Pre-attentive processing of room acoustics

DISSERTATION

zur Erlangung des Grades eines Doktors der Philosophie der Faktultät für Geistes- und Sozialwissenschaften der Helmut Schmidt Universität / Universität der Bundeswehr Hamburg

vorgelegt von
Johannes Daniel Frey
aus Hamburg

Hamburg 2017

Hamburg, den 15. Januar 2017
Erstgutachter: Prof. Dr. Thomas Jacobsen

Zweitgutachter: Prof. Dr. Dr. Thomas Köhler

Tag des Abschlusses der Disputation: 25.4.2017
Johannes Daniel Frey

Pre-attentive processing of room acoustics
Acknowledgements

I would firstly like to thank my wife for her continuous support and encouragement. I am looking forward to our lifelong journey and I look back gratefully to our journey we already enjoyed together. Thank you for always being there for me since I first met you. Thank you to all my friends and my whole family for further emotional support.

I would like to thank Prof. Dr. Dr. Köhler for evaluating my work as an expert. Thank you to everyone from my university for assistance with the study for this project or other research experiments: Susan Beudt, Philipp Dehmel, Dr. Thomas Konstantin Jacobsen, Christian Kassyda, Svantje Kähler, Dr. Andreas Löw, Dr. Aquiles Luna-Rodriguez, Stephan Möller, Kirsten Neumann, Prof. Dr. Mike Wendt, Prof. Dr. Udo Zölzer,

I would like to thank the whole Mismatch Negativity community for providing the basis for my research through their commitment to the topic and related topics. Thanks to the subjects who voluntarily participated in the experiments.

Lastly, thank you Prof. Dr. Jacobsen introducing me into this exciting research topic and giving me the opportunity to become part of the Mismatch Negativity community. Without your extensive knowledge and support, my research would not have become what it now is. Especially for the publication in peer-reviewed journals your assistance was of great help.
# Table of contents

1 Preface 08

I Theoretical Section: Introduction of topic and methods 11

2 Auditory processing 13

3 Attention 13

4 Room acoustics 16

5 Event-related potentials (ERPs) and components related to deviant processing 17

5.1 P1 18

5.2 N1 18

5.3 P2 18

5.4 The Mismatch Negativity (MMN) 18

5.5 P3 19

5.6 Late Discriminative Negativity (LDN) 19

6 Research Question and Preview 20

II Empirical Section 22

7 Experiment 1: Automatic detection of unattended changes in room acoustics 24

7.1 Material and Methods 24

7.1.1 Participants 24

7.1.2 Materials 24

7.1.3 Experimental design and procedure 28

7.1.4 Electrophysiological recordings 28

7.1.5 Data analysis 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2</td>
<td>Results</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td><strong>Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences</strong></td>
<td>33</td>
</tr>
<tr>
<td>8.1</td>
<td>Material and Methods</td>
<td>33</td>
</tr>
<tr>
<td>8.1.1</td>
<td>Participants</td>
<td>33</td>
</tr>
<tr>
<td>8.1.2</td>
<td>Materials</td>
<td>33</td>
</tr>
<tr>
<td>8.1.3</td>
<td>Experimental design and procedure</td>
<td>34</td>
</tr>
<tr>
<td>8.1.4</td>
<td>Electrophysiological recordings</td>
<td>35</td>
</tr>
<tr>
<td>8.1.5</td>
<td>Data analysis</td>
<td>36</td>
</tr>
<tr>
<td>8.2</td>
<td>Results</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td><strong>Experiment 3: Replication of Experiment 2 with single tone stimuli</strong></td>
<td>41</td>
</tr>
<tr>
<td>9.1</td>
<td>Material and Methods</td>
<td>41</td>
</tr>
<tr>
<td>9.1.1</td>
<td>Participants</td>
<td>41</td>
</tr>
<tr>
<td>9.1.2</td>
<td>Materials</td>
<td>41</td>
</tr>
<tr>
<td>9.1.3</td>
<td>Experimental design and procedure</td>
<td>42</td>
</tr>
<tr>
<td>9.1.4</td>
<td>Electrophysiological recordings</td>
<td>43</td>
</tr>
<tr>
<td>9.1.5</td>
<td>Data analysis</td>
<td>44</td>
</tr>
<tr>
<td>9.2</td>
<td>Results</td>
<td>45</td>
</tr>
<tr>
<td>III</td>
<td>Discussion and outlook</td>
<td>49</td>
</tr>
<tr>
<td>10.</td>
<td>Discussion</td>
<td>51</td>
</tr>
<tr>
<td>10.1</td>
<td>Discussion Experiment 1</td>
<td>51</td>
</tr>
<tr>
<td>10.2</td>
<td>Discussion Experiment 2</td>
<td>53</td>
</tr>
<tr>
<td>10.3</td>
<td>Discussion Experiment 3</td>
<td>55</td>
</tr>
</tbody>
</table>
# Table of contents

