frequencies in the white noise band compared to phonemes' frequencies might have been registered in an early processing stage and led to an exclusion from the attentional focus even at the to-be-attended location. Thus, "attended" white noise probes were processed in a lesser degree equally to unattended white noise. This latter conclusion was confirmed by the absence of an Nd-effect as attended white noise probes were compared to the same as they were unattended.

Additionally, this result would also highly conform to the object based attention theory (see review: Alain & Arnott, 2000; Scholl, 2001) which assumes that all features that belong to an attended object, even though they are not task-relevant, will be processed preferentially. Thus, phonemes (because of physical similarities) and not white noise probes may have been perceived as a part of the to-be-attended prose stream whereas white noise was rather seen as different. This would explain why more mental resources were allocated to this specifically attended story with its phonemes but not to white noise probes.

Apart from a missing Nd, one single component was modulated indeed when A+noise and A-noise were compared. The N300 component was larger for unattended white noise compared to the attended one. White noise stimuli in general could have been very obvious, because they were excluded from the prose stream and thus from attentional focus due to feature differences in contrast to phonemes, and thus they could have initially distracted attention more likely. With regard to this speculation, it is also very explicable why unattended noise probes have caused a larger N300 than A+noise stimuli. A possible associated switch towards the distracting stimulus was spatially larger for unattended than noise probes on the attended side. In other words, attended white noise was on the same side as attention resulting in a smaller distraction shift, whereas A-noise probes were presented on the contrary side of attention requesting a larger shift in space reflected in a larger N300. This is speculation only so far.

The N300 is a not yet well understood component. The precise functional significance remains to be elucidated. Visual findings may parallel equivalent mechanism in the auditory field. McPherson and Holcomb (1992), for example, suggest in their visual study that the N300 is related to the processing of "object-specific information". They found that the N300 is sensitive to object structure information (difference between pseudo-objects and scrambled objects, which violate structural principles). If one generalizes this suggestion and if one therefore assumes that white noise stimuli like in this third experiment because they are non-linguistic, were a kind of pseudo- or scrambled object compared to the residual linguistic stimuli then the N300 could indicate a notice of a main difference within the experimental stimuli. Therefore, it could reflect unconscious analysis of additional information with respect of the spatial location of presentation (stronger effect on the unattended side where a larger shift was necessary compared to the attended stream, as speculated above).

In auditory research regarding the N300, Bostanov and Kotchoubey (2004) reported a N300 in ERPs to contextually incongruous exclamations, associated with emotional affects. They interpreted this component as analogous to the well-known N400 response to semantically inappropriate words. What if the N300 is not necessarily evoked by emotional differences, but context incongruence in general? This interpretation then would be applicable to the present data very well. The nonlinguistic white noise probes do not fit in an otherwise linguistic context with speech streams and phonemes which finally has caused the N300 for noise stimuli. This general assumption of a stimulus' "incongruence" also would conform to the visual study by McPherson et al. (1992). They only talked about violation in an object's structure which basically is nothing else than incongruence between structures of objects. As mentioned above, noise stimuli may have not been processed as though they belong to the prose streams, but instead to another auditory object (see review regarding object based attention: Alain & Arnott, 2000; Scholl, 2001) causing this assumed incongruence.

Later on, McPherson and Holcomb (1999) assumed that N300 may index processes specifically associated with high-level analysis of representations. This means, if the subjects in the current experiment noticed the white noise as a task irrelevant deviant stimulus then they may also have created a special mental representation of it with the intention of ignoring this stimulus. Whenever this to-be-ignoring stimulus appeared – either on the attended or unattended side, a neural response (N300) occurred that shows recognition of this stimulus. This assumption would confirm the lack of a negative difference (Nd) as well. White noise in general was not included in the attentional focus due to being very different to the to-be-attended prose stream whereas phonemes were more embedded into the prose stream and perceived as one auditory object. Furthermore, this would rather argue for an oddball paradigm than an equiprobable design in the present experiment. The phonemes were seen as standards because of the support by the story itself, whereas white noise probe were standing alone and thus perceived as deviant stimuli like in a typical oddball paradigm.

However, this incongruence theory gets more support by a study of Federmeier and Kutas (2001). They reported that only between-category mismatches and not within a category cause an N300 effect. Therefore, the N300 seems to reflect the categorization process. As soon as an incongruent category is involved a N300 is evoked. This led to a simple transfer to the present study in this chapter. It could simply mean, as discussed above, that white noise was noticed to belong to another category (non-lingual) than the prose and phonemes (lingual) at the same presented side. Thus, the N300 is sensitive to categorical-level mismatches. Furthermore, the N300 was also seen as a relatively automatic neural response to perceptual change (West, Herndon, & Crewdson, 2001). According to them, the N300 is modulated by the requirement to monitor the environment for changes (detection of deviant stimuli).

Apart from the fact that those studies, mentioned above, mainly considered pictures or written words, the N300 effect was furthermore reported for non-coherent audiovisual dubbings (audiovisual speech processing) with a modulation of the N300 in amplitude with a larger amplitude for the incongruent condition (Lebib et al., 2004).

The bottom line is that all studies mentioned above argue that the N300 reflects neural activity associated with the detection process of incongruence causing an increment of sensitivity to those noticed stimuli. White noise stimuli in the current experiment could have been noticed and therefore demanded for attentional resources in this timeframe before switching back to the primary task. This claim for resources was confirmed by the likewise P3a component (discussed later on) observed for white noise stimuli compared to phonemes.

### ***Deviant effects***

Within the next comparison – phoneme vs. white noise, whether attended or unattended -, only noise stimuli revealed an N1 component. This deflection acted more like a mismatch negativity (MMN in an oddball paradigm) similar to the N1 component in experiment 2.1. The reason for this difference could be the linguistic vs. non-linguistic character of the probes in the present investigation. Maybe the subjects integrated the phonemes into the prose stream because they are linguistic (Alain & Arnott, 2000; Scholl, 2001). The nonlinguistic noise probes on the other hand might have been excluded. Hagoort and Brown (2000) also reported a reduced N1 in continuous speech compared to reading of sentences. They supposed that the reduction or even absence of the N1 may be due to the continuous nature of acoustical speech signals with no clear physical boundaries between words in the speech waveform. Thus, especially phonemes with the same fundamental frequency at the same location as the speaker’s voice respectively were maybe processed within that speech stream resulting in an absence of an N1. In contrast, white noise could have been separated easily from this speech stream and therefore able to evoke this negativity within the ERP. If so, by assuming an integration of phonemes within the prose streams, white noise stimuli were in minority (“deviant”) compared to those “standard stimuli” (phonemes + prose), similar to an oddball-paradigm. This would explain why the observed N1 for the noise probes in this case equals to a typical MMN, although phonemes and noise probes were presented equiprobable.

For the same reason, the positivity following the white noise N1/MMN deflection can only be the P3a component. The P3b as an alternative assumption for this component can be excluded because the current experimental setting contained task-irrelevant probe stimuli, and a P3b can only be elicit by active attention and a required response (Katayama & Polich, 1998; Sutton et al., 1965). Moreover, it could be shown that a P3a cannot only be evoked by a ‘novel’ sound, but also by ‘typical’ deviant stimuli as long as they are distracting (for an overview see Polich, 2004). The P3a is considered to be related to a switching of attention to a deviant event (Friedman et al., 2001; Münte et al., 1995). Comerchero and Polich (1998) suggested that the P3a may be generated by deviant-standard discrimination rather than by stimulus novelty as first assumed. Additionally, the observed central scalp distribution in the present study is more typical for a P3a than a P3b component especially for task-irrelevant stimuli (e.g. Cycowicz & Friedman, 2004). The latter one shows more often a centro-parietal predominance (Comerchero & Polich, 1999).

However, the N1/MMN (deviance detection (for a review see Picton et al., 2000)) and the P3a (attentional switch (Friedman et al., 2001)) support the initially made conclusion: White noise probes with its wide frequency range are not included in the actual focus of attention, even when presented at the same location. White noise was treated differently on both the attended and unattended side compared to phonemes. Both components have demonstrated that white noise was recognized and processed as a kind of deviant (non-linguistic) stimulus that distracted attention temporarily compared to the otherwise linguistic probes and prose streams. None of those effects would have been observed if white noise features were accounted for the same to-be-attended object (prose) and therefore the same attentional focus.

The fact that there was no difference between attended and unattended MMN or P3a suggests that indeed no attentional modulation occurred. This means that “attended” white noise was not preferentially processed or more elaborated than unattended noise probes. Furthermore, this can be taken as an evidence for the restriction of the attentional focus on the to-be-attended frequency spectrum only. That is, white noise with a much wider frequency range than the to-be-attended frequency band is excluded from the actual focus of attention and thus treated as unattended stimulus even when it was presented on the to-be-attended side.

Moreover, deviant effects as the MMN and P3a besides attentional effects such as the Nd provide evidence that attentional selection may take part in two separate stages. There is an early preattentive stage (e.g. MMN) wherein a stimulus’ feature determines primarily whether more mental resources are provided for its processing. In the later stage of selective attention (attentive, e.g. Nd), top-down processes are more important to include a stimulus into attentional focus. Paavilainen et al. and Winkler et al. (1989; 1998) hypothesized that even when attention is focused on a specific task, our brain pre-attentively browses other sensory input for which it may be of importance to (involuntarily) orient to. Furthermore, these deviant effects would also support the conclusions made for the N300.

### ***Summary***

Taken together, the current study yielded the typical attention effect with facilitating processing of attended compared to unattended stimuli. Furthermore, evidence was found that the attentional focus is restricted to the spectral range of the to-be-attended speaker’s voice only. It does not allow stimuli to pass the filter that have a different frequency spectrum. Hence, stimuli with a wider spectral range such as the white noise were outside of the attentional focus and therefore processed as unattended stimuli even on the to-be-attended side. This kind of restriction with regard to the attentional focus conforms to the finding of the preceding two experiments. Moreover, it could be shown that, besides focusing and sustaining attention to a certain point in space and to specified stimulus features, there is still the ability of the brain to scan pre-attentively the residual sensory input for possibly important stimuli which may be relevant to (involuntarily) orient to. Because of being independent of attention, there is no difference between those deviant detection processes on the attended or unattended side.

# One story out of four – a four-speaker setting as a more complex cocktail party situation

## Chapter 3

### Abstract

In the preceding experiments, it was shown that humans are able to distinguish between two separate but simultaneous streams of speech. The findings suggested that listeners are able to pay attention to the to-be-attended story and that the actual attentional focus is restricted to the features of the attended story only. Even slight changes in fundamental frequency, spectral content or spatial location of probe stimuli caused an exclusion from the attentional focus.

The present experiment investigates the same attentional processes in a more complex cocktail party situation. Four human speakers at spatially separate locations are used. Therefore, the listener needs not only to discriminate between two sides of presentation (between ears) but also needs to differentiate between two speech streams on each side. Again, the unique virtual-spatial representation of different stories and the superimposed probe stimuli ear was based on HRTFs. Task-irrelevant probes were always spatially and frequency congruent to the prose voice that it was superimposed on.

ERPs revealed attention-dependent modulations of the P1, Pd, and Nd/PN- components between attended and unattended stimuli as well as between stimuli on a “target” or “non-target” position. Larger amplitudes or more negative-going components were revealed for attended probes and for phonemes in target streams in general. Thus, the maximum effect was found for target probes on the attended side whereas phonemes in a non-target stream on the unattended side displayed the least brain activity. Attentional effects were comparable to the two-speaker setting.

### Introduction

At a beginning of a cocktail party, when only a couple of people have arrived forming separate groups, it is quite easy to understand a conversation within one of these groups. The background noise is also still relatively low. However, at the party’s climax, the noise level has increased tremendously, and separate chatting groups are drowning each other out, so that it is very difficult to follow the content of a single conversational circle.

The challenge is to actively select one of those simultaneously ongoing conversations and to ignore or to filter out all others. In order to manage this task, the focus of attention should be quite narrow to allow processing of only these stimulus features matching the to-be-attended object (e.g. the speech of a certain guest at a certain location) to pass the attentional filter. In the previous experiments (chapter 2) an efficient restriction of the attentional focus has been demonstrated for a two-speaker situation. The question whether those selection mechanisms would be still effective in a similar manner in a more complex situation can only be answered if a more realistic situation with more than two human speakers is applied. Therefore, individual HRTFs were used to simulate a virtual auditory environment with four distinct prose streams (uttered by different human speakers). Each of those prose streams was superimposed by probe stimuli (phonemes: frequency and spatially congruent) each.

As demonstrated in chapter 2, attentional selection of a stimulus is primarily based on the location where it is presented. Stimuli from another location are processed to a lesser degree than the actual to-be-attended object as evidenced by the ERP effects to the probe stimuli. In subsequent processing stages, changes in fundamental frequency or in the frequency spectrum lead to an exclusion from the focus of attention even though a stimulus is presented at the to-be-attended location. According to Bregman (1990), stream segregation is a basis to identify meaningful sounds correctly.

The general question addressed in this experiment is how humans are able (brain activity wise) to select an object’s attention in a more complex situation. In a simple dichotic listening task, it is only confirmed that if a stimulus is selected once and passes the selective filter then more resources are available for its processing (more intense processing of attended compared to unattended stimuli) (Suzuki et al., 2005; Wickens et al., 1983). Would it be the same in a more complex setup? It might be argued that with an increasing number of concurrent auditory streams and therefore with increasing noise level it is harder to segregate between the concurrent auditory streams. To my knowledge, there is no study using more than three concurrent ongoing prose streams at the same time investigating attentional neural correlates.

The goal of the present study was to extend earlier efforts in a simple cocktail-party-situation evaluating how directing attention affects the ERPs elicited by irrelevant auditory probes on either the attended or the unattended prose stream. Furthermore, it should be clarified whether findings made in a simple dichotic setting are transferable and can be generalized to more complex situations or whether complex situations require more or different factors in stimulus processing.

## *Experiment 3.1:*

### Method

#### Subjects

The same 32 subjects as in the previous two-speaker experiments participated as well in the present one.

#### Stimuli

To reduce eye movements, subjects fixated on a cross on a computer screen located in front of them during the recording. The experimental set-up comprised four different stories placed on four distinct spatial locations in a hemispheric manner at the same time. The virtual spatial angles of the four narratives were 60° left, 15°left, 15° right and 60° right respectively. The stories were spoken by four different native speakers (two male and two female speakers; see appendix A1) and presented via headphones. Only the 60° stories were attended with left and right being focused on in separate runs. Thus, the 60° location will be called target location from now on. The stories at 15° on both sides were never targets (non-target stories). Otherwise, the procedure and stimulus preparation was the same as for the previous experiments – including HRTF modulation, and usage of probe stimuli and their presentation.

The superimposed probes (phoneme “da”; 100 ms; also convolved with individualized HRTFs) with the same fundamental frequency and the same spatial angle as the speaker’s voice of each prose were employed. Hence, there were four different probe stimuli depending on the presented location: A+S+L60, A-S-L60, A-S+L15 or A-S-L15. A+/- stands for an attended or unattended stimulus, S+/- represents the side with regard to the attended stream (S+/-: same or opposite side), and L60/15 refers to the location (60° or 15°).

For this experiment, 4000 probe stimuli in total (500 probes for each of the eight conditions; see Table 8) were presented in 4 runs (2 attended right, 2 attended left) lasting about 11 min each. The subject’s task consisted of actively listening to one of four parallel stories as instructed before each run, and to keep the content of that story in their mind. No button press was requested. After each run, the participants were asked several questions about the attended story to make sure that they really directed their attention to that story only. The entire session had a length of 1.25 hours (excluding the EEG setup of approximately 30 minutes).

**Table 8: Distribution of probe stimuli within the eight experimental conditions. S+/- represents probe stimuli either on the same side as the attended prose or on the opposite side. The spatial location is encoded in L60 (target story) or L15 (non-target prose). A+ stands for “attended” and A- for “unattended”.**

| | **Left** | | | **Right** | | |
|---|---|---|---|---|---|---|
| | ***A+*** | ***A-*** | | ***A+*** | ***A-*** | |
| | ***S+*** | ***S+*** | ***S-*** | ***S+*** | ***S+*** | ***S-*** |
| **L60** | 500 | | 500 | 500 | | 500 |
| **L15** | | 500 | 500 | | 500 | 500 |

## EEG-Recording and data analysis

The EEG was recorded by using an elastic-cap with integrated tin electrodes (positions: Fp1, Fp2, F3, F4, F7, F8, C3, C4, P3, P4, Fpz Fz, Cz, Pz, T7, T8, Fc5, Fc6, Fc1, Fc2, Cp5, Cp6, P7, P8, P3, P4, Po1, Po2, O1, O2 in the international 10-20 system (Jasper, 1958)). All other parameters of EEG-recording were identical to those used in experiments reported in chapter two.

The ERPs were averaged separately for attention condition (attended (A+) or unattended (A-) probes), side (the same side (S+) or opposite side (S-) as the attended target story), spatial angle (60° (target) or 15° (non-target)), and location (left or right). After preliminary analyses had indicated no difference between effects to left and right-sided probe stimuli, their ERPs were collapsed to yield waveforms for electrode positions ipsi- and contralateral to the location of probe presentation.

The ERPs were generally quantified by mean amplitude measures, in some cases by local peak latency measures (mentioned separately in the text), and the resulting data were subjected to repeated measures analyses of variance (ANOVAs). If an overall two-way ANOVA with factors condition, and electrodes (all 30 head sites) did not reveal any significances, some regional ANOVAs were calculated to find potentially smaller effects at certain electrode groups or single head sites: with the factors condition, and site(s); if possible with additional factors laterality (2 levels), and anterior-posterior each (3-5 levels).

## Results

### Subjects

The results of twenty subjects (17 women, mean age 27.0, range 20-35) out of the original 32 were statistically analyzed eventually. The data of 4 of the remaining subjects had to be discarded, because of insufficient behavioral performance. Another 8 subjects were lost because of too many artifacts or technical failures. If more than one third of epochs were rejected due to artifacts (exclusion criteria), a person was excluded from further analyses.

### Behavioral data

It can be concluded from the questionnaire that this experiment was not very easy to accomplish. In this case, variation in individual capabilities became more evident. Not all subjects were able to concentrate on the specified story and to ignore all others completely. This was evident in the level of correct responses (average 58% correct; range: 18% to 84%; see appendix A3d).

### Electrophysiology

The following ERP figures display group average ERPs (n = 20) and difference waves for attended probes at the target location (A+S+L60), unattended probes at target location (A-S-L60), unattended stimuli at non-target location of the same side as the attended stream (A-S+L15), or unattended stimuli at non-target location of the opposite side of the attentional focus (A-S-L15). ERPs from left and right stimuli are pooled to yield ipsi- (i) and contra-lateral (c) derivations.

In general, the data was statistically evaluated in several time-windows separately for each subject, condition and electrode site. These were entered into repeated measures analyses of variance. Significant results ($p < .05$) are shown.

In the following, attentional and spatial effects were analyzed separately. If significant, ANOVAs with factors condition (2 levels: a pair of either A+S+L60, A-S-L60, A-S+L15, or A-S-L15), and electrode site (30 levels) as the within-subject factors are shown. Separate regional ANOVAs for selected electrode groups (frontal, midline, parasagittal, and temporal) for regional effects only are also reported: with factor condition, and if possible with factor laterality (left, right), and anterior-posterior.

## *Attentional effects – attended vs. unattended side*

### 1) Phonemes at the target story's location

First, let us take a look at the comparison between probe stimuli on both target positions – one attended and one unattended (A+S+L60 vs. A-S-L60). Similar to the results of experiment 1 of chapter 2, instead of a clear fronto-central Nd effect, Figure 36 shows first an anterior positivity for attended probe stimuli, whereas the ERPs to unattended probes are more negative than the same but attended probes between 280-350 ms. This positive difference (labeled Pd to distinguish this component from the Nd) can be better seen in the difference waves respectively. Difference waves were obtained by subtracting unattended stimuli from their equivalent attended probes. The Pd was followed by a negative difference (Nd) between attended and unattended L60 stimuli in the time window of about 400-500 ms. This Nd is mainly prevailing on centro-parietal sites compared to the more frontal oriented positive deflection.

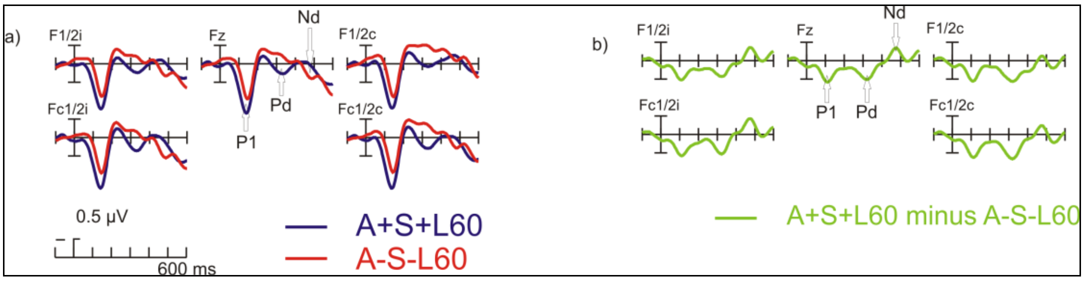


**Figure 36:** a) **Group average ERPs for probe stimuli at the target location (L60) on the attended (A+S+) and unattended (A-S-) side respectively, whereas ipsi- (i) and contra-lateral (c) sites were accounted for both, right and left probe stimuli combined.** b) **Difference waves (attended minus unattended stimuli).**

Statistical analyses yielded support for these visual observations of this first comparison between the attended and unattended probes at the target location (L60): 280-350 ms (Pd), and 400-500 ms (Nd effect) after stimulus onset.

A significant main effect of Condition (overall (30 electrodes): $F(1,19) = 5.65$, $p = 0.0282$) was obtained for the positive difference with a more positive-going waveform for A+S+L60 probes compared to A-S-L60 stimuli. Similar to experiment 2.1, the Pd differed significantly when the first half of the experiment was compared to the last part. The positive deflection was reduced towards the end of this experiment; especially at frontal sites (Fp1/2, F3/4/7/8, Fc1/2/5/6: F(1,19) = 5.95, p = 0.0247).

With regard to the Nd, there was also a significant ANOVA with a clear overall main effect of Condition (F(1,19) = 5.22, p = 0.0340) indicating a negative deflection for the A+S+L60 condition compared to A-S-L60 probes. Furthermore, an anterior-posterior effect between conditions with a centro-parietal predominance was observed (midline electrodes (Fpz, Fz, Cz, Pz): F(3,57) = 4.02, p = 0.0251).

### 2) Phonemes at the non-target story's location

The comparison between phonemes at the non-target story's location on both sides (S+ vs. S-) also displays a slight attentional effect (Figure 37). The first endogenous modulation between both conditions (A-S+L15 and A-S-L15) was seen between 170 and 330 ms. The A-S+L15 showed a more negative-going wave form than the A-S-L15 condition. This negative difference (Nd) became clearer in the respective difference waves.

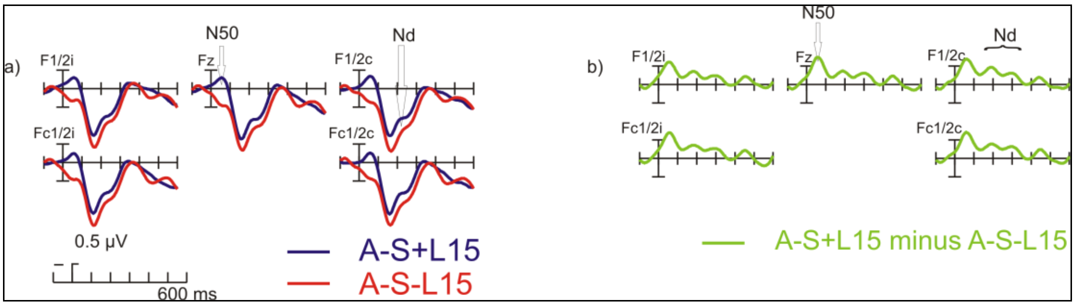


**Figure 37:** a) **Group average ERPs for probe stimuli at the non-target location (A-; L15) on the attended (S+) and unattended (S-) side respectively; ipsi- (i) and contra-lateral (c) sites were accounted for both, right and left probe stimuli.** b) **Difference waves (non-target (S+) stimuli minus non-target (S-) stimuli).**

Statistical analysis revealed a significant Nd effect between A-S+L15 and A-S-L15 probe stimuli. The ANOVA for the time window between 120 and 240 ms after stimulus onset attained a significant overall main effect of Condition (F(1,19) = 6.96, p = 0.0162). Thus, the ERP waveforms of A-S+L15 were more negative than the A-S-L15 ERPs.

## *Spatial effects – target location probes vs. non-target location probe stimuli*

### 1) Attended side

In Figure 38, both kinds of probe stimuli on the same side as the to-be-attended prose stream are plotted. A+S+L60 were compared to A-S+L15 resulting in main ERP differences in two time frames. It is the PN that can be divided into an earlier (PNe) and a later (PNl) subcomponent (see difference waves). The PNe between 160 and 300 ms evidences a more negative-going waveform for A+S+L60 than for A-S+L15 probes, which was slightly more apparent at contralateral electrodes. The later PN component, PNl, occurred between 400 and 510 ms and shows again a negative deflection for probes at the to-be-attended story's location compared to the probes at the non-target story's location. Therein, the PNl was larger on the contralateral side. Nevertheless, both PN subcomponents were more predominant at anterior sites compared to posterior electrodes.

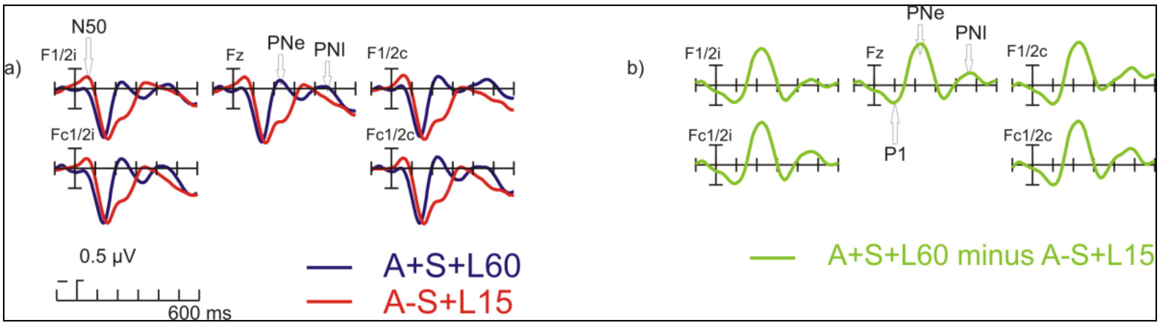


**Figure 38:** a) **Group average ERPs for probe stimuli at the target location (L60 (A+)) and non-target location (L15 (A-)) on the attended (S+) side; ipsi- (i) and contra-lateral (c) sites were accounted for both, right and left probe stimuli.** b) **Difference waves (attended target minus non-target stimuli).**

The two different effects (PNe and PNl) were statistically confirmed by comparing A+S+L60 and A-S+L15 probe stimuli. In the PNe latency window (160-300 ms), the ANOVA yielded a significant main effect of Condition (overall: F(1,19) = 5.16, p = 0.0349) indicating that A+S+L60 stimuli elicited a negative deflection compared to stimuli at the unattended location at the same hemi-side/ear in space. The slight contralateral shift did not become significant.

The second PN subcomponent, PNl (400-510 ms), unveiled a focal significant difference. There was a main effect of Condition at fronto-central sites (Fp1/2, F3/4, F7/8, Fc1/2, Fc5/6: F(1,19) = 4.50, p = 0.0473) reflecting a more negative ERP for A+S+L60 compared to A-S+L15 probes in this time frame. A contralateral predominance was confirmed by a significant condition x laterality interaction (frontal (Fp1/2, F3/4/7/8, Fc1/2/5/6): F(1,19) = 4.76, p = 0.0437).

2) Unattended side

When the two different stimuli (L60 and L15) on the unattended side were compared (Figure 39), then a similar ERP pattern as on the attended side could be observed. Again, two subcomponents of the PN were seen. The PNe was present at 190 to 340 ms with a contralateral emphasis, especially at frontal sites. The later component, PNl, was salient in the time window of 400-450 ms; predominant on anterior and contralateral electrodes. Difference waves highlight these effects.

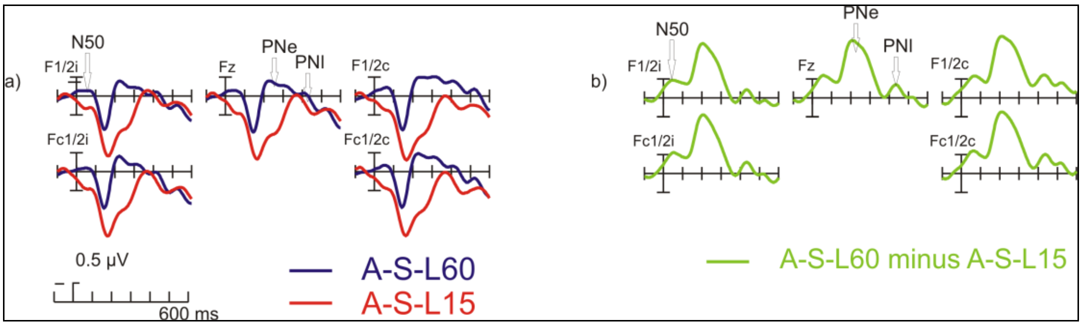


**Figure 39:** a) **Group average ERPs for probe stimuli at the target location (L60) and non-target location (L15) on the unattended (A-S-) side; ipsi- (i) and contra-lateral (c) sites were accounted for both, right and left probe stimuli.** b) **Difference waves (unattended target minus non-target stimuli).**

Statistical analyses revealed as well significant differences between A-S-L15 and A-S-L60; similar to the differences between both conditions on the attended hemispace (Table 9).

**Table 9. Significant results (main effect of condition) of pair-wise ANOVAs between stimuli from the target and non-target location on the unattended side.**

| components | max. effect | F-value (df) | p-value (pHF) |
|---|---|---|---|
| **PNe** | overall | 19.58 (1,19) | 0.0003 |
| **PNl** | Frontal (FP1/2, F3/4/7/8, FC1/2/5/6) | 4.75 (1,19) | 0.0422 |

## Discussion

In this more complex auditory experiment, four instead of two simultaneous stories were presented to investigate whether similar underlying attentional selection processes as in simpler (two-speaker) settings can be found. Additionally, attentional effects in dependence on the spatial distance to the to-be-attended location were examined.

The stories were virtually arranged in a semi-circle ahead of each subject. Otherwise, the procedures were identical to preceding experiments (chapter 2). Probe stimuli (equiprobable) consisted of simple “da” phonemes matching the fundamental frequency and spatial location regarding the prose stream that they were superimposed on. Only one of the two stories at the outer end of the virtual hemi-circle (60° right or left) needed to be attended to. Thus, phonemes at the target’s location (L60) could reach two distinct states of attention (attended and unattended) whereas non-target probes (L15) were always unattended, even on the same side as the to-be-attended target story. All four kinds of probe stimuli differed in spatial location and also in the fundamental frequency from one another, but not in probability of occurrence.

The same attentional effects were observed as in the preceding three simpler designs and as in previous studies without HRTFs. There was an Nd effect between attended and unattended probes – for either target or non-target positions. A+S+L60 stimuli compared to A-S-L60 phonemes showed an additional effect, the positive difference (Pd) preceding the Nd.

When target probes were compared with non-target stimuli on either side, a broad PN with an earlier and a later subcomponent could be seen.

### ***Target location***

The main aim was to compare the attended and unattended probe stimuli at the target story's location on both sides (A+S+L60 vs. A-S-L60). Similar to the first experiment in the previous chapter (equiprobable frequency congruent vs. incongruent probe stimuli), a positive difference (Pd; 280-350 ms) with a more positive A+S+L60 waveform than the A-S-L60 ERP was present. This positive deflection also occurred with a frontal to fronto-central scalp distribution as in experiment 2.1. As speculated in the first experiment of this dissertation, this positivity could be evidence for a first active (top-down; endogenous) selection process, or a reflection of a rejection process, an active inhibition or suppression of the unattended stimulus. Nevertheless, the well known Nd (negative difference (Jemel et al., 2003)) finally dominated the observed Pd and could indicate a beginning shift back to the to-be-attended stream. The Nd's onset then would reflect the time at which a particular feature or object was discriminated and selectively processed according to its attentional level. Again, as in the first experiment, the latency of the following Nd component is much later than normally reported in the literature (about 100 ms or even earlier (D. L. Woods & Alain, 1993) instead of the current 400 ms onset time) because of the Pd occurrence first.

Not much is known about this Pd component and its possible coactions with other components. Therefore, only speculations can be made at this point. However, the same question as in experiment 2.1 arises: Why was the Pd never found in other auditory stimulations with continuous speech and superimposed task-irrelevant probe stimuli (e.g., Teder et al., 1993) similar to the present ones? And why did the Pd only appear in two out of four similar experimental settings whereas the second and third experiment did not show such a Pd-Nd combination? Differences in task characteristics may not be sufficient to evoke a Pd (see experiment 2.1: varying stimulus probability or features). The stimuli were selected carefully, so that the speakers' voice of simultaneously presented stories had similar characteristics (none of them dominated in intonation or volume). Therefore, this argument of different stimulus material is not very probable as seen in Beer and Röder's study (2005). Rather, the working memory load in a certain task could be the decisive factor for a Pd elicitation. As soon as two modalities were important like in Beer and Röder's research the working memory is more used to its capacity. The same would be true for a more complex or complicated context such as spatially close presented stories with the same speaker's gender in the current investigation. In general, previous auditory selective listening tasks were basically quite simple and less demanding regarding the working memory, which could be the reason why no Pd for this modality was found yet.

Moreover, it could be argued that the Pd-Nd complex only occurs during a new and unfamiliar auditory setting wherein a subject is not able to revert to automated process mechanisms to relieve the working memory. Experiment 2.1, which was presented as the subjects' first experiment of the whole multi-experimental study, showed both selection effects whereas the second and third investigation reported in chapter 2 elicited only a clear Nd effect. As subjects become trained on this kind of task, which was common from experiment to experiment, the Pd ceased. One can speculatively conclude that, the Pd may be the result of unfamiliarity with a very demanding dichotic listening task.

However, why is the Pd-Nd complex seen in this experiment which was conducted as the subject's third experiment, before experiment 2.3 was measured? Nevertheless, the argument of unfamiliarity with a task could be still valid. All investigations of chapter 2 differed from the present one in one important characteristic: the number of concurrent and simultaneously presented stories (two vs. four). The more complex cocktail-party situation of this current study could have been a new and untrained setting for all subjects, despite their experience with the two-speaker setting beforehand. That explains the presence of the Pd, which was assumed to represent novelty or unfamiliarity in this more complex experiment. This hypothesis was also confirmed by a decreasing positivity amplitude towards the end of the experiment. Subjects got used to the new demands with time.

The doubling of presented simultaneous prose streams may have created an exponential increase in required selective attention skills. The drop-out rate due to task difficulty also shows that not all subjects were able to effectively select the attended story above the background noise of the residual three prose streams, whereas two speakers at once were easier to distinguish for all subjects. The differentiation of selective attention skills between subjects within a more complex situation could be seen as evidence that the present experiment was more difficult and therefore new and untrained with regard to its task demanding processes. No coping strategies for this kind of situation might have been developed so far. Thus, the occurrence of the Pd preceding the Nd for this relative complex and demanding set-up can be explained. Moreover, the assumption that the Pd-Nd complex is restricted to unfamiliar settings only, in which no strategies for stimulus processing and for the relief of the working memory load are available, was also confirmed by these findings.

Nevertheless, the presence of the Pd and Nd as attentional modulations in the present study indicated that subjects had indeed directed their attention to the to-be-attended story (fundamental frequency and spatial location of the speaker's voice) besides three other different and concurrent prose streams. It also shows that more brain resources were allocated to process the to-be-attended compared to unattended stimuli (M. Giard et al., 1988; Suzuki et al., 2005; Woldorff et al., 1993).

### ***Non-target location***

Another comparison between the attended (S+) and unattended (S-) side concerned probe stimuli at the non-target story's location (L15). Both S+L15 and S-L15 conditions were always unattended (A-) compared to the to-be-attended story (A+S+L60). Nevertheless, a slight attentional modulation between ERP waveforms of the two non-target probes could be observed. A small Nd effect (endogenous) was found with a more negative-going waveform for A-S+L15 than A-S-L15 stimuli, and with an emphasis at central electrodes. This result argues for an advantage in stimulus processing by being presented on the same side where the actual attention is focused on (Näätänen & Alho, 2004). In contrast, the same non-target probe stimuli on the opposite side of the attentional focus are rather attenuated (evidenced by the Nd).

Usually, this Nd effect would indicate that non-target stimuli on the attended side were still in the attentional focus. The preceding studies, however, have demonstrated that spatial deviants were excluded from focus. This contradicting result on first sight might be explicable by peripheral versus central resolution. One has to keep in mind that the first three experiments employed rather peripheral stories with even more peripheral spatial deviants. In contrast, the non-target location of the present fourth experiment was closer to azimuthal areas than the target stories. This would suggest the following for our data. Even though non-target streams were out of focus, but because of their vicinity to azimuth, they have been analyzed a bit more distinct between attended and unattended side. Spatial localization is most precise at azimuthal positions and gets weaker in the auditory periphery (W. Teder-Sälejärvi & Hillyard, 1998; W. A. Teder-Sälejärvi et al., 1999). Teder-Sälejärvi, however, reported an attentional gradient around attended stimuli only, but did not compare unattended stimuli in both auditory hemispaces. Maybe, a differentiation between unattended stimuli on both sides appears towards azimuth. Nevertheless, this can only be speculated as no study has investigated this possible phenomenon so far.

Furthermore, non-target phonemes in the present study were integrated into a separate speech stream each in contrast to the stand-alone spatial deviant in the preceding experiments. Additionally, they were spoken in opposite gender's voice. Thus, spatially deviating stimuli differed not only in location, but also in the speaker's fundamental frequency and gender. These characteristics may have influenced the processing of non-target stimuli as a separate speech class and therefore different attentional trace causing an Nd effect for non-target probes. To validate the present results, further studies addressing the question of differences between unattended stimuli around central and peripheral locations, or studies using the same speaker's gender for all four speech streams are needed.

### ***Target vs. non-target location***

#### a) Attended side

What about both kinds of probe stimuli (L60 and L15) on the same side as the attentional focus? Are they processed similarly, because they are at the attended side (S+), or differently, because they are either in (L60) or outside (L15) the focus of attention? Answers to those questions were unveiled by comparing A+S+L60 probes with A-S+L15 stimuli. According to preceding experiments of this dissertation, an assumption would be that phonemes within the second, but unattended prose streams on the same side in space differed too much (spatial angle and fundamental frequency of the speaker's voice) from the to-be-attended stream to be processed at the same rate. Evidence for this postulate was provided by the PN effect with its two subcomponents: PNe and PNl. The early and late PN showed a more negative ERP for A+S+L60 than for A-S+L15 stimuli confirming that the later stimuli received less brain resources for their processing than stimuli within the actual attended prose stream (Araki et al., 2005). This also reflects that the probe stimuli from the non-target story's location are outside of the attentional focus and therefore treated as unattended stimuli. The enormous role of a stimulus' spatial location in determining whether it will be included into the attentional focus was confirmed in experiment 2.2 of this doctoral dissertation.

The observed early PN and late PN, equally to the Nd subcomponents, are distinguishable by different functional meanings (Näätänen, 1992). The early PN simply reflects a temporary feature recognition system. It determines the suitability of the stimulus for further processing. Therefore, the basic physical characteristics of each stimulus are compared to a template stored in the memory. As soon as a stimulus occurs that differs from the template, the matching process terminates. The early PN is also thought to be really closely related to the discriminability of attended and unattended sequences (Hansen & Hillyard, 1980). This assumption would suggest that both kinds of probe stimuli on the same attended side might have been easy to discriminate, because they did not differ only in the spatial distance, but also in the fundamental frequency. This may explain the clear PNe between A+S+L60 and A-S+L15.

The second PN subcomponent, the PNl, is associated with selective rehearsal and the maintenance of the attentional trace (Näätänen, 1990). The PN is a generic term for subtraction waveforms that reveal target-related processing. The Nd, on the other hand, is a more precise measure making use of the same physical stimulus with and without attention directed to it. Thus, the underlying assumptions for the function of PN and Nd are similar. Nevertheless, due to the findings of an early and late component each, Näätänen developed the "attentional-trace" model of selective attention (Näätänen, 1982, 1990, 1992). In this model, the early PN represents an initial selection by comparing a sensory input with and attentional trace in auditory cortex. The late PN on the other hand is thought to be related to a frontal mechanism to control and maintain the attentional trace (M.-H. Giard, Fort, Mouchetant-Rostaing, & Pernier, 2000). In general, the observed PN effect confirms that only A+S+L60 probes are in the attentional focus and therefore processed preferentially compared to the A-S+L15 stimuli (Alho, Donauer et al., 1987; Alho et al., 1994; Hansen & Hillyard, 1980, 1983).

Apart from this PN observation between probes from the attended hemispace, there was another distinct feature: Probe stimuli from another location than the attentional focus did not evoke a N1/MMN-like or P3a component as reported in all previous experiments in chapter 2. How can this difference be explained? The experimental designs were similar. The only distinction regards the number of simultaneously presented speech streams. In chapter 2, only two speakers were present at a time. Thus, all spatial deviants were perceived as being outside of each stream – stand-alone. In the present four-speaker experiment, however, spatial deviants were integrated into another separate speech stream. Therefore, they were not outstanding anymore, but rather "masked". This latter fact could have prevented a deviant effect for spatially different stimuli in the current case compared to previous experiments.

### b) Unattended side

The same comparison between L60 and L15 stimuli was conducted for the unattended side (opposite side to the attentional focus). Thereby, the PN effect with its two subcomponents (early and late PN, Näätänen, 1992) was found as well. This more negative-going waveform for A-S-L60 compared to A-S-L15 stimuli reflects more allocated mental resources for phonemes at the unattended target location. Thus, there is no "all-or-nothing" principle, but instead a gradation from slight to extreme attentional dependent inhibition of unattended information. Attention effects decrease as a function of distance to the attended location (Teder & Näätänen, 1994; W. Teder-Sälejärvi & Hillyard, 1998; W. A. Teder-Sälejärvi et al., 1999).

The arguments accounted for the PNe and PNl occurrence mentioned in the previous paragraph, as A+S+L60 and A-S+L15 probes were compared, hold their validity in the current comparison as well. Both kinds of probe stimuli (L60 and L15) differed largely to cause a large PNe effect. The experimental setup is also still the same, so that the need to monitor a possible selected auditory object within this complex cocktail-party situation legitimates the occurrence of the PNl (M.-H. Giard et al., 2000; Wijers et al., 1996).

Moreover, this similar pattern between target and non-target phonemes on either side – attended or unattended – suggest a similar processing of stimuli from spatial locations on the attended side and paralleling position on the unattended side. The big differences in stimulus processing are seen on the same side between different spatial angles. Thus, a speculation would be that the spatial hemispheres (left, right) may simply mirroring each other and therefore causing a similar processing of stimuli appearing at a paralleling position.

Furthermore, attention may be able to increase the sensitivity of one hemisphere/side compared to the other one. This could explain the observed hierarchy in allocation of mental resources for stimulus processing. The most preferentially processed stimulus was the one presented at exactly the same location as the attentional focus (A+S+L60), followed by the stimulus that occurs at the same/paralleling position on the opposite side (A-S-L60) because it is less contrasting than a largely and spatially varying stimulus within the same hemisphere (like the L15 probe). Nevertheless, between both paralleling conditions a clear attention effect was observed. The next stimulus within the processing hierarchy belongs to the A-S+L15 condition which does not represent a stimulus from a parallel position but a stimulus from another spatial angle at the same hemisphere in space.