11 General Discussion and Outlook 57

11.1 General Discussion 57

11.2 Outlook 59

References 60

List of abbreviations 66

Appendix 67

A Informed written consent 69

B Questionnaire after the experiments 70

C EEG-Recording report 72

D Matlab code for room simulation 74

E Statistical analyses 76

F Time domain representations of the acoustic stimuli from Experiment 2 and 3 86

G Peer-reviewed papers 87
Preface

1 Preface

The human ability to hear by sensing auditory signals is one of the basic, traditional senses (sight, hearing, taste, smell, touch). A lot of research is conducted on auditory processing in general and on specific auditory dimensions like pitch, duration, loudness and sound source location. Room acoustical changes produce complex interaction with different auditory dimensions of pre-attentive processing of room acoustics. Psychological research on room acoustics is sparse and most research uses behavioral measures. More fundamental research is needed to validate and improve knowledge in this area. When focusing on auditory processing most research focusses on acoustic stimulation in the focus of attention. By looking at room acoustics in attended situations, attention and many other factors can influence the results. So research without the focus of attention on the auditory signals can add new valuable insights.

By using the human event-related brain potential (ERP), via recording of an electroencephalogram (EEG), it is possible to focus research on acoustic stimulation outside the focus of attention. In this thesis, this tool was used as a main research method. Using this method, it can be looked at fundamental and specific aspects of room acoustics. For the three experiments of this thesis different computational methods are used to simulate differing room acoustics.

Previous research has shown that the human auditory system continuously monitors its acoustic environment, detecting a variety of irregularities (e.g., deviance from prior stimulation regularity in pitch, loudness, duration, and (perceived) sound source location). Detection of irregularities can be inferred from a component of the ERP, referred to as the Mismatch Negativity (MMN). The detection can be shown even when participants are instructed to ignore the auditory stimuli or are involved in other tasks (Näätänen, Paavilainen, Rinne, & Alho, 2007).

The three experiments investigated:

1) The first experiment investigated how unattended room acoustical changes can be detected pre-attentively and how the human auditory system monitors its acoustic environment.
2) The second experiment replicated the results from the first experiment and provided additional evidence for the notion of pre-attentive detection of unattended changes in room acoustics.
3) The third experiment replicated previous findings from Experiment 1 and 2 using only single tone stimuli, instead of a sequence of tones per stimulus.

All studies used EEG recordings to generate ERPs used for evaluation of different ERP components. The participants were seated in an electrically and acoustically shielded experimental chamber watching a silent movie with subtitles. For acoustically auditory room stimulation headphones were used, to have better control on how stimuli are perceived.

Part I of this thesis gives a detailed introduction to the topic. Furthermore, the methods are explained in detail. Information about auditory processing in general and automatic auditory processing specifically are given. Subsequently, different relevant EEG components will be presented, especially the MMN.

Part II contains the empirical section presenting all experiments of this thesis with method and results.

Part III combines the findings of each experiment in a discussion. Further a general discussion gives an elaborate concluding evaluation of the results from the different experiments. This section concludes with an outlook for further research.

The central research question is: Do humans process room acoustics pre-attentively?

The basis of this thesis are two articles published in Neuroscience Letters. The experiments from these articles are also the foundation for this thesis:


In the first article Experiment 1 from this thesis is described, in the second article Experiment 2 and 3 are the basis (in the article these experiments are named Experiment 1B and
Preface

Experiment 1A). The formulation of the articles was done by myself, so in the thesis many parts from the articles can be found in this thesis. Some are changed a little and others are literally adopted. Of course, the co-authors of the articles improved some formulations of the articles that I also adopted for my thesis.
Part I

Theoretical Section:

Introduction of topic and methods
2 Auditory processing

Auditory signals are perceived through the outer ear. The pinna focuses sound waves through the ear canal to the eardrum. As a consequence, the eardrum is vibrating like the waveform of the sound. In the middle ear the vibrations are transferred to the inner ear with the help of three bones (ossicles). In the inner ear hair cells release neurotransmitters to the auditory nerve, the signals are transmitted via action potentials and processed in different brain regions. Previous research has shown that auditory processing happens on different levels, on a pre-attentive level and on an attentive level. Auditory processing happens to a certain extent during sleep and serves as an alarm sense. Auditory processing is possible when stimuli are behind a person and not in sight (Müsseler, 2008, p. 65). More about the auditory system can be found in Müsseler (p. 70ff).