The stimulus with the fewest allocated resources is the A-S-L15 probe that differs in spatial hemisphere as well as spatial position with regard to the attentional focus. Thus, this latter stimulus is the most contrasting probe and therefore processed less than any other stimulus in this hierarchy. This again shows the importance of a stimulus' spatial location for its processing and selection (Sonnadara et al., 2006).

### ***Summary***

Taken together, attentional effects between attended and unattended probe stimuli were found in line with the literature and the results of the three preceding experiments in chapter 2. Thus, these results confirm that the underlying processes within a more complex cocktail-party situation are roughly the same as under a simpler condition.

Moreover, stimuli that vary in spatial location compared to the to-be-attended story show a reduced processing on both the attended and unattended side. This indicates that they are all processed as unattended stimuli, even when presented on the attended side. Interestingly, there is a hierarchy in stimulus processing for those spatially deviant probe stimuli arranged in a semi-circle. This hierarchy does not simply follow the semi-circle arrangement or as a function of deviance relative to the to-be-attended story. In contrast, it is more like a zigzag gradation between two auditory hemifields.

The highest processing priority is given to stimuli within the focus of attention. The stimulus at the non-target location on the unattended side will get the least processing power.

Hence, the environmental space must be divided into two hemifields (left and right) allowing stimuli at the same equivalent position on both hemispheres to be processed similarly. Only attention may increase the sensitivity of one hemisphere causing attentional effects between parallel locations of both hemispheres.

Stimuli with a different spatial position compared to the focus of attention are processed less. Nevertheless, the spatially deviants in the attended hemispace experience a little attentional advantage over its paralleling stimulus in the other hemifield.

# Summary and Conclusions

## Chapter 4

*One of our most important faculties is our ability to listen to, and follow, one speaker in the presence of others. This is such a common experience that we may take it for granted; we may call it "the cocktail party problem." No machine has been constructed to do just this, to filter out one conversation from a number jumbled together.* – (Cherry, 1957).

Before a speech stream can be followed and completely understood in a cocktail-party like situation, the listener has to select one stream first and to ignore other competing ones. Thus, two main problems have to be solved by the listener: first, the problem of segmenting the entire incoming auditory signal into separate and distinguishable sound sources or auditory objects, and second, the problem of selecting one stream while filtering out the others. The latter problem cannot be solved without a successful solution for the first problem.

The present study employed a multi-speaker setting in a realistic environment to investigate brain activity during selective listening. Subjects are clearly able to spatially segment incoming information streams and to selectively attend to just one of them. The most important aim of this dissertation, however, is to find out how the brain distinguishes between this competing information. There is much research on auditory selective attention, which is based on very simple dichotic listening tasks rather than a more valid free-field stimulation. Thus, the overall question for the current investigation was how humans solve these problems under more realistic conditions, in both a simple cocktail-party situation and in a more complex setting.

### *Summary: General effects*

Four experiments in total (chapter 2 and 3) were conducted: Within a two- or a four-speaker setting, attentional effects in a more natural auditory environment were examined by using HRTFs for a virtual spatial stimulus presentation. In these experiments, it was shown that only attended stimuli are processed to the full extent whereas unattended stimulus processing was attenuated. This attenuation does not follow an all-or-none law. Unattended stimuli in the attended hemispace were processed differently compared to unattended stimuli in the to-be-ignored hemifield. In general, two different effect types were found: attention effects (Nd – between hemifields) and deviant effects (MMN, and P3a, or RON – within each hemispace); see Table 10. Furthermore, these different effects have been shown as well that selection takes place in two different stages. There is an earlier, still preattentive selection (Pd, N1, MMN) and a later attentive filtering (Nd, P3a, RON). The first selective stage seemed to be much wider whereas the later stage was more finely tuned because different deviants elicited either a P3a or a RON, and not both. This observation matches the results of Teder-Sälejärvi et al. (1998; 1999) pretty well.

**Table 10: Overview: components elicited by specific auditory conditions**

| | | P1 | Nd | PN | N1 | MMN | P3a | N300 | RON |
|---|---|---|---|---|---|---|---|---|---|
| *Congruent* | Standards (A+ vs. A-) | x | x | | | | | | |
| *Incongruent*<br>(deviants or deviating stimuli) – 2-speaker setting | Spatial (slight) - standards | | | | | x | | | x |
| | Spatial (extreme) - standards | | | | | x | | | x |
| | Frequency (slight) - standards | | | | | | x | | |
| | Frequency (extreme) - standards | x | | | x | x | x | | |
| | White noise (spectrum) | | | | x | | x | x | |
| *Incongruent*<br>(deviants or deviating stimuli) – 4-speaker setting | Spatial & frequency | | x | x | | | | | |

Importantly, and in accordance with the literature, the attentional focus is highly precise regarding all to-be-attended features. Any changed, missing or additional features, even on the same side as the to-be-attended stimulus, lead to exclusion from focus causing reduced processing. This was demonstrated by results with fundamental frequency, spatial location and spectral content in the present experiments. Thus, just one matching feature would not be sufficient to be included in the focus.

The same attentional effects (Nd) were found in auditory environments of different complexity. Simple (two-speaker) as well as more complex (four-speaker) situations revealed similar ERP differences between attended and unattended stimuli, suggesting that the same attentional mechanism underlies stimulus selection and processing in complex auditory environments. Nevertheless, the four-speaker compared to the two-speaker experiment revealed an apparent mirror effect for ERP results between the attended and the unattended hemispace. The only difference between hemifields was a slight shift in processing level, in parallel below those for each counterpart on the attended side. This result, never reported before, indicates that even unattended stimuli are distinguished and not treated as the same stimulus. Even unattended stimuli are not completely inhibited, but still processed in a just slightly attenuated form that still allows them to be distinguished.

## *Validity of the applied HRTF method combined with the probe technique*

The HRTF method combined with the probe technique has not been used before to study auditory selective attention. Woods et al. (1984) employed probe stimuli superimposed on speech but without application of HRTFs. They found reliable Nd effects in their selective dichotic listening task. Thus, the first goal of this present study was to show that this combination of probes with the HRTF technique is a valid as well as useful addition and extension to previous rather simple dichotic designs. All four experiments, especially experiment two and three, have demonstrated clear attentional effects between stimuli from the attended and unattended side, or between standard probes and deviants. Attending to one of the two or four simultaneous stories in the current investigation increased the tonic neural activity in sensory areas responsible for processing of incoming auditory stimuli. The reason for this tonic increase might be ‘bias signals’ from higher attentional control areas (see review of Hillyard et al. (1998)). However, the superimposed probe stimuli are of phasic nature. By averaging brain waves of these probe stimuli, the tonic activity by the stories is canceled out and only the specific, phasic, modulation by the probes was observed. As the results show, systematic attention related ERP modulations occurred between attended and unattended probes.

However, typical attentional effects, such as an Nd/PN between conditions (Araki et al., 2005; Münte et al., 2001; Näätänen & Picton, 1987; Woldorff & Hillyard, 1991), have not been observed as clear in all present experiments. An additional Pd component (see discussion below) preceded and partially suppressed the typical Nd effect in two of four experiments (the first of the two-speaker (chapter 2.1) and the four-speaker experiment (chapter 3.1)). However, this more negative-going ERP (Nd/PN) for to-be-attended stimuli, especially in the other two experiments, is a clear replication of classic ERP results reported in selective listening studies (Hansen & Hillyard, 1980; Hillyard et al., 1973). Furthermore, behavioral data has shown that the applied method was quite sensitive to reveal individual differences in selection skills. The answer accuracy differed between subjects as well as between experiments. The individual difference was much more pronounced in the more complex environment (four-speaker) confirming an increased degree of difficulty from the two-speaker to the four-speaker experiment. These behavioral findings, as well, argue for the method’s validity in two quite realistic but different complex auditory environments.

Previous investigations of auditory selective attention used predominantly simple dichotic listening tasks (e.g., Hillyard et al., 1973 as one of the earliest selection studies). Only a few studies tried to apply a more realistic auditory situation by employing ILD, and/or ITD (Darwin & Hukin, 1999; Shinn-Cunningham & Ihlefeld, 2004), but no research was done with HRTFs so far.

In addition to the mentioned attention effects, another primary result of the current investigation was the observation of deviant effects. Spatial and frequency deviants caused the MMN and P3a components typically found in oddball paradigms (Escera et al., 2000; Friedman et al., 2001; Näätänen, 1992; Picton et al., 2000, for reviews). A RON occurred as well indicating distraction by deviant stimuli (Rüsseler et al., 2002).

Thus, by this current replication of typical attention as well as deviant effects, two important things have been demonstrated. First, these experiments have shown that the method of combining HRTFs and the probe technique is a valid procedure to investigate selective auditory attention in a more realistic auditory environment. Based on this result, the second main conclusion that can be drawn from this research is the sufficiency of a standard auditory design for this kind of inquiry. The current effects confirm the validity of previous results. Therefore, simple dichotic studies could investigate selective auditory attention in a valid way as well, without the effort of simulating a rather realistic auditory environment.

In the following paragraphs let us take a closer look at some specific attentional effects of the present study, and then also at the observed deviant effects just mentioned.

## Active inhibition in novel and unfamiliar situations

Usually, there is a clear negative difference (Nd or PN) between attended and unattended stimuli regarding their ERPs (Hillyard et al., 1973; Näätänen et al., 1978). However, this was not the case in all of our four experiments. The first and last experiment in the present investigation showed a main distinction in ERPs compared to previous research. There was an additional positive difference (Pd) that preceded the typically observed Nd between attended and unattended stimuli. No other study has detected this effect before. Nevertheless, this Pd-Nd co-occurrence only appeared in two out of four experiments (2.1 and 3.1).

How can this additional Pd effect be explained? Where does it come from, and why did it occur in this investigation and not in others? The only reasonable cause is the novelty and unfamiliarity of a situation or task. The participants took part in all four experiments and should therefore have become more experienced from one to the next round. Experiment 2.1 was the subject’s first test during this whole multi-experimental study. Experiment 3.1, on the other hand, was the last but one test out of the four. Thus, the participants should have been trained very well by the preceding trials. It is known that practice has an enormous effect on neural activity in specific brain areas (Petersen, van Mier, Fiez, & Raichle, 1998; Raichle et al., 1994). This training effect may be true for the other two 2-speaker-settings. Experiment 3.1, however, contained four talkers instead of just two as in the other three experiments. Therefore, the transfer of skills to the fourth experiment may not have been successful (Lockhart, 2002). There is evidence that skills learned under simple conditions are not always easy to transfer to complex skills (Wulf & Shea, 2002). The fourth experiment may have been quite unfamiliar to the subjects because of two additional concurrent talkers. Thus, this more complex cocktail-party situation may have functioned as a new and untrained setting despite participants’ experience with the two-speaker setting beforehand.

This novelty or unfamiliarity could then have elicited the Pd in these two experiments, because with more training (experiment 2.2 and 2.3 compared to experiment 2.1) no Pd was evoked. Instead a pure Nd effect occured, in these latter similar selective listening tasks. Thus, a speculation would be that this Pd might indicate a first active suppression of unattended stimuli when no procedure for automated filtering or initial discrimination between unattended and attended stimuli is available due to novelty or unfamiliarity. In simple or even trained complex tasks, the identification of unattended and irrelevant stimuli could be easier or supported by a parallel automatism that inhibits unwanted information from early stages on (e.g., Näätänen, 1990). Differences in Pd amplitude between the first and second half of experiment 2.1 (first two-speaker setting) and 3.1 (four-speaker environment) support this assumption. The Pd was larger at the beginning, and reduced towards the end of the experiment. This suggests that a "novelty" effect is represented by the positive difference.

Perhaps the Pd and Nd, whose onsets indicates the time at which a particular feature or object is discriminated and selectively processed according to its task relevance (Näätänen, 1992) may differ in just one aspect. The Nd may reflect the specific attentional selection of an auditory object whereas the Pd may represent a general selection process after an initial orientation and possibly distraction period (Alho, Tottola et al., 1987; Iwanami, Kanamori, Isono, Okajima, & Kamijima, 1996; Kamio et al., 2001). Furthermore, it may be that the more concurrent and distracting streams are present, the longer the Pd is. Then, the Pd may increasingly overlay the Nd. This assumption arises from the observation that the Nd in experiment 3.1 (four talkers) was quite small compared to the Pd and in contrast to the Nd in experiment 2.1. Because the Pd is not yet well understood, further research will be needed to disambiguate processes reflected by this component. Additionally, to avoid such unfamiliarity effects in future experiments, practice trials shortly before the actual experiment may prevent effects seen in our ERPs resulting from unfamiliarity in the more complex task.

## Precision of the attentional focus

After seeing that attended and unattended stimuli are processed differently, it was another interesting question to observe which stimulus falls into attentional focus, and how strongly restricted the focus would be. To address this question, we analyzed in detail which stimulus features are included in the attentional focus and to which degree variations in feature values lead to exclusion from the focus of attention.

The primary result was a "modification hypothesis", because any feature or object changes – especially changes in a stimulus' location – led to an exclusion from the focus of attention even when presented in the same hemifield on which attention was focused. Probe stimuli that differed either in spectral content, fundamental frequency, or spatial location compared to the to-be-attended stimulus – whether they were just slightly or extremely deviant – evoked modulated ERP components (MMN, P3a, RON, N300). These effects demonstrate change detection and a following attentional switch respectively (Friedman et al., 2001; Münte et al., 1995). As deviants on the attended side were compared with their counterparts in the unattended hemispace, no attentional modulation (Nd) became apparent.

Thus, the focus of attention is tightly restricted to **all** to-be-attended stimulus features (e.g., fundamental frequency, spectrum, **and** location) determined by the current situation's demand (e.g. an individual intention, a given task, or stimulus driven if no endogenous goal is set up). Any small and especially large deviance was treated as a different input than the to-be-attended stimulus and thus processed as such – as an unattended stimulus. These results extend Paavilainen's study (1993) who found a strong attentional focus for frequency deviants as well. The current study demonstrated furthermore that also the spectral content of a speaker's voice plays an important role in whether a stimulus is included in attentional focus. It is not sufficient to have the same fundamental frequency; it is necessary to have exactly the same and not more frequencies as the to-be-attended stimulus stream. If the frequency spectrum exceeds a to-be-attended frequency range, the stimulus will be recognized as different and processed as being irrelevant outside of the attentional focus. White noise, that contained the talker's spectrum for example, was filtered out.

However, besides frequency and spectral deviations, spatial location played an important role in auditory selection. In the present study, the minimum difference between spatially different stimuli was 15° (slight spatial deviants). This distance was pretty large regarding a very restricted spatial focus. Teder and Näätänen (1994) were the first authors who reported a narrow focus in spatial attention with neurophysiologic evidence. They found a steep decline in the N1 effect from an attended location to adjacent locations that were 3° in space apart. Thus, the spatial resolution in selective auditory attention equals or might be smaller than 3°. This would explain why there was no significant ERP difference between slight and extreme spatial deviants. Both kind of stimuli were "far" away from focus and therefore attenuated similarly. With fifteen or more degrees difference, and because of the apparent floor function (Teder & Näätänen, 1994; W. Teder-Sälejärvi & Hillyard, 1998; W. A. Teder-Sälejärvi et al., 1999), there was no significant distinction between these locations further away from focus.

Furthermore, Teder-Sälejärvi et al. (1998) found evidence for two distinct stages of spatial focusing of attention. The early selection in form of an N1 modulation was a more broadly tuned filtering of all inputs (between different auditory channels) around 80-200 msec after stimulus onset. Thereafter, the selection was more narrowly focused on spatial deviants

(> 250 msec; within an auditory channel). Both effects were observed for spatial deviants in the present investigation as well. Our data revealed these indicators additionally for frequency and spectral deviants as well as for standard probes. Thus, the N1 modulation or MMN respectively as well as the PNe reflect the initial input filter in general whereas the P3a, RON as well as the PNl image the subsequent selection for all types of deviating stimuli. The latter stage would explain why attentional focus is strongly restricted. After rough filtering of information (early stage), a later finely tuned selection may easily detect any small deviations in the reduced input. In the case of standard stimuli, this two-stage-filtering became most obvious in the two experiments showing the Pd-Nd complex. As discussed in the previous paragraph about active inhibition in novel or unfamiliar situations, the Pd may represent an initial general selection process whereas the Nd may reflect the specific attentional selection of an auditory object.

Altogether, the modification hypothesis specifies that **any** changes in a stimulus' physical characteristics – whether by adding or changing – lead to exclusion from the strong focus of attention. Thus, the deviants will be perceived separately and not as belonging to the attended object (this is an example of object-based perception). Hence, they will be processed in a more attenuated way as well. The first indications for this assumption were seen in the late seventies. Hillyard et al. (1975; 1976) for example investigated N1 enhancement by different pitch frequencies and/or locations. They argued that this increased N1 reflects finely tuned selective attention to one stimulus among several concurrent and competing stimuli/channels. Thus, their results could be taken as support for the here mentioned modification hypothesis. Any changes in either location or pitch frequency in Hillyard's studies led to the discovery of a distinct auditory channel perception that is excluded from the actual attentional focus.

## Two speakers vs. four speakers

In the present study two different auditory environments were employed to investigate whether differences in stimulus processing would depend on the complexity of a situation. One would expect that unattended stimuli in a four-speaker setting would be attenuated more than in a two-speaker environment. Furthermore, from classical simple dichotic listening investigations of selective attention, especially those focusing on Nd/PN research, one might conclude that all stimuli on the unattended side are markedly attenuated compared to their attended counterparts (Michie, Bearpark, Crawford, & Glue, 1990; Näätänen, 1982). They are in fact processed to a lesser degree, but not in the manner that all ERP amplitudes from the unattended side are all below the lowest amplitudes from the attended ear as the present data revealed. Thus, there is no rank order in the sense that all stimuli from the attended channel are processed predominantly, and then all others from the unattended side are processed.

### Two-speaker

All three two-speaker experiments in the current investigation have shown that the stimuli from the ignored hemifield are indeed attenuated compared to the attended side, but unattended congruent probes elicited still a more negative ERP than attended incongruent stimuli. The same ERP components between congruent and incongruent probes were visible on both sides whereas their appearance was not or was just slightly larger for the attended hemifield. Thus, the assumption of a simple and direct rank order (first attended, then unattended) is not supported, because stimuli from the irrelevant hemispace are not processed as a single class of deviant. They are somehow still distinguished.

No research has addressed this matter to date. Only direct comparisons between an attended stimulus or component and the same component when unattended, or between two different stimuli on the same side (attended or unattended) have been studied so far, as is the case in the present investigation. No previous investigation has attempted yet to analyze differences between attended congruent and unattended incongruent stimuli, or vice versa. Nevertheless, the current two-speaker data suggests that the two congruent auditory channels – attended vs. unattended hemifield – with all their input information might simply be mirroring each other. Thus, by attending to one ear, information from the other side might be processed in a similar and partially, but not completely, attenuated way. One could picture these attentional effects for the unattended side as a combined parallel shift to slightly below the level of processing of their attended counterparts, but not below the minimum level of all attended stimuli in ERP amplitudes.

### Four-speaker

Further support for this mirror characteristic in stimulus processing in the two auditory hemifields was provided by the four-speaker experiment (3.1). With four simultaneous speech streams, it was clearly shown that stimuli from the unattended side are not suppressed entirely and not processed to the same degree, below all attended stimuli. There is rather a gradual suppression of equivalent, mirroring locations. Thus, the two unattended streams were not treated in the same way, but in a way that depends on their location. If the location of the unattended stream mirrored the location of the to-be-attended stream in the contralateral hemifield, then more mental resources were given to this unattended stream than to one whose location did not mirror the location of the attended stream. That is why a PN was observable, even between the two locations (target vs. non-target) in the unattended hemispace and not only in the attended hemifield. This similar deviance or attentional pattern, with slightly smaller components for the unattended hemifield only, suggested that both unattended spatial locations mirroring the spatial position on the attended side are processed in a similar manner with attentional modulation (Näätänen, 1982, 1988, 1990, 1992; Näätänen & Alho, 2004). The differences in stimulus processing are seen within each hemifield between different spatial angles. That was the reason to assume that spatial hemispheres (left vs. right) may simply mirror each other, causing a similar processing of stimuli appearing at contralateral positions. This observation can further be seen as an evidence for a division of our auditory environment into two distinct hemispaces, which further supports the conclusion that within-hemifield spatial discriminations are completely different from laterality (across-hemifield) discriminations (Boehnke & Phillips, 1999; Heffner & Heffner, 1990).

One could further speculate that attention itself is able to attenuate the sensitivity of one auditory hemifield independently from the other side. This differentiation between two auditory hemi-spaces and attentional shift in one hemispace could explain the observed hierarchy in allocation of mental resources. The most preferentially processed stimulus was the one presented at exactly the same location as the attentional focus, followed by the stimulus that occurs at the position in the opposite hemifield because there is less contrast between it and spatially varying stimulus within the same hemisphere. A clear attention effect was shown for stimuli at the non-target locations in hemifields. Thus, the next stimulus within the processing hierarchy is a non-target condition on the attended side. The stimulus with the fewest allocated resources is a non-target probe that differs not only in the spatial hemisphere but also in the spatial position with regard to the attentional focus. Thus, this latter stimulus contrasts the most relative to the target probe and is therefore processed less than any other stimulus in this hierarchy.

However, another possible reason for the observed main differences within each hemifield could lie in the speaker's gender. Female voices were used for target streams whereas the non-targets were spoken by male speakers. It has been shown that gender difference can cause an increase in discriminability (D. S. Brungart, Simpson, Ericson, & Scott, 2001). Brungart et al. also reported that the ability to extract information from the target phrase in a 3- or 4-talker setting decreases when the voice characteristics of the masking or unattended stimuli are very similar. This was the primary reason why we chose to use different genders for target- and non-target talkers. This should make it easier for the subjects to segregate the four simultaneous speech streams. Brungart's behavioral study (2001) showed that monaural factors might be important in the segregation of speech signals in multichotic environments. Thus, it could have been that the effects we observed were caused by gender difference between target and non-target stimuli rather than by the different locations itself. Further investigations with the same four-speaker setting, but also the same talker's gender – either all male or all female – are needed to make more explicit conclusions.

No research has been done so far that has investigated a multi-speaker setting with linguistic probe stimuli and inter-hemispatial attentional effects. There are a few ERP studies analyzing three or more noise (pink noise) sources (Münte et al., 2001; Nager, Kohlmetz et al., 2003; Röder et al., 1999; W. A. Teder-Sälejärvi et al., 1999), but competing sounds were either presented in one hemifield only to investigate pure segregation effects, or the right/left processing was compared with the azimuth (center position) to look for the position of sharpest focus, and thus ignored right-left attentional differences (see Excursus). Thus, there is no data available that can help to interpret the present observations directly.

**Excursus:**

Spatial location, more than other features of an auditory object, seems to play a very important role in selective auditory attention. This is why we will take a closer look at spatial attention at this point.

As seen in the different experiments, spatial deviants showed a reduced stimulus processing compared to the attended stream. None of the deviant effects, such as MMN or P3a, were significantly larger in either of the two auditory hemifields (attended or unattended) when two speech streams were presented. While selecting one out of four competing speech messages, not only differences between target locations (L60) but also between stimuli from the non-target locations (L15) on either side were observed. Thus, information from deviant and not mirroring positions relative to the to-be-attended speech stream is processed differently and not as one class of deviating stimuli. L15 stimuli on the attended side were slightly less attenuated than the same stimuli when they were in the unattended hemifield.

This observation indicates that attenuation of processing for stimuli outside the attentional "spotlight" does not follow an all-or-nothing principle. Rather, attention effects decrease as a function of distance to the attended location (Teder & Näätänen, 1994; W. Teder-Sälejärvi & Hillyard, 1998; W. A. Teder-Sälejärvi, Hillyard, Röder, & Neville, 1999). Teder-Sälejärvi et al. (1994; 1998; 1999) conducted several experiments revealing a gradient of attentional focus rather than discrete boundaries. Overall, they found a finely tuned gradient around the attended sound source. The attended location is primarily processed. Stimuli at immediately adjacent locations, however, are notably attenuated not to mention stimuli at far distant location. Furthermore, Teder-Sälejärvi's data (1998; 1999) suggested that the gradient is steeper around attended central stimuli than around attended peripheral stimuli. Thus, the spatial localization is most precise at azimuthal positions. Nevertheless, this azimuthal advantage is not absolute. In a study by Röder et al. (1999) it was shown that blind participants have superior localization abilities than those of sighted control subjects when attending to sounds in peripherical auditory space.

For peripheral sound sources, seeing subjects evoked an Nd effect for sound from the loudspeaker next to the attended location as well, indicating a worse spatial selectivity in peripheral space. Blind people, in contrast, showed a more pronounced Nd effect for the attended loudspeaker than for adjacent locations in the auditory periphery, similar to azimuthal positions. One conclusion, therefore, would be that the spatial resolution for peripheral sound sources is improved in blind people. Not only blind subjects seem to have an increased peripheral spatial resolution but also people who professionally need to be able to distinguish between auditory information at peripheral locations. Professional music conductors for example have to monitor the performance of the entire orchestra as well as the performance of each single musician. After years of experience, conductors developed an increased spatial resolution for peripheral auditory sound sources as well. Nager et al. (2003) reported that only conductors showed attentional selectivity in form of an Nd effect for peripheral speakers whereas the Nd was quite similar for all three peripheral speakers in pianist and non-musicians.

The last two experiments indicated profound changes in attention related processing of sounds from different spatial locations in blind and professional conductors compared to seeing subjects without musical experience. Is it due to neuroplasticity? This question cannot be answered surely at the moment. However, there are indications that compensatory reorganization of brain areas in the blind or experience related functional and anatomical changes in conductors may have contributed to the improved spatial resolution for the auditory periphery. Further investigation addressing the question about neuroplasticity in this domain is needed.

However, applied to the present results of spatial deviants, these studies mentioned above cannot explain why, in the four-speaker setting, the non-target stimuli in the unattended hemispace (A-S-L15) are slightly less processed than the irrelevant stimuli in the attended hemifield (A-S+L15). Though, there is evidence for a gradient characteristic of auditory attentional focus as for example observed by Teder-Sälejärvi et al. (1998; 1999). Considering that all stimuli are distributed along a semicircle in front of each subject, A-S-L15 stimuli are indeed further away from the actual focus than A-S+L15 stimuli. This could have led to the more attenuated processing for these stimuli further apart reported in experiment 3.1. Nevertheless, one could also see in Teder-Sälejärvi et al.'s studies (1998; 1999) that the gradient effect has a floor function for stimuli with increasing distance to the to-be-attended location. Only very adjacent locations benefit in their processing from mental sources provided for the attentional focus. According to this, one would not expect any significant differences between deviant stimuli that are far and further away. With a distance of 45 spatial degrees between a target and non-target in one auditory hemispace, there is most likely no benefit left for the processing of non-target stimuli in the same or the opposite hemispace. Rather, a similar stimulus processing would be expected which was not the case. Actually, attention effects should have been minimized with this distance between speech streams. Nevertheless, experiment 3.1 showed attention effects between target and non-target streams. What could explain this phenomenon? Perhaps, it may have been caused by the better spatial resolution at azimuthal locations (W. Teder-Sälejärvi & Hillyard, 1998; W. A. Teder-Sälejärvi et al., 1999). In the present study, the to-be-attended message was presented rather in the auditory periphery whereas the non-target but competing streams occurred more towards the central locations at 15° to either side. Thus, even though the non-targets were absolutely irrelevant, they could have been analyzed more distinctively; simply because of the higher spatial resolution around azimuth compared to peripheral locations. This fact may explain why there was still, or again, an attention effect for stimuli 45° apart from focus. Previous research looked at two to three adjacent locations that covered either the periphery or the central region only (W. Teder-Sälejärvi & Hillyard, 1998; W. A. Teder-Sälejärvi et al., 1999). There is no report of the whole range from periphery to azimuth.

Furthermore, the observed stronger attenuated stimulus processing for non-targets in the unattended hemispace may be due to a general attentional disadvantage for this hemifield relative to the attended one. This would be true for the comparison between the two target streams – attended vs. unattended. The auditory environment might be mentally divided into two hemispaces – right and left hemifields. Thus, stimuli from each hemispace are processed in parallel but still distinctively depending on attentional modulation. Perhaps there is a relationship between a location in one hemifield and a location situated symmetrically across the midline, or “mirrored”, in the other hemispace. Not all stimuli from the attended ear are processed predominantly over all others from the other hemifield. Instead, unattended but competing stimuli from a mirror position to a to-be-attended stimulus/location seem to get more mental resources than a stimulus from a different location stimulus on the same attended side. It is as though attention slightly but incompletely decreased the sensitivity to, but not the discriminability between, stimuli on the unattended side. This means that, even in a very demanding and complex auditory environment, all incoming stimuli are processed at least at a low level.

The aforementioned studies, however, differed from the present investigation in one important thing. They used noise bursts (pink noise) only as stimuli without speech streams as in the present study. Furthermore, they conducted an active oddball task with deviants that varied in bandwidth from standard tones. Attended deviant stimuli were clearly distinct to unattended deviants by spatial location. The same was true for standard noise bursts. Thus, the spatial attention was directly accessible. In contrast, we used linguistic probes (phonemes), with exception of white noise in experiment 2.3, superimposed on speech streams. All experiments were passive in nature, because the subjects’ task was to attend one of two or four simultaneous prose streams. The superimposed probes were task-irrelevant. Our attended standard probe stimuli shared location, fundamental frequency as well as spectral content of these speech streams respectively. Therefore, they were embedded into and processed as the speech message. Deviating stimuli in experiment 3.1, where we found an apparent attentional modulation between the two auditory hemifields, differed both in location and fundamental frequency, and were furthermore embedded in another speech stream. Most importantly, the spatial distance between stimuli in the present experiment were 30°-45° whereas it was very small (3°-9°) in previous studies. Altogether, these differences could have led to different processing of stimuli in the present investigation that revealed a possible relationship between the two auditory hemifields.

Nonetheless, if the mirror characteristic between auditory hemispaces is true then it will not contradict classical auditory selection theories and their effects (Näätänen, 1982; Näätänen et al., 1978; D. L. Woods & Clayworth, 1987). Those previous studies explored only two simultaneous auditory streams presented either directly from either side (180° spatial separation) or from a certain angle in space equally for the right or left ear. This means that they always compared two mirroring locations in space when investigating auditory hemifields. From that point of view, it is not wrong to say that attending selectively to one auditory stream leads to attenuation in stimulus processing for unattended ear stimuli. It would be a false assumption to think that all stimuli from that unattended side – whether from a mirroring location compared to the to-be-attended position or not – are processed in the same way or suppressed with the same strength. The present results indicated that selective attention might be not that simple. Even within the unattended hemifield, there is differentiation between stimuli – for example, for different positions, deviance in frequency or spectral content. It is definitely not the case that all stimuli from the unattended side are lumped together and therefore processed equally as though they belong to one and the same channel and without any distinction in stimulus features.

However, the same ERP components and attentional effects were found in the more complex situation as in a simpler set-up, indicating that the same selection mechanism is used in any auditory environment – no matter how complex a given situation is. The only obvious distinction is that the subject's ability to listen selectively to one out of several concurrent streams and keep their attention there for a longer period differed greatly between all participants. Individual differences between subjects were significant; not all subjects could fulfill the task adequately. The drop-out rate was chiefly influenced by inability to cope with the excessive demands in the more complex cocktail-party situation whereas in a simpler setting no such discrimination was possible. Brungart and colleagues (2002; 2004) reported that ability to segregate informational channels decreases with competing speech streams in the same ear, especially when there was also an additional talker on the other side. Therefore, the current results do not argue against the experimental design, but rather for the complexity of the presented auditory situation.

### *Summary: two-speaker vs. four-speaker*

Taken together, the observed attentional differences between attended and unattended probe stimuli (stronger processing of attended than unattended phonemes) in the more complex environment are consistent with to the literature and the results of the preceding experiments with a simpler cocktail-party situation. Thus, the same underlying selection mechanism can be assumed. Moreover, there is a hierarchy in stimulus processing for spatially different speech streams arranged in a semi-circle. This hierarchy does not simply follow this semi-circle arrangement in front of a listener or as a function of deviance from the to-be-attended story. Rather, it is more like a zigzag gradation with similar processing for stimuli at mirroring locations in both hemispaces whereas attention alone is able to increase reactivity to a stimulus within the attended hemifield and thereby causes the zigzag effect. Thus, there are good reasons to assume two concurrent auditory hemispaces that mirror each other. The brain seems to be able to process stimuli in both hemifields separately by attentional demands, but still in a similar way. Nevertheless, from the present results, it cannot be excluded that the observed effects may be due to different talker's gender. Therefore, the observed mirror characteristic might only be valid for this specific four-speaker setting in the current investigation. Future studies using the same gender for all competing talker may unveil whether the mirror theory is the true cause of these effects.

## *The experimenter's dilemma*

For a long time, researchers have been interested in one of the most important but incomprehensible human cognitions: the phenomena of attention and attentional selection. What is attention? Is it sensory specific?

In order to learn more about this intriguing topic, the researcher must create an overloading situation, which means a setting with more stimuli than a human can process at the same time. Thus, selection of certain stimuli is necessary so that they can be processed properly without risking an overload of working memory. To investigate basic mechanisms underlying attention, situations are needed that are not too complex, but do they still represent the attentional processes in the daily life? There are different and sometimes amazing results. What do the results of the simple attentional task really tell us? Attention has been studied on several levels over the time: cognitive, neurosystem, cellular, synaptic, and genetic level. None of these provide a sufficient analysis of the role of attention which prepares the way for a reevaluation of experimental conditions. The term attention is still not completely understood. Only some links between those different levels are made that allow attention to be viewed as an independent system with its own anatomy, circuitry, functions, and deficits (specific attentional disorders). Therefore, attentional research still does not have a better definition of attention than William James had a century ago.

Without a real definition and understanding of attention, it is quite difficult to conceive an adequate experimental design. Till now, these paradigms have been developed in a more stepwise approach. Nevertheless, the safest way to study auditory attentional phenomena is in a cocktail-party situation with overlapping discussions and a noisy background. Only by selecting a specific conversation, voice, or person a listener can understand what is said. By far, only one and not all simultaneous discussions in such a cocktail party can be followed at a time. This does not exclude a switch of attention to another conversation. But what switches – the attention itself or the perceptional system directed by attention? This and even more questions are still unanswered. Therefore, there is still much to investigate regarding attention and its mechanisms. All that can be said now is that attention is highly focused on a specified auditory information stream characterized by certain features. Any changes of these features lead to an exclusion from the actual attentional focus.

# References

Alain, C., Achim, A., & Woods, D. L. (1999). Separate memory-related processing for auditory frequency and patterns. *Psychophysiology, 36*, 737–744.

Alain, C., & Arnott, S. R. (2000). Selectively Attending to Auditory Objects. *Frontiers in Bioscience, 5*, 202-212.

Alain, C., & Izenberg, A. (2003). Effects of Attentional Load on Auditory Scene Analysis. *Journal of Cognitive Neuroscience, 15*(7), 1063–1073.

Alain, C., & Woods, D. L. (1994). Signal clustering modulates auditory cortical activity in humans. *Perception and Psychophysics, 56*, 501-516.

Alain, C., & Woods, D. L. (1997). Attention modulates auditory pattern memory as indexed by event-related brain potentials. *Psychophysiology, 34*, 534-546.

Alho, K., Donauer, N., Paavilainen, P., Reinikainen, K., Sams, M., & Näätänen, R. (1987). Stimulus selection during auditory spatial attention as expressed by event-related potentials. *Biological Psychology, 24*, 153-162.

Alho, K., Medvedev, S. V., Pakhomov, S. V., Roudas, M. S., Tervaniemi, M., Reinikainen, K., et al. (1999). Selective tuning of the left and right auditory cortices during spatially directed attention. *Brain Res. - Cognitive Brain Research, 7*, 335-341.

Alho, K., Teder, W., Lavikainen, J., & Näätänen, R. (1994). Strongly focused attention and auditory event-related potentials. *Biological Psychology, 38*, 73-90.

Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R. J., et al. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology, 33*, 369–375.

Alho, K., Tottola, K., Reinikainen, K., Sams, M., & Näätänen, R. (1987). Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalography and Clinical Neurophysiology, 68*(6), 458-470.

Anllo-Vento, L., Schoenfeld, M. A., & Hillyard, S. A. (2004). Cortical Mechanisms of Visual Attention. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 180-193). New York, London: The Guilford Press.

Araki, T., Kasai, K., Nakagome, K., Fukuda, M., Itoh, K., Koshida, I., et al. (2005). Brain activity for active inhibition of auditory irrelevant information. *Neuroscience Letters, 374*, 11-16.

Baldeweg, T., Klugman, A., Gruzelier, J. H., & Hirsch, S. R. (2002). Impairment in frontal but not temporal components of mismatch negativity in schizophrenia. *Int. J. Psychophysiol., 43*, 111 –122.

Basil, M. D. (1994). Secondary reaction-time measures. In A. Lang (Ed.), *Measuring Psychological Responses to Media Messages* (pp. 85- 98). Hillsdale, NJ: Lawrence Erlbaum.

Beer, A. L., & Röder, B. (2004). Attention to motion enhances processing of both visual and auditory stimuli: an event-related potential study. *Cognitive Brain Research, 18*, 205–255.

Beer, A. L., & Röder, B. (2005). Attending to visual or auditory motion affects perception within and across modalities: an event-related potential study. *European Journal of Neuroscience, 21*, 1116-1130.

Begault, D. R. (2000). *3-D Sound for Virtual Reality and Multimedia*. Moffett Field, California: NASA/TM; Ames Research Center.

Begault, D. R., Wenzel, E. M., & Anderson, M. R. (2001). Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal Audio Eng. Soc., 49*(10), 904-916.

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Arch. Psychiatry and Clinical Neurosciences, 87*, 527–570.

Berti, S., Roeber, U., & Schröger, E. (2004). Bottom-up influences on working memory: behavioral and electrophysiological distraction varies with distractor strength. *Exp. Psychol., 51*(4), 249-257.

Blauert, J. (1997). *Spatial Hearing: "The Psychophysics of Human Sound Localization"* (Revised Edition ed.). Cambridge, MA: MIT Press.

Boehnke, S. E., & Phillips, D. P. (1999). Azimuthal tuning of human perceptual channels for sound location. *Journal of Acoustic Society of America, 106*(4), 1948-1955.

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research, 25*, 107-116.

Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology, 41*(2), 259-268.

Bregman, A. S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. *MIT Press, Cambridge, MA.*

Broadbent, D. E. (1958). *Perception and Communication*. Oxford: Pergammon Press.

Brungart, D., & Rabinowitz, W. (1996). *Auditory Localization in the Near-Field.* Paper presented at the Third International Conference on Auditory Display, Santa Fe Institute.

Brungart, D. S., & Simpson, B. D. (2002). Within-ear and across-ear interference in a cocktail-party listening task. *Journal of Acoustic Society of America, 112*(6), 2985-2995.

Brungart, D. S., & Simpson, B. D. (2004). Within-ear and across-ear interference in a dichotic cocktail party listening task: effects of masker uncertainty. *Journal of Acoustic Society of America, 115*(1), 301-310.

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of Acoustic Society of America, 110*(5), 2527-2538.

Busse, L., Roberts, K. C., Crist, R. E., Weissmann, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences, 102*(51), 18751-18756.

Carlille, S., & Pralong, D. (1994). The location-dependent nature of perceptually salient features of the human head-related transfer functions. *Journal of Acoustic Society of America, 95*(6), 3445-3459.

Carlyon, R. P., Cusack, R., Foxton , J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology Human Perception Performance, 27*, 115-127.

Cherry, C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of Acoustic Society of America, 25*, 975-979.

Cherry, C. (1957). *On Human Communication: A Review, a Survey and a Criticism*. Cambridge, MA: MIT Press.

Coch, D., Sanders, L. D., & Neville, H. J. (2005). An Event-related Potential Study of Selective Auditory Attention in Children and Adults. *Journal of Cognitive Neuroscience, 17*(4), 605-622.

Colin, C., Radeau, M., Soquet, A., Dachy, B., & Deltenre, P. (2002). Electrophysiology of spatial scene analysis: the mismatch negativity (MMN) is sensitive to the ventriloquism illusion. *Clinical Neurophysiology, 113*, 507-518.

Comerchero, M. D., & Polich, J. (1998). P3a, perceptual distinctiveness, and stimulus modality. *Brain Res. - Cognitive Brain Research, 7*, 41-48.

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology, 110*, 24-30.

Cowan, N. (1995). Attention and Memory: An Integrated Framework. *Oxford University Press.*

Crispien, K., Fellbaum, K., Savidis, A., & Stephanidis, C. (1996). *A 3D-Auditory Environment for Hierarchical Navigation in Non-visual Interaction.* Paper presented at the International Conference on Auditory Display.

Cycowicz, Y. M., & Friedman, D. (2004). The old switcheroo: when target environmental sounds elicit a novelty P3. *Clinical Neurophysiology, 115*, 1359-1367.

Darwin, C. J., & Hukin, R. W. (1999). Auditory objects of attention: the role of interaural time differences. *Journal of Experimental Psychology Human Perception Performance, 25*(3), 617-629.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cogn. Psychol., 3*, 255–267.

de Bruin, N. M., Ellenbroek, B. A., van Schaijk, W. J., Cools, A. R., Coenen, A. M., & van Luijtelaar, E. L. (2001). Sensory gating of auditory evoked potentials in rats: effects of repetitive stimulation and the interstimulus interval. *Biol. Psychol., 55*(3), 195-213.

Deutsch, J., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80-90.

Doeller, C. F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., & Schröger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *NeuoImage, 20*, 1279-1282.

Donald, M. W., & Little, R. (1981). The analysis of stimulus probability inside and outside the focus of attention, as reflected by the auditory N1 and P3 components. *Can. J. Psychol., 35*, 175-187.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 manifestation of context updating? *Behav. Brain Sci., 11*, 357-374.

Duda, R. O., & Martens, W. L. (1997). *Range Dependence of the HRTF for a Spherical Head.* Paper presented at the IEEE ASSP Workshop.

Eason, R. G. (1981). Visual evoked potential correlates of early neural filtering during selective attention. *Bull. Psychonom. Soc., 18*, 203-206.

Eimer, M., & Van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: Evidence from event-related potentials. *Psychophysiology, 39*, 437-449.

Escera, C., Alho, K., Schröger, E., & Winkler, I. (2000). Involuntary Attention and Distractibility as Evaluated with Event-Related Brain Potentials. *Audiology Neuro-Otology, 5*, 151-166.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *J. Cogn. Neurosci., 10*, 590–604.

Fabiani, M., karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology, 23*, 298-308.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: influence of context, semantic memory organization, and perceptual predictability on picture processing. *J. Exp. Psychol. Learn. Mem. Cogn., 27*, 202–224.

Fisher, H., & Freedman, S. J. (1968). The role of the pinnae in auditory localization. *Journal of Auditory Research, 8*, 15-26.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. Rev., 25*, 355-373.

Frodl-Bauch, T., Kathmann, N., H.J., M., & Hegerl, U. (1997). Dipole localization and test-retest reliability of frequency and duration mismatch negativity generator processes. *Brain Topography, 10*, 3–8.

García-Larrea, L., Lukaszewicz, A., & Mauguiére, F. (1992). Revisiting the oddball paradigm. Non-target vs. neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia, 30,* 723-741.

Gevins, A. S., Brickett, P., Costales, B., Le, J., & Reutter, B. (1990). Beyond topographic mapping: towards functional–anatomical imaging with 124-channel EEGs and 3-D MRIs (Review). *Brain Topography, 3,* 53–64.

Giard, M., Perrin, F., & Pernier, J. (1990). Brain generators implicated in processing of auditory stimulus deviance: A topographical ERP study. *Psychophysiology, 27,* 627–640.

Giard, M., Perrin, F., Pernier, J., & Peronnet, F. (1988). Several attention related waveforms in auditory areas: A topographic study. *Electroencephalography and Clinical Neurophysiology, 69,* 371–384.