3 Attention

In “The Principles of Psychology” James (1890) defines attention as follows:

Everyone knows what attention is. It is the taking possession of 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, and is a condition which has a real opposite in the confused, dazed, scatter brained state which in French is called distraction, and Zerstreuheit in German. (p. 403-404)

Even if everyone knows what attention is, it is still far from being a clear construct. Lots of research still needs to be done to get a deeper insight into this construct. Näätänen (1992) pointed out the importance of Attention: “Attention is one of the most popular constructs in modern cognitive psychology, psychophysiology, and related fields; in fact it is hard to understand how we could do without it” (p. 2).

Not much research on attention was done until 1950, involving research on human performance, looking at how human beings are performing at certain tasks. Another important aspect is how attention is directed. There is conformity that attention can be actively directed
to different sensory input. Näätänen (1992) states the following: “Apparently, we can attend to any aspect of stimuli or their combinations, within the limits of discriminability” (p. 5). Sustained attention or vigilance is the long-term performance of human beings. Looking at declining performance over time relying on attention. This is another important aspect of attention but will not be relevant for this thesis, because our participants do not perform in any way to different stimuli, in fact they are supposed to do the exact opposite and direct their attention away from the auditory stimuli presented via headphones. Telling people not to attend to stimuli usually results in the opposite behavior, participants direct their attention to the stimuli supposed to be ignored. Therefore, in the experiments the participants were instructed to attend to other stimuli known not to influence the measures. The other stimuli are from a different sense (vision) to reduce the effect of divided attention. Näätänen (1992) describes divided attention as follows:

Divided attention occurs when the subject is instructed to monitor simultaneously two or several input sources such as to listen to two dichotically presented stories, that is, one story to one ear and the other story concurrently to the other ear. (p. 5)

Kahneman & Treisman (1984) call the two main problems that an organism must solve corresponding to the adaptive function of selective attention “selective processing” and the “adoption of an appropriate set”. The first mainly having problems with perceptual overload and the second with response incoherence (Kahneman & Treisman 1984). The focus of this thesis will be on “selective processing” because the focus of this thesis is not how human beings respond to stimuli but on how they perceive information. The revival of interest in attention in the 1950s was motivated at least in part by the discovery of surprising limitations in the handling of simultaneous messages by air-traffic controllers and by subjects in dichotic listening tasks (Kahneman & Treisman 1984).

Another aspect is the capacity that is needed for different attention processes. In contrast, automatic processing is believed to be free from capacity limitations and attentive processing is believed to need different amounts of attention depending on many different factors and the perceptual domains. Näätänen (1992) states about processing in audition: “Physical features of auditory stimuli are probably extensively processed even in the absence of attention” (Näätänen, p. 24, 1992).
Treisman (1960) showed that physical stimulus features are processed in the absence of attention. Human beings were better in reporting targets from a concurrent auditory stream when the stimuli were different regarding to some acoustical feature, while the subjects were orally repeating the speech stimuli from the other ear and reporting targets from this ear. “Broadbent’s (1958) original version of the filter theory proposed that firstly, all stimuli are briefly stored and analyzed in parallel for their elementary physical properties (preattentive level, or the S-system)” (Näätänen, p. 29, 1992).