Giard, M.-H., Fort, A., Mouchetant-Rostaing, Y., & Pernier, J. (2000). Neurophysiological Mechanisms of Auditory selective Attention in Humans. *Frontiers in Bioscience, 5,* 84-94.

Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representations of stimulus frequency, intensity and duration in auditory stimulus deviants. *J. Cogn. Neurosci., 7,* 133–143.

Goldberg, M. E., & Wurtz, R. H. (1972). Activity of superior colliculus in behaving monkey. II. The effect of attention on neuronal responses. *J. Neurophysiol., 35,* 560-574.

Goldstein, A., Spencer, K. M., & Donchin, E. (2002). The influence of stimulus deviance and novelty on the P300 and Novelty P3. *Psychophysiology, 39,* 781-790.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: semantic ERP effects. *Neuropsychologia, 38,* 1518-1530.

Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and Clinical Neurophysiology, 106,* 156-164.

Handy, T. C., & Mangun, G. R. (2000). Attention and spatial selection: Electrophysiological evidence for modulation by perceptual load. *Perception and Psychophysics, 62,* 175-186.

Hansen, J. C., Dickstein, P. W., Berka, C., & Hillyard, S. A. (1983). Event-related potentials during selective attention to speech sounds. *Biological Psychology, 16*(3-4), 211–224.

Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology, 49,* 277-290.

Hansen, J. C., & Hillyard, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology Human Perception Performance, 9,* 1-19.

Hansen, J. C., & Hillyard, S. A. (1984). Effects of stimulation rate and attribute cuing on event-related potentials during selective auditory attention. *Psychophysiology, 21,* 394-405.

Hansen, J. C., & Hillyard, S. A. (1988). Temporal dynamics of human auditory selective attention. *Psychophysiology, 25,* 316-329.

Heffner, H. E., & Heffner, R. S. (1990). Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *Journal of Neurophysiology, 64,* 915-931.

Heinze, H. J., Luck, S. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. I. Evidence for early selection. *Electroencephalography and Clinical Neurophysiology, 75,* 511-527.

Hershkowitz, R., & Durlach, N. (1969). Interaural Time and Amplitude JNDs for a 500-Hz Tone. *Journal of the Acoustical Society of America, 46,* 1464-1467.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science, 182,* 177-180.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. London, 353,* 1257-1270.

Hink, R. F., Frenton, W. H. J., Pfefferbaum, A., Tinklenberg, J. R., & Kopell, B. S. (1978). The distribution of attention across auditory input channels: an assessment using the human evoked potential. *Psychophysiology, 15*, 466-473.

Hink, R. F., & Hillyard, S. A. (1976). Auditory evoked potentials during selective listening to dichotic speech messages. *Perception and Psychophysics, 29*, 36-242.

Hruby, T., & Marsalek, P. (2003). Event-Related Potentials - the P3 Wave. *Acta Neurobiologiae Experimentalis, 63*, 55-63.

Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *Journal of the American Statistical Association, 65*, 1582-1589.

Iwanami, A., Kanamori, R., Isono, H., Okajima, Y., & Kamijima, K. (1996). Impairment of inhibition of unattended stimuli in schizophrenic patients: event-related potential correlates during selective attention. *Neuropsychobiology, 34*, 57–62.

Iwanami, A., Shinba, T., Sumi, M., Ozawa, N., & Yamamoto, K. (1994). Event-related potentials during an auditory discrimination task in rats. *Neuroscience Research, 21*(1), 103-106.

James, W. (1890). *The Principles of Psychology*. New York: Holt.

Jäncke, L., Mirzazade, S., & Shah, N. J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neuroscience Letters, 266*, 125-128.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 370-375.

Jemel, B., Oades, R. D., Oknina, L., Achenbach, C., & Röpcke, B. (2003). Frontal and Temporal Lobe Sources for a Marker of Controlled Auditory Attention: The Negative Difference (Nd) Event-Related Potential. *Brain Topography, 15*(4), 249-262.

Johnson, R. (1989). Developmental evidence for modality-dependent P300 generators: a normative study. *Psychophysiology, 26*, 651-667.

Kamio, S., Nakagome, K., Murakami, T., Kasai, K., Iwanami, A., Hiramatsu, K., et al. (2001). Impaired suppression of processing in schizophrenic patients suggested by ERPs obtained in a selective attention task. *Schizophr. Res., 49*, 213–221.

Kastner, S. (2004). Attentional Response Modulation in the Human Visual System. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 144-156). New York, London: The Guilford Press.

Katayama, J., & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology, 35*, 23-33.

Katayama, J., & Polich, J. (1999). Auditory and visual P300 topography from a 3 stimulus paradigm. *Clinical Neurophysiology, 110,* 463-468.

Kayser, J., Bruder, G. E., Tenke, C. E., Stewart, J. E., & Quitkin, F. M. (2000). Event-related potentials (ERPs) to hemifield presentations of emotional stimuli: differences between depressed patients and healthy adults in P3 amplitude and asymmetry. *Int. J. Psychophysiol. Spec. Rep., 36*, 211–236.

Knight, R. T. (1996). Contribution of human hippocampal region to novelty (P3a) detection. *Nature, 383*, 256-259.

Knight, R. T., Scabini, D., Woods, D. L., & Clayworth, C. (1988). The effects of lesions of superior temporal gyrus and inferior parietal lobe on temporal and vertex components of the human AEP. *Electroencephalography and Clinical Neurophysiology, 70*(6), 499–509.

Kok, A. (1997). Event-related potentials ( ERP) reflections of mental resources: a review and synthesis. *Biol. Psychol., 45,* 19-56.

Kramer, A. F., Wickens, C. D., & Donchin, E. (1985). Processing of stimulus properties: evidence for dual-task integrality. *Journal of Experimental Psychology Human Perception Performance, 11*, 393– 408.

Kujala, T., Myllyviita, K., Tervaniemi, M., Alho, K., Kallio, J., & Näätänen, R. (2000). Basic auditory dysfunction in dyslexia as pinpointed by brain-activity measurements [Special Report]. *Psychophysiology, 37*, 262–266.

Kulkarni, A., & Colburn, H. S. (1998). Role of spectral detail in sound-source localization. *Nature, 396*(6713), 721-724.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation. *Science, 197*, 792-795.

Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: emotion, attention, capacity, and memory. *J. Broadcast. Electron. Media, 40*, 460-477.

Lange, K., Krämer, U. M., & Röder, B. (2006). Attending points in time and space. *Exp. Brain Research*.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol: Hum. Percept. Perf., 21*(48), 451-468.

Lebib, R., Papo, D., Douiri, A., de Bode, S., Downes, M. G., & Baudonnière, P.-M. (2004). Modulations of ‘late’ event-related brain potentials in humans by dynamic audiovisual speech stimuli. *Neurscience Letters, 372*, 74-79.

Leppert, D., Goodin, D. S., & Aminoff, M. J. (2003). Stimulus recognition and its relationship to the cerebral event-related potential. *Neurology, 61*, 1533-1537.

Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory, 10*(5/6), 397-403.

Luck, S. J. (2005a). Averaging, Artifact Rejection, and Artifact Correction. In M. S. Gazzaniga (Ed.), *An Introduction To The Event-Related Potential Technique* (pp. 131-174). Camebride, Massachusetts, London, England: MIT Press.

Luck, S. J. (2005b). An Introduction to Event-Related Potentials and Their Neural Origins. In M. S. Gazzaniga (Ed.), *An Introduction To The Event-Related Potential Technique* (pp. 1-50). Camebride, Massachusetts, London, England: MIT Press.

Luck, S. J. (2005c). The Operation of Attention - Millisecond by Millisecond - Over the First Half Second. In H. Ogmen & B. G. Breitmeyer (Eds.), *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes*. Cambridge, MA: MIT Press.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and Clinical Neurophysiology, 75*, 528-542.

Luck, S. J., & Hillyard, S. A. (1995). The role of attention in feature detection and conjunction discrimination: an electrophysiological analysis. *Int. J. Neurosci. Biobehav. Rev., 80*, 281– 297.

Luck, S. J., & Hillyard, S. A. (1999). The operation of selective attention at multiple stages of processing: evidence from human and monkey electrophysiology. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (2nd ed., pp. 687-700). Boston: MIT Press.

Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology Human Perception Performance, 20*, 887-904.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4*(11), 432-440.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention in humans. *Psychophysiology, 32*, 4-18.

Mangun, G. R., & Hillyard, S. A. (1991). Modulation of sensory-evoked brain potentials provide evidence for changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology Human Perception Performance, 17*, 1057-1074.

McCarthy, G., & Donchin, E. (1981). A metric for thought: A comparison of P300 latency and reaction time. *Science, 211*, 77-80.

McKenzie, D. N., & Barry, R. J. (2005). The Independence of Memory Traces of Attended and Unattended Stimuli. *Cerebral Cortex*.

McPherson, W. B., & Holcomb, P. J. (1992). Semantic priming with pictures and the N400 component. *Psychophysiology, 29*, 51.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology, 36*, 53–65.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). Invited review - EEG source imaging. *Clinical Neurophysiology, 115*, 2195-2222.

Michie, P. T., Bearpark, H. M., Crawford, J. M., & Glue, L. C. (1990). The nature of selective attention effects on auditory event-related potentials. *Biol. Psychol., 30*(3), 219-250.

Middlebrooks, J., & Green, D. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology, 42*, 135-139.

Moray, N. (1970). *Attention: Selective Processes in Vision and Hearing*. New York: Academic Press.

Müller, B. W., Jüptner, M., Jentzen, W., & Müller, S. P. (2002). Cortical activation to auditory mismatch elicited by frequency deviant and complex novel sounds: a PET study. *NeuroImage, 17*, 231–239.

Münte, T. F., Berger, D., Terkamp, C., Schofl, C., Johannes, S., & Brabant, G. (1995). Cognitive functioning in experimental hypoglycaemia assessed with event-related potentials. *NeuroReport, 6*(11), 1509-1512.

Münte, T. F., Kohlmetz, C., Nager, W., & Altenmüller, E. (2001). Superior auditory spatial tuning in conductors. *Nature, 409*, 580.

Näätänen, R. (1982). Processing negativity: an evoked-potential reflection of selective attention. *Psych. Bull., 92*, 606-640.

Näätänen, R. (1988). Implications of ERP data for psychological theories of attention. *Biological Psychology, 26*, 117-163.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related brain potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences, 13*, 201-288.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Näätänen, R., & Alho, K. (2004). Mechanisms of Attention in audition as Revealed by the Event-Related Potentials of the Brain. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 194-206). New York

London: The Guilford Press.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologia, 42*, 313-329.

Näätänen, R., & Michie, P. T. (1979). Early selective attention effects on the evoked potential: A critical review and reinterpretation. *Biological Psychology, 8*, 81-136.

Näätänen, R., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375–425.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends Neuroscience Letters, 24*, 283–288.

Nager, W., Estorf, K., & Münte, T. F. (2006). Crossmodal attention effects on brain responses to different stimulus classes. *BMC Neuroscience, 7*(31), 1-8.

Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornells, A., & Münte, T. F. (2003). The fate of sounds in conductors' brains: an ERP study. *Cognitive Brain Research, 17*, 83-93.

Nager, W., Teder-Sälejärvi, W., Kunze, S., & Münte, T. F. (2003). Preattentive evaluation of multiple perceptual streams in human audition. *NeuroReport, 14*, 871-874.

Naumann, E., Huber, C., Maier, S., Plihal, W., Wustmanns, A., Diedrich, O., et al. (1992). The scalp topography of p300 in the visual and auditory modalities: a comparison of three normalization methods and the control of statistical type II error. *Encephalography and clinical Neurophysiology, 83*, 254-264.

Neisser, U. (1967). *Cognitive Psychology*. New York, NY: Appleton-Century-Crofts.

Nittono, H., Hamada, A., & Hori, T. (2003). Brain potentials after clicking a mouse: a new psychophysiological approach to human–computer interaction. *Human Factors, 45,* 591– 599.

Novak, G., Ritter, W., & Vaughan, H. G. J. (1992). The chronometry of attention-modulated processing and automatic mismatch detection. *Psychophysiology, 29*, 412-430.

Oades, R. D., Ditmann-Balcar, A., & Zerbin, D. (1997). Development and topography of auditory event-related potentials (ERPs): Mismatch and processing negativity in individuals 8-22 years of age. *Psychophysiology, 34*, 677-693.

Paavilainen, P., Karlsson, M. L., Reinikainen, K., & Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and Clinical Neurophysiology, 73*(2), 129-141.

Paavilainen, P., Tiitinen, H., Alho, K., & Näätänen, R. (1993). Mismatch negativity to slight pitch changes outside strong attentional focus. *Biol. Psychol., 37*(1), 23-41.

Papanicolaou, A. C., & Johnstone, J. (1984). Probe evoked potentials: theory, method and applications. *International Journal of Neuroscience, 24*, 107-131.

Parasuraman, R. (1980). Effects of information processing demands on slow negative shift latencies and N100 amplitude in selective and divided attention. *Biological Psychology, 11*, 217-233.

Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences, 95*(3), 853-860.

Pfefferbaum, A., Christensen, C., Ford, J. M., & Kopell, B. (1986). Apparent response incompatibility effects on P300 latency depend on the task. *Electroencephalography and Clinical Neurophysiology, 64*, 424-437.

Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: different water in the same river. *Audiol. Neurootol., 5*(3-4), 111-139.

Polich, J. (1996). Meta-analysis of P300 normative aging studies. *Psychophysiology, 33*, 334-353.

Polich, J. (2004). Neuropsychology of P3a and P3b: A Theoretical Overview. In N. C. Moore & K. Arikan (Eds.), *Brainwaves and mind: Recent developments* (pp. 15-29). Wheaton, IL: Kjellberg Inc.

Posner, M. I. (2004). *Cognitive Neuroscience of Attention*. New York, London: The Guilford Press.

Pralong, D., & Carlille, S. (1994). Measuring the human head-related transfer functions: a novel method for the construction and calibration of a miniature "in-ear" recording system. *Journal of Acoustic Society of America, 95*(6), 3435-3444.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M., Pardo, J. V., Fox, P. T., et al. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex, 4*(1), 8-26.

Rif, J., Hari, R., Hämäläinen, M., & Sams, M. (1991). Auditory attention affects two different areas in the human supratemporal cortex. *Electroencephalography and Clinical Neurophysiology, 79*, 464–472.

Rinne, T., Alho, K., Ilmeniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity source. *NeuoImage, 12*, 14-19.

Ritter, W., Sussman, E., & Molholm, S. (2000). Evidence that the mismatch negativity system works on the basis of objects. *NeuroReport, 11*, 61-63.

Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature, 400*(6740), 162-166.

Rosenfeld, J. P., Bhat, K., Miltenberger, A., & Johnson, M. (1992). Event-related potentials in the dual talk paradigm: P300 discriminates engaging and non-engaging films when film-viewing is the primary task. *Int. J. Psychophysiol., 12*, 221-232.

Rugg, M. D., & Coles, M. G. H. (1995). The ERP and cognitive psychology: conceptual issues. In M. D. C. Rugg, M.G.H. (Ed.), *Electrophysiology of the mind: Event-related brain potentials and cognition* (pp. 27-39). Oxford: Oxford University Press.

Rüsseler, J., Kowalczuk, J., Johannes, S., Wieringa, B. M., & Münte, T. F. (2002). Cognitive brain potentials to novel acoustic stimuli in adult dyslexic readers. *Dyslexia, 8*, 125-142.

Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research, 12,* 171– 180.

Scherg, M., & Berg, P. (1991). BESA - brain electromagnetic source analysis.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *J. Cogn. Neurosci., 1*, 336– 355.

Scholl, B. J. (2001). Object and attention: the state of the art. *Cognition, 80*, 1-46.

Schröger, E. (1994). Human brain potential signs of selection by location and frequency in an auditory transient attention situation. *Neuroscience Letters, 173*, 163-166.

Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *J. Cogn. Neurosci., 8,* 527–539.

Schröger, E. (1997). On the detection of auditory deviations: a pre-attentive activation model. *Psychophysiology, 34*(3), 245-257.

Schröger, E., Giard, M.-H., & Wolff, C. (2000). Auditory distraction: event-related potential and behavioral indices. *Clinical Neurophysiology, 111*, 1450-1460.

Schröger, E., & Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related brain potentials. *NeuroReport, 9*, 3355-3358.

Schwent, V. L., & Hillyard, S. A. (1975). Evoked potential correlates of selective attention with multi-channel auditory inputs. *Electroencephalography and Clinical Neurophysiology, 38,* 131-138.

Schwent, V. L., Hillyard, S. A., & Galambos, R. (1976). Selective attention and the auditory vertex potential.II. Effects of signal intensity and masking noise. *Electroencephalography and Clinical Neurophysiology, 30*, 615-622.

Schwent, V. L., Snyder, E., & Hillyard, S. A. (1976). Auditory evoked potentials during multichannel selective listening: Role of pitch and localization cues. *Journal of Experimental Psychology, 2*, 313-325.

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J. Neurosci., 23*(13), 5545–5552.

Shelley, A. M., Ward, P. B., Michie, P. T., Andrews, S., Mitchell, P. F., Catts, S. V., et al. (1991). The effect of repeated testing on ERP components during auditory selective attention. *Psychophysiology, 28,* 496-510.

Shestakova, A., Ceponiene, R., Huotilainen, M., & Yaguchi, K. (2002). Involuntary attention in children as a function of sound source location: evidence from event-related potentials. *Clinical Neurophysiology, 113*, 162-168.

Shinn-Cunningham, B., & Ihlefeld, B. (2004). *Selective and Divided Attention: Extracting Information from Simultaneous Sound Sources.* Paper presented at the ICAD 04 - Tenth Meeting of the International conference on Auditory Display, Sydney, Australia.

Smid, H., Jakob, A., & Heinze, H. J. (1997). The organization of multidimensional selection on the basis of color and shape: an event-related brain potential study. *Perception and Psychophysics, 59*, 693-713.

Smith, E. M., Halgren, E., Sokolik, M., Baudena, P., Musolino, A., Liegeois-Chauvel, C., et al. (1990). The intracranial topography of the P3 event-related potential elicited during auditory oddball. *Electroencephalography and Clinical Neurophysiology, 76*, 235-248.

Sonnadara, R., Alain, C., & Trainor, L. J. (2006). Effects of spatial separation and stimulus probability on the event-related potentials elicited by occasional changes in sound location. *Brain Research, 1071*, 175-185.

Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology, 28*, 343-358.

Spitzer, A. R., Cohen, L. G., Fabrikant, J., & Hallett, M. (1989). A method for determining optimal interelectrode spacing for cerebral topographic mapping. *Electroencephalography and Clinical Neurophysiology, 72*, 355–361.

Squires, N. K., Squires, K. C., & Hillyard, A. (1975). Two varieties of long latency positive waves evokes by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology, 38*, 387-401.

Srinivasan, R. (2005). High-Resolution EEG: Theory and Practice. In T. C. Handy (Ed.), *Event-Related Potentials. A Methods Handbook*. Cabridge, Massachusetts, London, England: The MIT Press.

Srinivasan, R., Nunez, P. L., & Tucker, D. M. (1998). Estimating the spatial Nyquist of the human EEG. *Behav. Res. Meth. Inst. Comp., 30*, 8–19.

Styles, E. A. (1997). *Psychology Of Attention* (1st ed.). Buckinghamshire College, Bucks, U.K.: Taylor and Francis (Psychology Press).

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked potential correlates of stimulus uncertainty. *Science, 150*, 1187–1188.

Suzuki, J., Nittono, H., & Hori, T. (2005). Level of interest in video clips modulates event-related potentials to auditory probes. *International Journal of Psychophysiology, 55*, 35-43.

Teder, W., Kujala, T., & Näätänen, R. (1993). Selection of speech messages in free-field listening. *NeuroReport, 5*, 307-309.

Teder, W., & Näätänen, R. (1994). Event-related potentials demonstrate a narrow focus of auditory spatial attention. *NeuroReport, 5*, 709-711.

Teder-Sälejärvi, W., & Hillyard, S. A. (1998). The gradient of spatial auditory attention in free field: An event-related potential study. *Perception and Psychophysics, 60*(7), 1228-1242.

Teder-Sälejärvi, W. A., Hillyard, S. A., Röder, B., & Neville, H. J. (1999). Spatial attention to central and peripheral auditory stimuli as indexed by event-related potentials. *Cognitive Brain Research, 8*, 213-227.

Tervaniemi, M., Radil, T., Radilova, J., & Kujala, T. (1999). Pre-Attentive Discriminability of Sound Order as a Function of Tone Duration and Interstimulus Interval: A Mismatch Negativity Study. *Audiology Neuro-Otology, 4*, 303-310.

Tiitinen, H., Alho, K., Huotilainen, M., Ilmoniemi, R. J., Simola, J., & Näätänen, R. (1993). Tonotopic auditory cortex and the magnetoencephalographic (MEG) equivalent of the mismatch negativity. *Psychophysiology, 30*, 537–540.

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature, 372*, 90–92.

Tinazzi, M., Zanette, G., La Porta, F., Polo, A., Volpato, D., Fiaschi, A., et al. (1997). Selective gating of lower limb cortical somatosensory evoked potentials (SEPs) during passive and active foot movements. *Electroencephalography and Clinical Neurophysiology, 104*(4), 312-321.

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*, 242-248.

Treisman, A. (1964). Selective attention in man. *British Medical Bulletin, 20*, 12-16.

Treisman, A., & Geffen, G. (1967). Selective attention: Perception or response? *Quarterly Journal of Experimental Psychology, 19,* 1-18.

Treisman, A., & Riley, J. (1969). Is selective attention selective perception or selective response? A further test. *Journal of Experimental Psychology, 79*, 27-34.

Trejo, L., Ryan-Jones, D. L., & Kramer, A. F. (1995). Attentional modulation of the mismatch negativity elicited by frequency differences between binaurally presented tone bursts. *Psychophysiology, 32*, 319-328.

Valdes-Sosa, M., Bobes, M. A., Rodriguez, V., & Pinilla, T. (1998). Switching Attention without Shifting the Spotlight: Object-Based Attentional Modulation of Brain Potentials. *Journal of Cognitive Neuroscience, 10*(1), 137-151.