Little is known about how room acoustical stimuli are processed in the brain. About most other auditory dimensions, we know that they are processed at least partly pre-attentively. Primary auditory dimensions that have been found to be processed pre-attentively are pitch (Jacobsen & Schröger, 2001; Näätänen, Gaillard and Mäntysalo, 1878; Sams, Paavilainen, Alho and Näätänen, 1985), duration (Jacobsen & Schröger, 2003; Kaukoranta, Sams, Hari, Hämäläinen and Näätänen, 1989; Näätänen, Paavilainen and Reinikainen, 1989), loudness (Jacobsen, Horenkamp and Schröger, 2003; Näätänen, Paavilainen, Alho, Reinikainen and Sams, 1987) and sound source location (Kaiser, Lutzenberger, Preissl, Ackermann and Birbaumer, 2000; Paavilainen, Karlsson, Reinikainen and Näätänen, 1989; Schröger & Wolff, 1996). Higher-order auditory dimensions have also been reported to be processed pre-attentively, for instance the omission of a tone in a recurrent pattern (Nordby, Hammerborg, Roth and Hugdahl, 1994; Tervaniemi, Saarinen, Paavilainen, Danilova and Näätänen, 1994), or speech stimuli violating abstract phonological rules following a sequence of standard stimuli (Näätänen, Paavilainen, Rinne and Alho, 2007). About secondary auditory dimensions lots of research focusses on how auditory stimuli are perceived. Different models are used to explain the auditory processes and give explanation on how auditory stimuli can be processed so fast. Models are getting more and more advanced as research advances. As a result it is not a mere question on pre-attentive or attentive processing, but the question how pre-attentive processing can be evaluated and experimentally researched.

The MMN, component of the human ERP, is a well-established measure for the investigation of pre-attentive auditory processing. The MMN is widely considered to reflect detection of violations of regularities extracted from the acoustic environment, even when the acoustic stimulation is not in the focus of attention (Näätänen, 1990). In the simplest form, it is observed when a repeated sound (standard) is followed by a differing sound (deviant) at an unpredictable time. A mechanism that constantly monitors the acoustic environment and detects changes is likely to govern this process (Näätänen & Winkler, 1999; Schröger,
Bendixen, Denham, Mill, Böhm and Winkler, 2014). Further detailed aspects of the MMN are explained under point 5.4., to give more theoretical and practical background. This thesis applies the MMN as the main tool for research on auditory processing.

4 Room acoustics

An auditory dimension that bears importance for perception and behaviour, particularly in real life contexts, relates to sound properties arising from the reflecting characteristics of objects that make up the environment of the sound source-perceiver system. In a built-up environment as well as in a considerable portion of the natural environment, virtually all sound is affected by these phenomena, referred to as room acoustics.

Behavioral studies have shown the relevance of differences in room acoustics. Previous research has shown that room acoustics impact both perceptual quality and behavioural performance. For instance, the reverberation time of a room influences the ability to localize sounds, especially for the localisation of continuous broadband noise (Hartmann, 1983). Behavioural relevance of room acoustics has been demonstrated for (musical) sound production (i.e., professional piano players adapted their playing style to varying room acoustics (Bolzinger, Warusfel, Kahle, 1994). Additionally, variations in room acoustics created by virtual rooms differing in size, influence the emotional valence of sounds (Tajadura-Jiménez, Larsson, Väljamäe, Västfjäll and Kleiner, 2010). Consequently, a metric to assess the mental and neural mechanisms underlying the processing of room acoustics could improve research in this field. Many difficult questions are unsolved. How does perception of room acoustics work? How can the brain extract information out of the action potentials to generate a room acoustical perception?

Changes in room acoustics provide important clues about the environment of sound source-perceiver systems. For example, indicating changes in the reflecting characteristics of surrounding objects or the different size of a room by opening or closing elements of the room (e.g. doors, windows). For the virtual simulation of room acoustics the ODEON software uses an image-source method combined with ray tracing to perform the room acoustical calculations for this thesis. This is explained in more detail in the method section. Other modelling with Matlab (see Appendix D) is based on the reverberation time using the room impulse response of a virtual room. Standards and deviants were created by modelling different room acoustics, keeping overall values of basic auditory dimensions constant to
make sure pre-attentive room acoustical detection is not due to other physical auditory differences.

Although the behavioural relevance of the perception of room acoustics has long been established (Blesser & Salter, 2007; Shield, Conetta, Dockrell, Connolly, Cox and Mydlarz, 2015), research on automatic detection of room acoustic related changes in auditory stimulation has not been shown in an experimental setup.

5 Event-related potentials (ERPs) and components related to deviant processing

Neural activity within the brain produces local electrical differences in voltages. These voltage differences can be measured by applying electrodes to the scalp. These electrodes are used for writing an electroencephalogram (EEG) by measuring electrical differences in voltages. The voltage difference recorded by a scalp electrode is the summation of all underlying voltage differences. The influence is reduced the further the source of the voltage difference is away from the scalp electrode. By recording the sum of all voltage differences under a scalp position an EEG is obtained. The more scalp electrodes are used the better the spatial resolution of the EEG.

Due to the fact, that the EEG only shows the summation of voltage differences, it is not clear where exactly the voltage differences are localized in the brain. This method has only low accuracies looking at the spatial dimension. Furthermore, not all changes can be detected due to the fact that voltage differences in one direction can be cancelled out by other voltage differences in the opposite direction. Only the summation of all changes is recorded and can be used to infer about synchronously activated underlying cortical neurons. The advantage of the method is that synchronous activity from most neurons underlying the scalp electrode is recorded with the EEG. A further advantage is the rapid detection of electrical differences in voltages resulting in high accuracy when looking at the temporal dimension of the EEG. Therefore, an EEG can be used to determine the precise time course of neuronal processing. Other methods looking at oxygen levels in the blood have a much less temporal accuracy compared to EEG, but a higher spatial resolution compared with an EEG recording. But still it is possible to relate ERPs to specific anatomical areas, sometimes the combination of different recording methods is used to improve accuracy (Eggermont & Ponton, 2002).
The next paragraphs are giving only a short overview on some ERP components that are important for this thesis. When describing the ERPs the main focus is on the MMN, because it is most relevant for this thesis.

5.1 P1
The P1 or P50 is categorized as a “late” component and peaks 50ms after stimulus onset. It is the first of the “late” components and resembles auditory inhibition. After detection of an auditory stimulus the P1 is elicited. If the stimulus is followed by another but redundant stimulus the amplitude of the ERP is reduced. This reduction seems to resemble a sensory gating mechanism. The stimuli passing the thalamus before reaching the cerebral cortex are controlled and analyzed (McCormick & Bal, 1994).

5.2 N1
The N1 or N100 resembles any auditory stimulus onset. The amplitude is modulated by attention. When participants are paying less attention to auditory stimuli the amplitude of the N1 component are significantly lower.

5.3 P2
The P2 often follows an N1 component at about 150 – 275ms after stimulus onset (Dunn, Dunn, Languis and Andrews, 1998). It is a prominent positive deflection of the cranial vertex, being sensitive to physical features of auditory stimuli (Novak, Ritter and Vaughan, 1992).

5.4 The Mismatch Negativity (MMN)
The Mismatch Negativity (MMN) is a component of the event-related potential. It usually peaks at 150ms to 250ms. And is usually accompanied by a polarity reversion at the mastoids in nose-referenced mastoid recordings (Decan, Gomes, Nousak, Ritter and Javitt, 2000). This component has consistently been observed when an infrequently presented sound, referred to as deviant, occurs in a sequence of repetitions of a different auditory signal, referred to as standard, even if participants attend to another (e.g., visual) source of stimulation and report no awareness for the occurrence of the deviant. The response is elicited by any discriminable change in auditory stimulation, so elicitation by changing room acoustics seems a reasonable hypothesis. This hypothesis still needs to be tested.
Different auditory dimensions have been shown to elicit MMN. So far, the four first-order auditory regularity violation dimensions that have been found to elicit a MMN are pitch (Jacobsen & Schröger, 2001; Näätänen et al., 1878; Sams et al., 1985), duration (Jacobsen & Schröger, 2003; Kaukoranta, et al., 1989; Näätänen, et al., 1989), loudness (Jacobsen, et al., 2003; Näätänen, et al., 1987) and sound source location (Kaiser, et al., 2000; Paavilainen, et al., 1989; Schröger & Wolff, 1996). Higher-order auditory regularity violation dimensions eliciting a MMN have also been reported, for instance the omission of a tone in a recurrent pattern (Nordby, et al., 1994; Tervaniemi, et al., 1994), or by speech stimuli violating abstract phonological rules followed by a sequence of standard stimuli (Näätänen, et al., 2007).