Van Voorhis, S. T., & Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Perception and Psychophysics, 22*, 54-62.

Wenzel, E. M., Wightman, F. L., Kistler, D. J., & Foster, S. H. (1988). Acoustic origins of individual differences in sound localization behavior. *Journal of the Acoustical Society of America, 84,* 79.

West, R., Herndon, R. W., & Crewdson, S. J. (2001). Neural activity associated with the realization of a delayed intention. *Cognitive Brain Research, 12,* 1-9.

Wickens, C. D., Kramer, A. F., Vanasse, L., & Donchin, E. (1983). performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science, 221,* 1080-1082.

Wightman, F. L., & Kistler, D. J. (1989a). Headphone simulation of free-field listening I: Stimulus synthesis. *Journal of the Acoustical Society of America, 85*(2), 858-867.

Wightman, F. L., & Kistler, D. J. (1989b). Headphone simulation of free-field listening II: Psychophysical validation. *Journal of the Acoustical Society of America, 85*(2), 868-878.

Wijers, A. A., Mulder, G., Gunter, T. C., & Smid, H. G. O. (1996). *Die hirnelektrische Analyse der selektiven Aufmerksamkeit* (2nd ed.). Göttingen: Hogrefe.

Winkler, I., & Näätänen, R. (1995). The effects of auditory backward masking on event-related brain potentials. In G. Karmos, V. Csépe, I. Czigler, M. Molnar & J. Desmedt (Eds.), *Perspectives of ERP Research, EEG Suppl.* (Vol. 44, pp. 185-189). Amsterdam: Elsevier.

Winkler, I., Tervaniemi, M., Schröger, E., Wolff, C., & Näätänen, R. (1998). Preattentive processing of auditory spatial information in humans. *Neuroscience Letters, 242,* 49-52.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences, 90,* 8722-8726.

Woldorff, M. G., Hackley, S. A., & Hillyard, S. A. (1991). The effects of Channel-Selective Attention on the Mismatch Negativity Wave Elicited by Deviant Tones. *Psychophysiology, 28*(1), 30-42.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology, 79*, 170–191.

Woldorff, M. G., Hillyard, S. A., Gallen, C. C., Hampson, S. R., & Bloom, F. E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology, 35*, 283-292.

Woods, D., Alho, K., & Algazi, A. (1994). Stages of auditory feature conjunction: An event-related potential study. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 81–94.

Woods, D. L. (1990). The physiological basis of selective attention: Implications of event-related potential studies. In J. W. P. Rohrbaugh, R.; Johnson, R.J. (Ed.), *Event-related brain potentials: Basic issues and applications*. NY: Oxford University Press.

Woods, D. L., & Alain, C. (1992). Feature processing during high-rate auditory selective attention. *Perception and Psychophysics*, 1-34.

Woods, D. L., & Alain, C. (1993). Feature processing during high-rate auditory selective attention. *Perception and Psychophysics, 53*(4), 391-402.

Woods, D. L., Alho, K., & Algazi, A. (1991). Brain potential signs of feature processing during auditory selective attention. *NeuroReport, 2*, 189-192.

Woods, D. L., & Clayworth, C. C. (1987). Scalp topographies dissociate N1 and Nd components during auditory selective attention. *Electroencephalography and Clinical Neurophysiology Supplement, 40*, 155-160.

Woods, D. L., Hillyard, S. A., & Hansen, J. C. (1984). Event-related brain potentials reveal similar attentional mechanisms during selective listening and shadowing. *Journal of Experimental Psychology Human Perception Performance, 10*, 761-777.

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon. Bull. Rev., 9*(2), 185-211.

Wurtz, R. H., & Goldberg, M. E. (1972). The primate superior colliculus and the shift of visual attention. *Investigative ophthalmology, 11*(6), 441-450.

# Appendix

## A1: A list of stories used in each experiment.

| | left stories | right stories |
|---|---|---|
| ***Exp. I*** | Angelika Schrobsdorff - **Von der Erinnerung geweckt**<br>*speaker: Angelika Schrobsdorff* | Noëlle Châtelet - **Die Dame in Blau**<br>*speaker: Marlen Diekhoff* |
| ***Exp. II*** | Hugo Verlomme - **Die Nacht der Delphine**<br>*speaker: Edgar M. Böhlke* | Theodor Storm - **Halligfahrt**<br>*speaker: Peter Gregor* |
| ***Exp. III*** | Ernest Hemingway - **Der alte Mann und das Meer**<br>*speaker: Rolf Boysen* | Antoine de Saint-Exupéry - **Nachtflug**<br>*speaker: Gert Westphal* |
| ***Exp. IV*** | Birgit Vanderbeke - **Geld oder Leben**<br>*speaker (target): Birgit Vanderbeke* | Tracy Chevalier - **Das Mädchen mit dem Perlenohrring**<br>*speaker (target): Stefanie Stappenbeck* |
| | Robert Schneider - **Schlafes Bruder**<br>*speaker (non-target): Fritz Hammel* | Sándor Márai - **Ein Hund mit Charakter**<br>*speaker (non-target): Charles Bauer* |
| ***Exp. V*** | Nuala O'Faolain - **Ein alter Traum von Liebe**<br>*speaker: Marlen Diekhoff* | Ingrid Noll - **Selige Witwen**<br>*speaker: Franziska Pigulla* |

## A2: Questionnaires of the control experiment as a sample.

Experiment V:

Right story

### „Ein alter Traum von Liebe“ von Nuala O'Faolain

| Zeit | Frage | Antwort | richtig | falsch |
|---|---|---|---|---|
| 0:16 | Wo lebten die Erzählerin und Hugo? | Mansardenwohnung; London | | |
| 0:45 | Was studierte die Erzählerin? | Journalistik | | |
| 1:38 | Was warf Hugo auf das Bett? | Die Unterlagen von einer Gerichtsverhandlung | | |
| 2:12 | Mit wem soll Frau Talvert die Ehe gebrochen haben? | Einem Hausangestellten; William Mullon; Stallbursche; Kutscher | | |
| 2:40 | Wo haben die Zeugen das Liebespaar gesehen? | Im Stall | | |
| 3:33 | Was interessierte die Erzählerin an dem Fall? | Die Leidenschaft der Liebenden | | |
| 4:05 | Lebt die Erzählerin gern in London? | Nein | | |
| 4:55 | Mit wem war die Erzählerin eng befreundet? | Jimmy aus Amerika | | |
| 5:40 | Mit wem kam die Erzählerin auf dem Flughafen ins Gespräch? | Mit einem Geschäftsmann | | |
| 6:20 | Was sagte der Geschäftsmann über die Afrikaner? | Sie haben Rhythmus im Blut | | |
| 8:00 | Wer rief an? | Der Schweizer Geschäftsmann | | |
| 8:12 | Wie nannte er die Erzählerin? | Kätzchen | | |
| 9:04 | Wen versuchte die Erzählerin in London zu erreichen? | Jimmy | | |
| 10:43 | Was sagt die Erzählerin zu Jimmy wie er aussehe? | Wie James Dean | | |
| 11:58 | Wohin möchte Jimmy mit der Erzählerin ziehen, wenn sie alt sind? | South Beach | | |
| 12:31 | Was erfuhr die Erzählerin während der Programmsitzung? | Dass Jimmy tot ist | | |
| 12:35 | Woran ist er gestorben? | An einem Herzanfall | | |
| 14:03 | Was tat die Erzählerin nach Jimmys Beerdigung? | Schrieb seine und ihre Artikel | | |
| 15:02 | Welchen Entschluss behielt die Erzählerin in London für sich? | TravelRight zu verlassen | | |
| 15:45 | Welche Tiere betrachtete die Erzählerin im | Affen | | |

| | | | | |
|---|---|---|---|---|
| | Zoo? | | | |
| 16:28 | Wo stammt die Familie der Erzählerin her? | Aus Irland | | |
| 17:38 | Was kam der Erzählerin in den Sinn als sie den Zoo verließ? | Ein Gedicht | | |
| 18:25 | Wem sagte die Erzählerin als erstes dass sie TravelRight verlassen wolle? | Alex; ihrem Chef | | |
| 19:30 | Was könnte die Erzählerin schreiben? | Buch | | |
| 20:46 | Was waren die einzigen Situationen, in denen die Erzählerin die Wohnung gemocht hatte? | Am Morgen nach langen Reisen | | |
| 21:17 | Woran erkannte Jimmy, dass die Erzählerin aus der Kellerwohnung kam? | Ihr hafte die Dunkelheit an | | |
| 21:37 | Wie oft ist die Erzählerin in den letzten 30 Jahren umgezogen? | Dreimal | | |
| 22:12 | Welches Auto fuhr die Freundin der Erzählerin? | Jeep | | |
| 23:50 | Was betete die Erzählerin? | „Mach, dass es nicht zu spät ist.“ | | |
| 24:38 | Was betrachtete die Erzählerin am Himmel? | Wolke | | |
| 26:23 | Welches Auto leiht sich die Erzählerin? | Audi | | |
| 27:28 | Wie heißt das Hotel, zu dem die Erzählerin fuhr? | „Half-Way“ | | |
| 28:22 | Welches Geschlecht hatte das Kind von Mrs. Talvert? | Weiblich | | |
| 30:07 | Wovon lenkte Irland die Erzählerin ab? | Von Jimmys Tod | | |
| 31:37 | Was dachte die Erzählerin beim Anblick der Bäume? | Dass das die Talvert Bäume sind | | |
| 32:20 | Welche Pflanzen standen in der Vase in der Empfangshalle? | Weidenkätzchen | | |
| 33:15 | Warum sind alle Zimmer des Talvert-Hotels belegt? | Eine Hochzeitsgesellschaft reist an | | |
| 34:04 | Wem ist der Hotelbesitzer noch nie begegnet? | Einem Talvert | | |
| 35:14 | Mit wem telefonierte die Erzählerin im Hotel? | Mrs. Leatch | | |
| 36:36 | Was bietet Mrs. Leatch der Erzählerin an? | Ein paar Unterlagen zum Talvert Fall | | |
| 36:51 | Wann soll die Erzählerin zu Mrs. Leatch kommen? | Zwei Uhr | | |
| 37:46 | Was zitierte die Erzählerin wörtlich nach der Talvert- Akte? | Die Zeugenaussagen | | |
| 39:00 | Wie viele Zeugenaussagen zitiert sie? | 5 | | |
| 40:32 | Warum ist die Erzählerin plötzlich so aufgewühlt? | Weil Jimmy tot ist | | |
| 41:33 | Was hatte Berti auf dem Arm? | Ein Kind | | |
| 42:17 | Worüber mussten Berti und die Erzählerin lachen? | Über den Hund | | |
| 44:40 | Was erfuhr die Erzählerin von Mrs. Leatch über die Familie Mullon? | Sie waren sehr angesehen | | |
| 45:02 | Mit welchem geschichtlichen Ereignis in Irland wollte sich die Erzählerin befassen? | Hungersnot | | |
| 45:08 | Was hatte Mrs. Leatch zum 150. Jahrestag der Hungersnot vorbereitet? | Eine Ausstellung dazu | | |
| 47:18 | Was schien für die Talvert Geschichte besser geeignet zu sein? | Roman | | |
| 47:43 | Was wollte sich Mrs. Leatch zusammen mit der Erzählerin ansehen? | Moorlandschaft um Mount Talvert herum | | |
| 48:30 | Worauf deutete Mrs. Leatch als sie aus dem Auto gestiegen war? | Überreste von Häusern | | |
| 49:09 | Wie lebten die Menschen damals am Rande des Moors? | Sehr dicht zusammen | | |
| 49:36 | Wie ist das Wetter über Mount Talvert? | schlecht | | |
| 50:21 | Wodurch wurde das Land nach Mrs. Leatchs Meinung ruiniert? | Kunst, Antiquitäten | | |
| 51:24 | Wen rief die Erzählerin im Hotel an? | Alex | | |
| 53:28 | Warum zog sich die Erzählerin an? Was wollte sie machen? | In der Stadt spazieren gehen | | |

…

Left story

## „Selige Witwen“ von Ingrid Noll

| Zeit | Frage | Antwort | richtig | falsch |
|---|---|---|---|---|
| 0:21 | Was geschah mit Coras Ehemann? | Sie haben ihn ermordet | | |
| 1:12 | Wo leben Cora und die Erzählerin? | Florenz | | |
| 1:20 | Wie heißt der Sohn der Erzählerin? | Bela | | |
| 1:47 | Worüber sprachen die Leute am Nachbartisch in dem Lokal? | Über den plötzlichen Unfalltod eines Engländers und den Verkauf seines Hauses | | |
| 3:09 | Was will Cora von dem Handwerker wissen? | Was das Haus kosten soll | | |
| 3:22 | Was denkt Cora über den Preis des Hauses? | Es ist fast geschenkt | | |
| 4:30 | Was wollte der Handwerker am nächsten Tag tun? | Er wollte den beiden das Haus bei Tageslicht zeigen und sie dann nach Florenz zurück fahren | | |
| 4:46 | Was taten die beiden nachdem der Handwerker ins Bett gegangen ist? | Sie wollten im Pool schwimmen | | |
| 5:08 | Wer schaltete plötzlich den Scheinwerfer an? | Der Handwerker Dino | | |
| 6:20 | Wohin sollte Dino nach dem Schwimmen gehen? | Ins Bett | | |
| 6:38 | Was dachte die Erzählerin als sie am nächsten Morgen auf die Terrasse trat? | Wie im Paradies, die Toskana sei der schönste Teil der Welt | | |
| 7:25 | Wen stellte sich die Erzählerin im Garten vor? | Ihren Sohn Bela | | |
| 7:35 | Was machte der Erzählerin im Garten in Bezug auf ihren Sohn Sorgen? | Der Pool | | |
| 8:17 | Wer lag in dem Bett bei Cora? | Dino | | |
| 8:46 | Warum dachte die Erzählerin, dass Cora und Dino sich über sie lustig machten? | Sie hörte ihren Namen und Gelächter | | |
| 9:05 | Welche Haarfarbe hat Cora? | rot | | |
| 9:58 | Warum fährt Dino kurz mit dem Wagen weg? | Er holt Frühstück | | |
| 10:25 | Wo befand sich Cora nach dem Frühstück? | Bibliothek | | |
| 11:52 | Für wen sollte das überdachte Gästehaus sein? | Emilia und Mario | | |
| 11:58 | Welcher Teil des Hauses gefiel den beiden am besten? | Wintergarten | | |
| 12:15 | Welche Anstellung hatte der Großvater von Dino in diesem Haus? | Gärtner | | |
| 12:50 | Was war das Hobby des Engländers? | Technische Geräte, vor allem sein Computer | | |
| 13:35 | Wie nannte man den Engländer im Dorf? | Il Barone | | |
| 14:07 | Wie ist der Engländer umgekommen? | Im Pool ertrunken | | |
| 14:40 | Was hat man bei der Obduktion angeblich gefunden? | Spuren eines Schlafmittels | | |
| 15:00 | Was vermutete Cora sofort? | Mord | | |
| 15:28 | Was lastet seit dem Tod des Engländers auf dem Haus? | Ein Fluch | | |
| 16:26 | Was markiert üblicherweise das Ende eines Bauernhofes? | Zwei Zypressen | | |
| 16:55 | Was wollte die Erzählerin im Haus genauer inspizieren? | Die Bibliothek | | |
| 17:43 | Was geschah als die Erzählerin in der Bibliothek Musik hörte? | Sie brach in Tränen aus | | |
| 17:50 | Wie heißt der Vater von Bela? | Jonas | | |
| 18:30 | Bis wann wollten sie auf dem Anwesen bleiben? | Montag | | |
| 18:46 | Was malte Cora? | Dino | | |
| 19:19 | Wie findet Dino den Neffen des Engländers? | unangenehm | | |
| 19:30 | Wen hat der Neffe des Engländers sofort entlassen? | Lucia, die Haushälterin | | |

| | | | | |
|---|---|---|---|---|
| 20:41 | Was hat Il Barone gesammelt? | Puppen | | |
| 20:46 | Welches Geschlecht hatten alle Puppen? | männlich | | |
| 21:43 | Wohin fuhren sie am Montagmorgen? | Sienna | | |
| 22:54 | Wohin flüchteten die beiden vor dem Regen? | In den Dom | | |
| 23:10 | Was faszinierte die Erzählerin besonders im Dom? | Fußboden | | |
| 24:34 | Mit welchem argument versuchte Cora den Preis zu drücken? | Ein Fluch lastet auf dem Haus | | |
| 26:26 | Wie heißt die Erzählerin? | Maja | | |
| 27:50 | Was sagte der Makler als Cora anrief um das Haus zu kaufen? | Es ist verkauft | | |
| 28:02 | Wer hat das Haus gekauft? | Eine Amerikanerin | | |
| 29:04 | Als was hatte die Erzählerin gearbeitet? | Deutschsprachige Stadtführerin in Florenz | | |
| 30:37 | An was war Coras Großmutter erkrankt? | Lungenentzündung | | |
| 31:45 | Wen wollte Cora ihrer Großmutter als ihren Sohn vorstellen? | Bela | | |
| 31:56 | Wie alt ist Bela? | vier | | |
| 32:38 | Wie heißt die Großmutter? | Charlotte Schwab | | |
| 32:40 | Wo wohnt die Großmutter? | Darmstadt | | |
| 32:56 | Wer öffnete die Tür? | Felix, Coras Vetter | | |
| 34:54 | Was schlägt Felix als Schlafplatz für Cora und Maja vor? | Seine WG | | |
| 36:00 | Was kochte Felix in der WG? | Pasta | | |
| 36:54 | Welches Hausstier lebte in der WG? | Hund | | |
| 38:00 | Wohin möchte Cora mit Felix fahren? | Toskana | | |
| 38:36 | Wer soll die Großmutter und den Hund betreuen? | Maja | | |
| 39:45 | Was hielt Maja davon, dass sie in Darmstadt blieb? | Fand es nicht gut | | |
| 40:23 | Wo wohnte Felix? | In einem ehemaligen Friseursalon | | |
| 41:13 | Als was arbeitete Andy? | Taxifahrer | | |
| 42:40 | Zu wem fuhr Maja als erstes? | Zur Großmutter | | |
| 43:06 | Welche Krankheit hatte die Großmutter in Wahrheit gehabt? | Bronchitis | | |
| 44:24 | Was wollte die Großmutter über Cora wissen? | Was sie arbeitet | | |
| 45:35 | Was stellt Maja beim Einkaufen fest? | Dass sie kein Geld hatte | | |
| 46:28 | Was hatte Maja bei der Großmutter mitgenommen? | Ein Buttermesser | | |
| 47:18 | Wo hatte Felix noch Geld versteckt? | In einem Schuhkarton | | |
| 47:32 | Wie viel Geld hatte Felix in dem Schuhkarton? | 300 Mark | | |
| 48:28 | Wer betrat abends die Wohnung? | Die Mitbewohnerin; Allerleirauh | | |
| 48:40 | Was war das Besondere an Allerleirauh? | Schnurrbart | | |
| 49:44 | Wie nannte ihr Bruder die Erzählerin? | Elefantin | | |
| 50:58 | Wen rief die Erzählerin am nächsten Morgen an? | Emilia | | |
| 51:28 | Wohin wollten Cora und Felix am nächsten Tag fahren? | Zum Meer | | |
| 52:48 | Wohin wollte Andy Maja und ihren Sohn hinfahren? | Ins Schwimmbad | | |
| 53:09 | Was fand Maja im Keller? | Ein Fahrrad | | |
| 54:25 | Wie kam Maja an ein besseres Fahrrad mit Kindersitz? | Sie stahl es vor einem Kindergarten | | |
| 54:57 | Wer rief Maja an? | Die Großmutter | | |

…

## A3: Answer probability of the questionnaires in each experiment.

### a) Experiment I:

| subject | right story | | left story | | sum correct | sum incorrect | % correct | % incorrect |
|---|---|---|---|---|---|---|---|---|
| | correct | incorrect | correct | incorrect | | | | |
| AB01_1 | 13 | 4 | 16 | 5 | 29 | 9 | 76.32 | 23.68 |
| AK28_1 | 15 | 1 | 21 | 1 | 36 | 2 | 94.74 | 5.26 |
| AP08_1 | 10 | 10 | 16 | 8 | 26 | 18 | 59.09 | 40.91 |
| AZ30_1 | 15 | 7 | 23 | 4 | 38 | 11 | 77.55 | 22.45 |
| CH24_1 | 16 | 4 | 16 | 4 | 32 | 8 | 80.00 | 20.00 |
| CJ04_1 | 13 | 6 | 16 | 8 | 29 | 14 | 67.44 | 32.56 |
| CK07_1 | 16 | 1 | 17 | 3 | 33 | 4 | 89.19 | 10.81 |
| DH29_1 | 17 | 5 | 14 | 9 | 31 | 14 | 68.89 | 31.11 |
| FB13_1 | 14 | 1 | 21 | 1 | 35 | 2 | 94.59 | 5.41 |
| FD27_1 | 13 | 6 | 20 | 2 | 33 | 8 | 80.49 | 19.51 |
| FK22_1 | 20 | 3 | 15 | 7 | 35 | 10 | 77.78 | 22.22 |
| IG03_1 | 18 | 3 | 21 | 2 | 39 | 5 | 88.64 | 11.36 |
| JB14_1 | 20 | 3 | 21 | 2 | 41 | 5 | 89.13 | 10.87 |
| JT02_1 | 6 | 9 | 14 | 6 | 20 | 15 | 57.14 | 42.86 |
| JT21_1 | 17 | 4 | 20 | 5 | 37 | 9 | 80.43 | 19.57 |
| KA16_1 | 13 | 6 | 18 | 12 | 31 | 18 | 63.27 | 36.73 |
| NH33_1 | 15 | 1 | 20 | 6 | 35 | 7 | 83.33 | 16.67 |
| NI23_1 | 17 | 3 | 16 | 1 | 33 | 4 | 89.19 | 10.81 |
| PK26_1 | 17 | 3 | 17 | 2 | 34 | 5 | 87.18 | 12.82 |
| RG09_1 | 17 | 2 | 21 | 2 | 38 | 4 | 90.48 | 9.52 |
| RG11_1 | 17 | 4 | 21 | 5 | 38 | 9 | 80.85 | 19.15 |
| RK19_1 | 22 | 3 | 22 | 12 | 44 | 15 | 74.58 | 25.42 |
| RM05_1 | 14 | 12 | 9 | 4 | 23 | 16 | 58.97 | 41.03 |
| SF25_1 | 16 | 3 | 21 | 3 | 37 | 6 | 86.05 | 13.95 |
| SN12_1 | 15 | 4 | 16 | 4 | 31 | 8 | 79.49 | 20.51 |
| SR10_1 | 16 | 6 | 12 | 12 | 28 | 18 | 60.87 | 39.13 |
| SS15_1 | 20 | 0 | 23 | 1 | 43 | 1 | 97.73 | 2.27 |
| SS18_1 | 12 | 4 | 22 | 1 | 34 | 5 | 87.18 | 12.82 |
| SS32_1 | 17 | 1 | 19 | 0 | 36 | 1 | 97.30 | 2.70 |
| VL17_1 | 17 | 2 | 19 | 2 | 36 | 4 | 90.00 | 10.00 |
| VR31_1 | 14 | 3 | 16 | 7 | 30 | 10 | 75.00 | 25.00 |
| YM06_1 | 12 | 7 | 13 | 14 | 25 | 21 | 54.35 | 45.65 |
| **total sum** | | | | | **1070** | **286** | **79** | **21** |