The MMN gives evidence for the presence of automatic intelligent processes at the level of auditory cortex. Näätänen et al. (2007) state that the MMN is generated by the brain’s automatic response to basically any change in auditory stimulation exceeding the behavioural discrimination threshold. The MMN is widely assumed to reflect the detection of a violation of an implicit prediction of impending auditory events, based on previously experienced regularity. Näätänen et al. (2007) add to this: “The MMN is rather composed of, at least mainly, the outcome of a discrimination process where the deviant event is found to be discongruent with the memory representation of the preceding stimuli (even in the absence of attention).” (p. 2546)

Given the particular sensitivity of the auditory system for perceiving distant events from all possible directions, such a monitoring mechanism seems particularly well-suited to identify unpredicted, potentially important changes in the environment at an early time.

5.5 P3
The P3 is an ERP component usually peaking at about 300 to 500 ms after stimulus onset in an active oddball paradigm. The target stimulus must be task-relevant to evoke a two peaked P3. The first peak is called P3a and resembles task irrelevant stimuli, resembling a neural correlate of an orienting response (Soltani & Knight, 2000). The second peak the P3b resembles task-relevant stimuli that are actively paying attention to or involuntarily turned attention to.

5.6 Late Discriminative Negativity (LDN)
Korpilahti, Lang, and Aaltonen (1995) described a late-latency waveform following an MMN at about 400 ms after stimulus onset, arguing that it reflects the automatic processing of
complex auditory stimuli. This late-latency Mismatch Negativity, sometimes called a late differentiating negativity (LDN). Most LDN research has focused on speech stimuli (Kuuluvainen, Alku, Makkonen, Lipsanen and Kujala, 2016), but there have also been LDN findings in non-speech studies (for an overview of the LDN research, see Cheour, Korpilahti, Martynova and Lang (2001)). Compared to the MMN the LDN is more strongly connected to age. Amplitude differs with age, but with no clear relationship. Furthermore, some LDN research does not find an LDN for some age groups while for other age groups an LDN is elicited. The findings regarding LDN responses are less robust compared to those for MMN responses.

6 Research Question and Preview

This thesis investigates whether the human auditory system monitors the acoustic environment regarding room acoustical changes pre-attentively. Given the sophisticated ability in detecting acoustic irregularities, on the one hand, and the perceptual-behavioural-emotional importance of room acoustics on the other, it seems likely that room acoustical-based sound changes are detected pre-attentively (i.e., in the absence of a corresponding goal and possibly without awareness). Humans are not necessarily aware of these changes, which renders measuring brain waves a good means of observing reaction to changing room acoustics (Blesser & Salter, 2007).

The first experiment extends previous findings by observing that auditory irregularities brought about by a change in room acoustics elicit a MMN in a passive oddball protocol (acoustic stimuli with differing room acoustics, that were otherwise identical, were employed as standard and deviant stimuli), in which participants watched a fiction movie (silent with subtitles). While most participants reported no awareness for any changes in the auditory stimulation. Given the novelty of these findings, however, corroboration of MMN elicitation by room acoustics-related changes seems warranted. Moreover, some peculiarities of the stimuli used in the first experiment deserve consideration. Of most importance, due to the manipulation of lateralized reflection properties, standards and deviants in this first experiment were associated with an asymmetry of sound intensity levels regarding the left and right ear. Given previous findings of MMN elicitation by inter-aural intensity difference (Paavilainen, et al., 1989), it cannot be completely dismissed that the MMN observed in the first experiment was driven by this asymmetry. Moreover, assuming difficulty of detecting
room-acoustics-related changes with short, homogeneous stimuli, the first experiment repeatedly presented sequences made of three different sounds (i.e., piano chords). An analysis of the time interval subsequent to the MMN (180 ms after the onset of the first chord) revealed a second negative deflection in the ERP, about 260 ms after the onset of the second chord of the tone triplet (visual inspection revealed a third negative deflection, which was not analysed further, however). Although it appears straightforward to assume multiple MMNs, each reflecting the detection of a room acoustics-related change for a particular chord, it is also conceivable that the two effects reflect qualitatively different processes. Specifically, because alterations of room acoustics are characterized by a complex pattern of changes, affecting various parameters of auditory stimuli, some minimal period of time may be necessary to evoke a mismatch response related to this pattern. Viewed from this perspective, it might be conjectured that only the second negative deflection indicates such a detection process whereas the first MMN might be elicited by low-level constituents of the pattern, that is, by one or more featural deviation which, in isolation, would not be indicative for a change in room acoustics. The inter-aural intensity difference in the first experiment is an obvious candidate for such a feature.