### b) Experiment II:

| subject | right story | | left story | | sum correct | sum incorrect | % correct | % incorrect |
|---|---|---|---|---|---|---|---|---|
| | correct | incorrect | correct | incorrect | | | | |
| AB01_2 | 12 | 19 | 33 | 4 | 45 | 23 | 66.18 | 33.82 |
| AK28_2 | 14 | 17 | 25 | 9 | 39 | 26 | 60.00 | 40.00 |
| AP08_2 | 18 | 13 | 24 | 8 | 42 | 21 | 66.67 | 33.33 |
| AZ30_2 | 18 | 10 | 34 | 6 | 52 | 16 | 76.47 | 23.53 |
| CH24_2 | 13 | 14 | 22 | 2 | 35 | 16 | 68.63 | 31.37 |
| CJ04_2 | 12 | 20 | 27 | 5 | 39 | 25 | 60.94 | 39.06 |
| CK07_2 | 13 | 17 | 26 | 9 | 39 | 26 | 60.00 | 40.00 |
| DH29_2 | 10 | 16 | 26 | 12 | 36 | 28 | 56.25 | 43.75 |
| FB13_2 | 12 | 19 | 27 | 6 | 39 | 25 | 60.94 | 39.06 |
| FD27_2 | 22 | 13 | 22 | 11 | 44 | 24 | 64.71 | 35.29 |
| FK22_2 | 21 | 19 | 21 | 15 | 42 | 34 | 55.26 | 44.74 |
| IG03_2 | 11 | 15 | 21 | 5 | 32 | 20 | 61.54 | 38.46 |
| JB14_2 | 12 | 17 | 32 | 4 | 44 | 21 | 67.69 | 32.31 |
| JT02_2 | 9 | 20 | 20 | 19 | 29 | 39 | 42.65 | 57.35 |
| JT21_2 | 15 | 17 | 28 | 6 | 43 | 23 | 65.15 | 34.85 |
| KA16_2 | 15 | 17 | 29 | 3 | 44 | 20 | 68.75 | 31.25 |
| NH33_2 | 16 | 14 | 35 | 4 | 51 | 18 | 73.91 | 26.09 |
| NI23_2 | 19 | 11 | 29 | 9 | 48 | 20 | 70.59 | 29.41 |
| PK26_2 | 21 | 10 | 31 | 1 | 52 | 11 | 82.54 | 17.46 |
| RG09_2 | 18 | 6 | 23 | 2 | 41 | 8 | 83.67 | 16.33 |
| RG11_2 | 22 | 10 | 32 | 7 | 54 | 17 | 76.06 | 23.94 |
| RK19_2 | 18 | 9 | 32 | 1 | 50 | 10 | 83.33 | 16.67 |
| RM05_2 | 12 | 20 | 22 | 11 | 34 | 31 | 52.31 | 47.69 |
| SF25_2 | 9 | 15 | 17 | 7 | 26 | 22 | 54.17 | 45.83 |
| SN12_2 | 19 | 15 | 19 | 5 | 38 | 20 | 65.52 | 34.48 |
| SR10_2 | 12 | 17 | 25 | 12 | 37 | 29 | 56.06 | 43.94 |
| SS15_2 | 15 | 8 | 16 | 9 | 31 | 17 | 64.58 | 35.42 |
| SS18_2 | 16 | 14 | 35 | 6 | 51 | 20 | 71.83 | 28.17 |
| SS32_2 | 6 | 23 | 24 | 3 | 30 | 26 | 53.57 | 46.43 |
| VL17_2 | 13 | 11 | 22 | 5 | 35 | 16 | 68.63 | 31.37 |
| VR31_2 | 12 | 21 | 23 | 12 | 35 | 33 | 51.00 | 49.00 |
| YM06_2 | 10 | 17 | 27 | 5 | 37 | 22 | 62.71 | 37.29 |
| **total sum** | | | | | **1294** | **707** | **65** | **35** |

### c) Experiment III:

| subject | right story | | left story | | sum correct | sum incorrect | % correct | % incorrect |
|---|---|---|---|---|---|---|---|---|
| | correct | incorrect | correct | incorrect | | | | |
| AB01_3 | 15 | 8 | 25 | 3 | 40 | 11 | 78.43 | 21.57 |
| AK28_3 | 12 | 12 | 20 | 3 | 32 | 15 | 68.09 | 31.91 |
| AP08_3 | 12 | 4 | 11 | 12 | 23 | 16 | 58.97 | 41.03 |
| AZ30_3 | 11 | 6 | 16 | 2 | 27 | 8 | 77.14 | 22.86 |
| CH24_3 | 4 | 16 | 24 | 0 | 28 | 16 | 63.64 | 36.36 |
| CJ04_3 | 13 | 17 | 18 | 8 | 31 | 25 | 55.36 | 44.64 |
| CK07_3 | 17 | 5 | 19 | 0 | 36 | 5 | 87.80 | 12.20 |
| FB13_3 | 12 | 6 | 14 | 3 | 26 | 9 | 74.29 | 25.71 |
| FD27_3 | 9 | 13 | 20 | 2 | 29 | 15 | 65.91 | 34.09 |
| FK22_3 | 10 | 12 | 23 | 1 | 33 | 13 | 71.74 | 28.26 |
| IG03_3 | 12 | 8 | 16 | 2 | 28 | 10 | 73.68 | 26.32 |
| JB14_3 | 9 | 15 | 25 | 2 | 34 | 17 | 66.67 | 33.33 |
| JT02_3 | 15 | 17 | 26 | 14 | 41 | 31 | 56.94 | 43.06 |
| JT21_3 | 3 | 16 | 19 | 6 | 22 | 22 | 50.00 | 50.00 |
| KA16_3 | 10 | 21 | 34 | 10 | 44 | 31 | 58.67 | 41.33 |
| NH33_3 | 8 | 9 | 24 | 1 | 32 | 10 | 76.19 | 23.81 |
| NJ23_3 | 21 | 5 | 21 | 5 | 42 | 10 | 80.77 | 19.23 |
| PK26_3 | 12 | 7 | 20 | 3 | 32 | 10 | 76.19 | 23.81 |
| RG09_3 | 20 | 4 | 27 | 1 | 47 | 5 | 90.38 | 9.62 |
| RG11_3 | 19 | 8 | 21 | 2 | 40 | 10 | 80.00 | 20.00 |
| RK19_3 | 12 | 21 | 29 | 4 | 41 | 25 | 62.12 | 37.88 |
| RM05_3 | 11 | 15 | 23 | 1 | 34 | 16 | 68.00 | 32.00 |
| SF25_3 | 17 | 15 | 24 | 3 | 41 | 18 | 69.49 | 30.51 |
| SN12_3 | 16 | 6 | 21 | 4 | 37 | 10 | 78.72 | 21.28 |
| SR10_3 | 15 | 15 | 21 | 3 | 36 | 18 | 66.67 | 33.33 |
| SS15_3 | 15 | 6 | 17 | 5 | 32 | 11 | 74.42 | 25.58 |
| SS18_3 | 5 | 5 | 8 | 1 | 13 | 6 | 68.42 | 31.58 |
| SS32_3 | 14 | 5 | 22 | 2 | 36 | 7 | 83.72 | 16.28 |
| VL17_3 | 12 | 6 | 17 | 2 | 29 | 8 | 78.38 | 21.62 |
| VR31_3 | 13 | 3 | 15 | 3 | 28 | 6 | 82.35 | 17.65 |
| YM06_3 | 11 | 8 | 15 | 2 | 26 | 10 | 72.22 | 27.78 |
| **total sum** | | | | | **1020** | **424** | **71** | **29** |

### d) Experiment IV:

| subject | right story | | left story | | sum correct | sum incorrect | % correct | % incorrect |
|---|---|---|---|---|---|---|---|---|
| | correct | incorrect | correct | incorrect | | | | |
| AB01_4 | 11 | 8 | 13 | 9 | 24 | 17 | 58.54 | 41.46 |
| AK28_4 | 8 | 13 | 13 | 9 | 21 | 22 | 48.84 | 51.16 |
| AP08_4 | 7 | 4 | 15 | 2 | 22 | 6 | 78.57 | 21.43 |
| AZ30_4 | 14 | 6 | 17 | 5 | 31 | 11 | 73.81 | 26.19 |
| CH24_4 | 11 | 8 | 15 | 10 | 26 | 18 | 59.09 | 40.91 |
| CJ04_4 | 14 | 6 | 10 | 11 | 24 | 17 | 58.54 | 41.46 |
| CK07_4 | 12 | 8 | 12 | 9 | 24 | 17 | 58.54 | 41.46 |
| DH29_4 | 4 | 12 | 2 | 15 | 6 | 27 | 18.18 | 81.82 |
| FB13_4 | 12 | 8 | 14 | 4 | 26 | 12 | 68.42 | 31.58 |
| FD27_4 | 6 | 11 | 10 | 11 | 16 | 22 | 42.11 | 57.89 |
| FK22_4 | 3 | 15 | 5 | 15 | 8 | 30 | 21.05 | 78.95 |
| IG03_4 | 9 | 5 | 12 | 5 | 21 | 10 | 67.74 | 32.26 |
| JB14_4 | 14 | 4 | 18 | 5 | 32 | 9 | 78.05 | 21.95 |
| JT02_4 | 5 | 7 | 1 | 14 | 6 | 21 | 22.22 | 77.78 |
| JT21_4 | 8 | 10 | 7 | 9 | 15 | 19 | 44.12 | 55.88 |
| KA16_4 | 5 | 13 | 10 | 11 | 15 | 24 | 38.46 | 61.54 |
| NH33_4 | 16 | 3 | 6 | 12 | 22 | 15 | 59.46 | 40.54 |
| NI23_4 | 13 | 6 | 12 | 6 | 25 | 12 | 67.57 | 32.43 |
| PK26_4 | 10 | 9 | 13 | 7 | 23 | 16 | 58.97 | 41.03 |
| RG09_4 | 8 | 7 | 15 | 1 | 23 | 8 | 74.19 | 25.81 |
| RG11_4 | 13 | 7 | 17 | 5 | 30 | 12 | 71.43 | 28.57 |
| RK19_4 | 10 | 10 | 10 | 12 | 20 | 22 | 47.62 | 52.38 |
| RM05_4 | 7 | 8 | 11 | 7 | 18 | 15 | 54.55 | 45.45 |
| SF25_4 | 10 | 5 | 13 | 4 | 23 | 9 | 71.88 | 28.13 |
| SN12_4 | 13 | 2 | 13 | 3 | 26 | 5 | 83.87 | 16.13 |
| SR10_4 | 11 | 8 | 12 | 9 | 23 | 17 | 57.50 | 42.50 |
| SS15_4 | 13 | 7 | 10 | 11 | 23 | 18 | 56.10 | 43.90 |
| SS18_4 | 8 | 8 | 8 | 8 | 16 | 16 | 50.00 | 50.00 |
| SS32_4 | 11 | 4 | 8 | 4 | 19 | 8 | 70.37 | 29.63 |
| VL17_4 | 13 | 4 | 14 | 6 | 27 | 10 | 72.97 | 27.03 |
| VR31_4 | 9 | 9 | 13 | 9 | 22 | 18 | 55.00 | 45.00 |
| YM06_4 | 15 | 3 | 11 | 13 | 26 | 16 | 61.90 | 38.10 |
| **total sum** | | | | | **683** | **499** | **58** | **42** |

**A4: Array of different conditions in the randomized stimulus presentation for the three experiments ((a) II, (b) IV and (c) V) where the early sensory effects occurred. Frequency of how often a certain condition follows one of the others; with an averaged ISI (in ms) between those stimuli.**

a) Experiment II (oddball-paradigm)

| | | FREQUENCY | AVERAGE ISI (in ms) |
|---|---|---|---|
| **standards *following* standards** | | | |
| | A+L+F+ _ after _ A+L+F+ | 14517 | 654.78 |
| | A+L+F+ _ after _ A-L+F+ | 14363 | 654.32 |
| | A-L+F+ _ after _ A-L+F+ | 14341 | 655.32 |
| | A-L+F+ _ after _ A+L+F+ | 14338 | 654.17 |
| **standards *following* deviants** | | | |
| | A+L+F+ _ after _ A+L+15F+ | 886 | 653.43 |
| | A+L+F+ _ after _ A+L+30F+ | 883 | 650.06 |
| | A+L+F+ _ after _ A-L+15F+ | 863 | 645.50 |
| | A+L+F+ _ after _ A-L+30F+ | 902 | 653.70 |
| | A-L+F+ _ after _ A+L+15F+ | 940 | 651.22 |
| | A-L+F+ _ after _ A+L+30F+ | 909 | 655.34 |
| | A-L+F+ _ after _ A-L+15F+ | 909 | 653.33 |
| | A-L+F+ _ after _ A-L+30F+ | 884 | 656.12 |
| | A+L+F+ _ after _ A+L+F+60 | 877 | 648.07 |
| | A+L+F+ _ after _ A+L+F+400 | 914 | 658.13 |
| | A+L+F+ _ after _ A-L+F+60 | 674 | 651.44 |
| | A+L+F+ _ after _ A-L+F+400 | 885 | 659.88 |
| | A-L+F+ _ after _ A+L+F+60 | 918 | 658.24 |
| | A-L+F+ _ after _ A+L+F+400 | 905 | 660.70 |
| | A-L+F+ _ after _ A-L+F+60 | 637 | 654.98 |
| | A-L+F+ _ after _ A-L+F+400 | 914 | 653.85 |
| **deviants *following* standards** | | | |
| | A+L+15F+ _ after _ A+L+F+ | 900 | 657.25 |
| | A+L+30F+ _ after _ A+L+F+ | 916 | 651.22 |
| | A+L+15F+ _ after _ A-L+F+ | 911 | 644.77 |
| | A+L+30F+ _ after _ A-L+F+ | 904 | 644.40 |
| | A-L+15F+ _ after _ A+L+F+ | 918 | 653.34 |
| | A-L+30F+ _ after _ A+L+F+ | 866 | 651.98 |
| | A-L+15F+ _ after _ A-L+F+ | 901 | 654.38 |
| | A-L+30F+ _ after _ A-L+F+ | 941 | 656.66 |
| | A+L+F+60 _ after _ A+L+F+ | 893 | 655.39 |
| | A+L+F+400 _ after _ A+L+F+ | 883 | 659.11 |
| | A+L+F+60 _ after _ A-L+F+ | 909 | 657.77 |
| | A+L+F+400 _ after _ A-L+F+ | 865 | 651.60 |
| | A-L+F+60 _ after _ A+L+F+ | 631 | 655.52 |
| | A-L+F+400 _ after _ A+L+F+ | 888 | 656.54 |
| | A-L+F+60 _ after _ A-L+F+ | 666 | 655.43 |
| | A-L+F+400 _ after _ A-L+F+ | 910 | 652.96 |
| **deviants *following* deviants** | | | |
| | A+L+15F+ _ after _ A+L+15F+ | 40 | 688.77 |
| | A+L+15F+ _ after _ A+L+30F+ | 58 | 651.60 |
| | A+L+15F+ _ after _ A+L+F+400 | 48 | 659.20 |
| | A+L+15F+ _ after _ A+L+F+60 | 60 | 658.87 |
| | A+L+15F+ _ after _ A-L+15F+ | 53 | 624.58 |
| | A+L+15F+ _ after _ A-L+30F+ | 60 | 643.96 |
| | A+L+15F+ _ after _ A-L+F+400 | 57 | 644.46 |
| | A+L+15F+ _ after _ A-L+F+60 | 45 | 668.78 |
| | A+L+30F+ _ after _ A+L+15F+ | 64 | 666.44 |
| | A+L+30F+ _ after _ A+L+30F+ | 42 | 667.36 |
| | A+L+30F+ _ after _ A+L+F+400 | 50 | 644.44 |
| | A+L+30F+ _ after _ A+L+F+60 | 51 | 662.09 |
| | A+L+30F+ _ after _ A-L+15F+ | 56 | 648.14 |
| | A+L+30F+ _ after _ A-L+30F+ | 63 | 649.45 |
| | A+L+30F+ _ after _ A-L+F+400 | 58 | 660.74 |
| | A+L+30F+ _ after _ A-L+F+60 | 30 | 686.45 |
| | A+L+F+400 _ after _ A+L+15F+ | 53 | 649.96 |
| | A+L+F+400 _ after _ A+L+30F+ | 61 | 648.48 |
| | A+L+F+400 _ after _ A+L+F+400 | 58 | 685.85 |
| | A+L+F+400 _ after _ A+L+F+60 | 53 | 634.53 |
| | A+L+F+400 _ after _ A-L+15F+ | 89 | 665.02 |
| | A+L+F+400 _ after _ A-L+30F+ | 53 | 718.86 |

|  |  |  |  |
|---|---|---|---|
|  | A+L+F+400 _ after _ A-L+F+400 | 64 | 640.93 |
|  | A+L+F+400 _ after _ A-L+F+60 | 43 | 675.68 |
|  | A+L+F+60 _ after _ A+L+15F+ | 61 | 658.29 |
|  | A+L+F+60 _ after _ A+L+30F+ | 49 | 674.55 |
|  | A+L+F+60 _ after _ A+L+F+400 | 55 | 684.86 |
|  | A+L+F+60 _ after _ A+L+F+60 | 65 | 665.84 |
|  | A+L+F+60 _ after _ A-L+15F+ | 46 | 650.90 |
|  | A+L+F+60 _ after _ A-L+30F+ | 59 | 653.48 |
|  | A+L+F+60 _ after _ A-L+F+400 | 54 | 610.41 |
|  | A+L+F+60 _ after _ A-L+F+60 | 41 | 660.88 |
|  | A-L+15F+ _ after _ A+L+15F+ | 36 | 703.03 |
|  | A-L+15F+ _ after _ A+L+30F+ | 71 | 628.39 |
|  | A-L+15F+ _ after _ A+L+F+400 | 48 | 704.00 |
|  | A-L+15F+ _ after _ A+L+F+60 | 52 | 652.13 |
|  | A-L+15F+ _ after _ A-L+15F+ | 66 | 644.47 |
|  | A-L+15F+ _ after _ A-L+30F+ | 58 | 648.08 |
|  | A-L+15F+ _ after _ A-L+F+400 | 54 | 654.61 |
|  | A-L+15F+ _ after _ A-L+F+60 | 33 | 631.67 |
|  | A-L+30F+ _ after _ A+L+15F+ | 52 | 625.39 |
|  | A-L+30F+ _ after _ A+L+30F+ | 59 | 661.36 |
|  | A-L+30F+ _ after _ A+L+F+400 | 50 | 628.66 |
|  | A-L+30F+ _ after _ A+L+F+60 | 53 | 636.93 |
|  | A-L+30F+ _ after _ A-L+15F+ | 50 | 622.24 |
|  | A-L+30F+ _ after _ A-L+30F+ | 59 | 667.98 |
|  | A-L+30F+ _ after _ A-L+F+400 | 54 | 632.73 |
|  | A-L+30F+ _ after _ A-L+F+60 | 48 | 663.39 |
|  | A-L+F+400 _ after _ A+L+15F+ | 62 | 670.92 |
|  | A-L+F+400 _ after _ A+L+30F+ | 58 | 696.80 |
|  | A-L+F+400 _ after _ A+L+F+400 | 55 | 651.74 |
|  | A-L+F+400 _ after _ A+L+F+60 | 57 | 669.19 |
|  | A-L+F+400 _ after _ A-L+15F+ | 55 | 668.63 |
|  | A-L+F+400 _ after _ A-L+30F+ | 53 | 658.37 |
|  | A-L+F+400 _ after _ A-L+F+400 | 48 | 708.09 |
|  | A-L+F+400 _ after _ A-L+F+60 | 43 | 687.74 |
|  | A-L+F+60 _ after _ A+L+15F+ | 43 | 646.00 |
|  | A-L+F+60 _ after _ A+L+30F+ | 38 | 680.98 |
|  | A-L+F+60 _ after _ A+L+F+400 | 44 | 690.34 |
|  | A-L+F+60 _ after _ A+L+F+60 | 44 | 621.51 |
|  | A-L+F+60 _ after _ A-L+15F+ | 39 | 653.07 |
|  | A-L+F+60 _ after _ A-L+30F+ | 44 | 645.70 |
|  | A-L+F+60 _ after _ A-L+F+400 | 44 | 631.89 |
|  | A-L+F+60 _ after _ A-L+F+60 | 54 | 653.65 |
| **total** |  | **88724** | **656.84** |

b) Experiment IV (equiprobable presentation)

|  |  | FREQUENCY | AVEARAGE ISI (ms) |
|---|---|---|---|
| **target-probes** ***following*** **target-probes** |  |  |  |
|  | A+S+L+ _ after _ A+S+L+ | 4890 | 653.93 |
|  | A+S+L+ _ after _ A-S-L+ | 4888 | 657.80 |
|  | A-S-L+ _ after _ A+S+L+ | 4937 | 654.54 |
|  | A-S-L+ _ after _ A-S-L+ | 4879 | 658.10 |
| **target-probes** ***following*** **non-target-probes** |  |  |  |
|  | A+S+L+ _ after _ A-S+L15 | 4746 | 657.34 |
|  | A+S+L+ _ after _ A-S-L15 | 4955 | 659.18 |
|  | A-S-L+ _ after _ A-S+L15 | 4918 | 657.15 |
|  | A-S-L+ _ after _ A-S-L15 | 4750 | 657.45 |
| **non-target-probes** ***following*** **target-probes** |  |  |  |
|  | A-S+L15 _ after _ A+S+L+ | 4758 | 655.71 |
|  | A-S+L15 _ after _ A-S-L+ | 4835 | 653.20 |
|  | A-S-L15 _ after _ A+S+L+ | 4895 | 658.01 |
|  | A-S-L15 _ after _ A-S-L+ | 4884 | 657.29 |
| **non-target-probes** ***following*** **non-target-probes** |  |  |  |
|  | A-S+L15 _ after _ A-S+L15 | 4963 | 654.81 |
|  | A-S+L15 _ after _ A-S-L15 | 4926 | 655.00 |
|  | A-S-L15 _ after _ A-S+L15 | 4852 | 656.84 |
|  | A-S-L15 _ after _ A-S-L15 | 4844 | 659.20 |
| **total** |  | **77920** | **656.60** |

### c) Experiment V (control experiment: equiprobable presentation)

| | | **FREQUENCY** | **AVERAGE ISI in ms** |
|---|---|---:|---:|
| **Phoneme-probes** *following* **phoneme-probes** | | | |
| | A+phon _ after _ A+phon | 5065 | 747.46 |
| | A+phon _ after _ A-phon | 5068 | 745.08 |
| | A-phon _ after _ A+phon | 5088 | 748.51 |
| | A-phon _ after _ A-phon | 5127 | 745.04 |
| **Phoneme-probes** *following* **click-probes** | | | |
| | A+phon _ after _ A+click | 5182 | 782.99 |
| | A+phon _ after _ A-click | 5140 | 780.25 |
| | A-phon _ after _ A+click | 5090 | 780.22 |
| | A-phon _ after _ A-click | 5146 | 775.29 |
| **Click-probes** *following* **click-probes** | | | |
| | A+click _ after _ A+click | 5131 | 780.28 |
| | A+click _ after _ A-click | 5047 | 780.98 |
| | A-click _ after _ A+click | 5049 | 781.06 |
| | A-click _ after _ A-click | 5122 | 780.77 |
| **Click-probes** *following* **phoneme-probes** | | | |
| | A+click _ after _ A+phon | 5181 | 747.35 |
| | A+click _ after _ A-phon | 5100 | 744.87 |
| | A-click _ after _ A+phon | 5123 | 743.23 |
| | A-click _ after _ A-phon | 5159 | 748.32 |
| **total** | | **81818** | **763.23** |

## A5: Early sensory effects found in experiment II and IV

### Experiment II

### *Deviants vs. standard*

Overall and interestingly, the deviant probe stimuli seemed to cause an initial sensory effect compared to the standard phonemes (L+F+) on both sides (A+ and A-). Shortly after stimulus presentation (< 100 ms), both types of spatial and frequency deviants (slight and extreme) on the attended side showed a negative deflection at around 50 ms after stimulus onset (N50). Almost the same pattern was the case for the unattended side, with exception of the slight frequency deviant probes which caused a positivity instead (P20). Furthermore, there was no difference between A-L+F+400 and A-L+F+60 probe stimuli. Detailed pair-wise analyses are reported in the next paragraphs.

1) Short-latency ERPs – spatial and frequency deviant stimuli – attended side

Before 100 ms after stimulus onset, the deviant stimuli showed a negative deflection compared to standard probes mainly. On the attended side, ANOVAs yielded statistical significances for the comparisons between the probe stimuli.

Extreme frequency deviants caused a N50 when compared to standard probes (main effect of Condition: F(1,14) = 27.42, pHF = 0.0001) with a frontal and ipsilateral distribution (maximal effect at electrode F3/4i (F(1,14) = 35.78 pHF = 0.0001)). In comparison with slight frequency deviants which were similar to standards, the A+L+F400 ERP was more negative in general (effect of Condition: F(1,14) = 4.79 pHF = 0.0460) with a fronto-central emphasis to stimulus presentation (maximal effect of Condition: Fc5/6i (F(1,14) = 7.86 pHF = 0.0141)).

Between spatial deviants (A+L+15F+ and A+L+30F+), no significant difference could be found.

2) Short-latency ERPs – spatial and frequency deviant stimuli – unattended side

On the unattended side, ANOVAs yielded statistical significances for the comparisons between probe stimuli as well.

Extreme frequency deviants caused an N50 compared to standard probes (effect of Condition: F(1,14) = 6.67 pHF = 0.0217) whereas slight frequency deviants elicited a significant P20 instead with a parietal maximum (effect of Condition at Pz: F(1,14) = 17.44 pHF = 0.0009; vs. Fz: F(1,14) = 10.91 pHF = 0.0052).

Again no significant difference between A-L+15F+ and A-L+30F+ was calculated.

3) Discussion

These components that have not been expected to differ between conditions represented the initial sensory effects before 100 ms after stimulus onset. The P20, and N50 components (marginally significant) which are thought to arise from the medial geniculate nucleus and the primary auditory cortex (Luck, 2005b) were found only for both kinds of deviant probe stimuli (spatial and frequency) with their two-level gradation (slight and extreme) compared to the standard (frequency and spatial congruent) stimuli on both sides, attended and unattended. How can these early effects be explained? Was it maybe a kind of overlay effect by a preceding presented stimulus? The ISI used in this experiment was quite short and varied between 250 and 750 ms randomly. By having a stimulus length of 100 ms each, it could have been possible that a previous stimulus was still processed while the next probe already occurred (Luck, 2005a). Nevertheless, some overlap may have still occurred apart from jittering the stimulus onset, but there are several arguments against this assumption. First, by using a high pass-filter (0.5 Hz) those remaining overlaps should have been filtered out as well unless the filter was not high enough. Furthermore, the lack of such early effects for the standard probe stimuli would also not fit in that overlap point of view. Normally, they would have been influenced by such overlaps as well, because of a jittered ISI and a randomized stimulus presentation.

By jittering of stimulus timing (interstimulus interval, ISI) the relatively sharp positive and negative peaks at the beginning of the ERP waveforms are thought to be eliminated. Thus, jittering works like a filter for high frequencies in the case of a small time range. By using a longer time range, the lower frequencies are reduced more and more from the overlap as well. Nevertheless, some low-frequency overlap may still occur even with a broad jitter. (For more information see Luck, 2005a).

Furthermore, only experiment 2 reported those early sensory differences between conditions. Experiment 1 with the same jittering of ISIs, stimulus randomization and high pass filtering processes did not cause these components, although it also contained frequency deviant stimuli. The only difference between both studies was that the probability of stimulus presentation differed: In experiment 1, frequency deviants occurred equiprobable to frequency-congruent probes, whereas the deviant stimuli in experiment 2 were less probable (5% each) compared to standards (80%). Therefore, a further suggestion would be that the deviant probe stimuli may have followed another deviant in a row. The deviant ERP, due to the stimulus' differing physical characteristics and probability compared to standards, may have contained more energy which could influence the following stimulus processing enormously. Because of a randomized stimulus presentation, deviants and standards in the current study were distributed in such a manner that the deviants followed a standard stimulus more often than another deviant (see also appendix A4a). Thus, the quite decent preceding standard ERP could not be the reason for those P20, or N50 effects in the following deviant ERP.

Nevertheless, a control experiment was conducted to investigate the possibility of an overlap effect or to rule it out as a possible explanation for the P20, and N50 occurrence. Therefore, a similar experimental setup was chosen whereby the probe stimuli were much shorter (60 ms) to eliminate an overlapping probability by reducing the perceiving ability of two different tones at the pre-attentive level. In other words, this would cause a lower processing of those pretty short stimuli (Tervaniemi, Radil, Radilova, & Kujala, 1999) with no determining effect to the following probes anymore.

Short (60 ms) equiprobable probe stimuli - a speaker's voice (frequency and spatial) congruent phoneme and a click-sound - were selected for this control experiment with a jittering ISI of 250-750 ms each. Two simultaneous prose streams with those randomized probes (phonemes and clicks) superimposed were presented. The EEG set-up consisted of 32 tin electrodes. The sampling rate was increased up to 1000 Hz instead of 250 Hz like in the other experiments to be able to inspect those early components in more detail and to make it easier to find signals with possibly fast transitions (results see appendix A6).

Furthermore, instead of a normal sampling rate of 250 Hz the EEG was recorded with 1000 Hz for a higher resolution in that early ERP time window. Results of this experiment with a higher temporal resolution showed still the same effects within the first 100 ms after stimulus onset, but slightly smaller, so that they did not reach significance anymore. Click sounds differed with a more negative ERP from phoneme-probes, whereas those phonemes showed no such an early effect. No fast transitions which may have caused the early effects by using the sampling rate of 250 Hz were observed either. Because of an evenly distributed stimulus array and the differences in the early components between both kinds of probes, an overlap effect was really not the reason for those early sensory effects; otherwise they would have been vanished completely or would have occurred for the two different probes in a similar way by this new control setup.

Therefore, another alternative explanation should be approached. It was conspicuous that only the deviants elicited such early effects. That is why one could conclude that their feature characteristics and maybe also their low probability led to a quite obvious and outstanding occurrence of those deviant probes compared to the more in the story embedded standard stimuli. Thus, the P20, and N50 could be a sign for a short term novel detection process. Thus, those rare deviant tones and not the standards could have been able to elicit such surface ERPs, like Iwanami and colleagues investigated on rats (Iwanami, Shinba, Sumi, Ozawa, & Yamamoto, 1994). Other researchers argued for being subcortical potentials and reflecting electrophysiology of the subcortical sensory pathway (Tinazzi et al., 1997). They investigated in the somatosensory field, but as pointed out because of attention being supramodal (top-down controlled (Kastner, 2004)) those findings could be transferred to the present auditory results as well. Furthermore, these researchers proposed a gating function at cortical level beyond those early ERP components. This would support the assumption that the deviants by reason of their physical characteristics are more obvious than the standard stimuli and therefore more able to open a gate for their processing at this early stage on their way along the auditory pathway. Underlying neural generators are congruently found in the auditory cortex like in the present study (see next paragraph) (Iwanami et al., 1994; Tinazzi et al., 1997). Thus, the P20, and N50 components showed an early gating for the present deviant stimuli only, unaffected by attention, because the effects did not differ between both attentional states in the current experiment.

## Experiment IV

### *Non-target vs. target location*

Similar to the second experiment of chapter 2, the spatial deviant probe stimuli (the phonemes within the non-target stream (L15)) of the present study seem also to cause an initial sensory effect compared to the phonemes at the target story's location (L60) on both sides (S+ and S-). Even between the S+L15 and S-L15 probes, the early ERPs appear to be different. Shortly after stimulus presentation (< 100 ms), the spatial deviant on the attended side showed a negative deflection at around 50 ms (N50) after stimulus onset whereas a positivity could be observed on the unattended side in the same time frame (~ 50 ms; P50) compared to the L60 stimuli on the same side respectively. Between both non-target phonemes (L15), the A-S+L15's early negativity contrasts to the positivity of the same stimuli on the unattended side (A-S-L15). No difference is salient between A+S+L60 and A-S-L60.

ANOVAs yielded statistical significances for the comparisons between probe stimuli. The stimuli at the non-target location showed a negative deflection compared to target probes on the attended side (main effect of Condition: $F(1,19) = 7.05$, $p = 0.0156$) whereas the L15 phonemes on the unattended side elicited a P50 compared to L60 probes (main effect of Condition: $F(1,19) = 9.29$, $p = 0.0066$). Comparing both non-target probe stimuli, a significant negative deflection for L15 on the same side as the attentional focus in contrast to the positivity of L15 on the unattended side was confirmed ($F(1,19) = 18.12$, $p = 0.0004$).

### *Discussion*

Similar to the effects in experiment 2.2 (chapter 2), some earlier effects before 100 ms after stimulus onset (P20, N50, P50 (e.g. in de Bruin et al., 2001)) between different kinds of probe stimuli (target vs. non-target probes) on both sides (attended vs. unattended, and target vs. non-target), but not between attended and unattended probe stimuli at the target story's location could be observed. In the oddball-experiment (2.2), it was shown that those early ERP components are not affected by attention, but rather by the stimulus' occurrence probability. In contrast to that oddball-paradigm of experiment 2.2, this current study used equiprobable probe stimuli. Thus, the argument of the stimulus probability is not sufficient. Nevertheless, the present investigation also differed from experiment 2.1 and 2.3 of chapter 2 which had an equiprobable stimulus presentation as well in one way. It contains a four-speaker setting instead of just two speakers which made the selection task more complex and difficult. Maybe there is an interaction between stimulus features and another factor. Changes in stimulus features could not have been enough to elicit such early gating effects (de Bruin et al., 2001), but in combination with another factor, those deviant stimuli became more obvious than frequency- and spatial-congruent probe stimuli. Thus, those deviants' stimulus features (deviant frequency, or spatial features) could have been strengthened by either their rare representation (experiment 2.2) or a more complex situation (more distracting streams; the present experiment 3.1) to be finally able to be gated preferentially.

One argument for the observed P20, and N50 in chapter 2 was that, due to the deviants' low probability, the neural system could not adapt to their rare representation compared to the frequent standard stimuli. Maybe a missing memory trace, the same reason for the occurrence of the MMN and P3, could have caused this sensitivity for those deviant stimuli at this early stage of stimulus processing. With regard to the current experiment, this argument is still valid from the point of view that the four simultaneous streams may have overloaded the short term memory. Thus, a memory trace only for stimuli at the target's location may have been created, because of the priority to attend rather to them than to the distracting non-target streams during the selective listening task. This again would explain why a non-target probe stimulus evoked such early ERP effects compared to the target stimuli although they were presented equiprobable. The neural system could not adapt to those "spatial deviant" stimuli (presented in another stream than the target and its paralleling stream on the opposite side), combined with the more complex and the memory overloading set-up. Thus, whenever the probes at the non-target story's location occurred, the neural system was still quite sensitive to those "deviants" and gated their processing more actively at this very early stage (< 100 ms) (de Bruin et al., 2001).

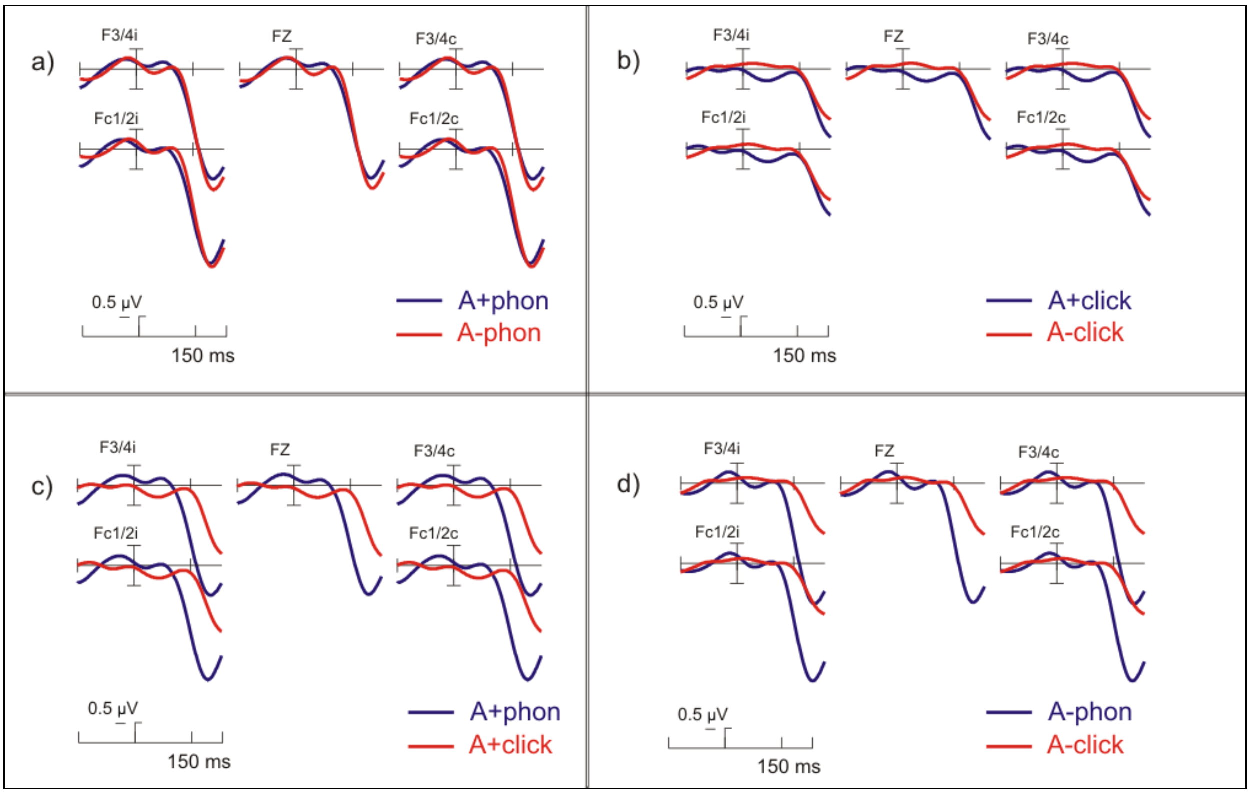
Again, the assumption of an overlap effect because of a too short ISI (Luck, 2005a) can be simply falsified by the observation of missing early components in the case of attended and unattended probe stimuli at the target story's location. The same jittered ISI, randomization and stimulus analyses as in the previous experiment were used which in total should have eliminated overlap effects (Luck, 2005a). Thus, rather the deviant stimulus' features in combination with another factor (insufficient working memory trace for those deviants) could have caused these early effects.

As in experiment 2.2, one alternative assumption was that an array effect may have been the reason for those early effects. This means that the non-target probe stimuli may have followed another non-target in a row more often than a target probe despite a full randomization. The non-target ERP, due to the stimulus' differing physical and spatial characteristics in the four-speaker setting compared to standards, may have contained more energy which could influence the following stimulus processing enormously. Nevertheless, the investigation of the stimulus array could not confirm any conspicuity with regard to this array effect assumption (see appendix A4b). All stimuli followed one of the other different kinds of probes relatively equally.

Altogether, this would support the assumption that the deviants by reason of their different physical characteristics combined with another factor (rare stimulus occurrence and the complexity of a given situation) cause a lack of a respective memory trace. Thus, these deviants, compared to the probes at the target story's location, are more able to open a gate for their processing at this early stage on their way along the auditory pathway (de Bruin et al., 2001).

## A6: Control-experiment: ERPs for the very early time frame (0-150 ms).

The stimuli of this control experiment were a short phoneme and a click sound of the same length (60 ms); ISI jittered between 250 and 750 ms and the sampling rate was 1000 Hz. **(a)** shows ERPs of the phoneme (phon) on either the attended (A+) or unattended (A-) side at frontal electrodes which differed slightly at around 50 ms, but less than 0.5 µV. **(b)** represents ERPs of the click sound (click) on either the attended (A+) or unattended (A-) side at frontal electrodes which differed slightly at around 50 ms, but not more than 0.5 µV. **(c)** maps ERPs of the phoneme (phon) compared to the click sound (click) on the attended (A+) side at frontal electrodes which differed slightly (about 0.5 µV) at around 50 ms. **(d)** maps ERPs of the phoneme (phon) compared to the click sound (click) on the unattended (A-) side at frontal electrodes which differed slightly at around 50 ms, but again less than 0.5 µV. Those visually seen differences **(a-d)** were not strong enough to yield statistical significances.



# Curriculum Vitae

## Personal information:

| | |
|---|---|
| Name: | Dörte Spring (née Kuhlicke) |
| Academic degree: | Diplom-Psychologin |
| Sex: | female |
| Birthday & -place: | Juli 10th 1979, Belzig, Germany |
| Nationality: | German |
| Family status: | married |
| Mail address: | 435 Ithaca Ct., Apt. F6; Las Cruces, NM 88011; USA |
| Email: | Doerte@spring-west.net |
| Phone: | +1 915 539 7312 |

## Education:

Otto von Guericke Universität, Department of Neuropsychology, Magdeburg, Germany
Doctoral program in psychology, April 2003 - present

Otto von Guericke Universität, Department of Neuropsychology, Magdeburg, Germany
Diplom in Psychology, Graduated March 2003
Cumulative GPA 1.2

## Experience:

Otto von Guericke Universität, Lehrstuhl für Neuropsychology
October 2000-March 2003
-Worked as an undergraduate research assistant on various projects within the workgroup
April 2003-December 2006
-Performed MRI, EEG examinations investigating various aspects of the cocktail phenomena.
-Taught undergraduate research methods seminar.
-Taught undergraduate neuropsychological disorders seminar.

Otto von Guericke Universität Medical Campus, Neuropsychology Unit
November 2003-February 2006
-Assisted in neuropsychological diagnosis and treatment of patients as an intern. (November 2003-November 2004)
-Worked as a substitute neuropsychologist, diagnosing and treating patients independently. (December 2004-February 2006)
Otto von Guericke Universität, Neuropsychological Outpatient Clinic
March 2004 – May 2006
-Wrote opinions of patients for the court in numerous cases.
-Continuously treated 2 patients on an outpatient basis.

New Mexico State University, Las Cruces, NM, Department of Psychology, Cognitive laboratory
June 2006 – present
-scientific exchange
-research about the anterior prefrontal cortex with fMRI and EEG

## Ad-hoc reviewer for the following journal:

Cognitive Brain Research

Dörte Spring

Magdeburg, January 2007

Dörte Spring