Based on these considerations, the second and third experiment set out to provide additional evidence for the notion of pre-attentive detection of unattended changes in room acoustics, controlling for differences of overall inter-aural intensity. To this end, left-right-symmetrical changes in room acoustics were simulated. Furthermore, we aimed to shed light on the role of multiple tone sequences by contrasting the presentation of three identical short sinusoidal tones (Experiment 2) and an uninterrupted single tone, extended to the same overall duration (Experiment 3). Whereas the notion of multiple MMNs, elicited by the onsets of constituent tones of the deviant sequence predicts the occurrence of additional negative deflections only in the three-tone-sequence condition, assuming consecutive detection of featural and room acoustic-specific pattern of deviance predicts a second negative deflection also in the single-tone condition.
Part II

Empirical Section:

Experiment 1: Automatic detection of unattended changes in room acoustics

Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

Experiment 3: Replication of Experiment 2 with single tone stimuli
7 Experiment 1: Automatic detection of unattended changes in room acoustics

7.1 Material and methods

7.1.1 Participants

Fourteen volunteers participated in the experiment (five male, mean age 25.5 years, range 19-60, one left handed). Handedness was assessed using an inventory adopted from Oldfield (1971). All participants were native German speakers, reported normal auditory and normal visual acuity and no neurological, psychiatric, or other medical problems. The experiment was carried out in line with ethical guidelines, in particular The Code of Ethics of the World Medical Association (Declaration of Helsinki) (2008). Informed written consent was obtained from all participants prior to the experimental session.

7.1.2 Materials

By playing a sequence of five piano chords in two variations of a simulated room, two stimuli were constructed. A room acoustics software program (Odeon 11.00 Combined Demo Version; Odeon A/S, Kgs. Lyngby, Denmark) was used for auralisation. Odeon was developed for simulating the interior acoustics of buildings and uses the image-source method combined with ray tracing. An ambisonic recorded stimulus (derived from “Piano Over the rainbow Mic2 SHORT.wav” from the Odeon package) consisting of a sequence of five piano chords was used: F7 (2793 Hz), E7 (2637 Hz), D7 (2349 Hz), C7 (2093 Hz), & G6 (1567 Hz; base frequencies given in the parentheses). To avoid difficulties in perceiving room acoustics based on a single tone, we chose a stimulus of considerable complexity and duration. The chord sequence had an overall duration of 1040 ms, including 5 ms rise and 5 ms fall times. The onset times of the chords were at approximately 5 ms, 350 ms, 510 ms, 660 ms, and 837 ms after stimulus onset, with no silent periods between consecutive chords. For the auralisation a virtual room (“example.Par” from the Odeon package) was used. The simulated room’s acoustic properties were altered to generate two auditory stimuli with different room acoustics but otherwise retained identical properties. The sound source was centred in front of the perceiver (point source; (x,y,z) = (1, 0, 2); see Figure 1 A), and the connection between the sound source and the receiver formed an imagined line dividing the room into two
symmetrical parts. The perceiver was seated in a central position with respect to the right and left walls (single point response receiver; \((x,y,z) = (20, 0, 5)\); see Figure 1A). The surface area of the room was 1268.23m², room temperature 20°C, relative humidity 50%. The reflective properties of walls were altered to generate two stimuli with different room acoustics (Figure 1A). For one stimulus ("right") a 90% absorbing material (equally absorbing all frequencies) was applied to the walls to the right of the receiver and for all other walls a 10% absorbing material (equally absorbing all frequencies) was used. This room setup produced the impression that the room was open to the right. For the second stimulus ("left") a 90% absorbing material was applied to the walls to the left of the receiver and all other walls were covered with 10% absorbing material, giving the impression that the room was open to the left, creating a fully symmetric counterpart.\(^1\) As a consequence, the total Root Mean Square (RMS) of "left" and "right" was the same. Remaining intensity differences due to stochastic aspects of the re-synthesis procedure were equalized using Adobe Audition CS5.5 Demo Version (Adobe Systems GmbH, München, Germany); mean RMS amplitude for "right" (left channel: -18.33dB; right channel: -21.91dB) and “left” (left channel: -22.07dB; right channel: -18.31dB). Different mean RMS amplitudes between the channels were essential in order to maintain the different acoustic properties of the two rooms (see fig. 1 C). As a consequence, there are intensity differences between the two channels. The frequency spectrum (see fig. 1 B) reveals that channel right of stimulus “left” is not identical as channel left of stimulus “right”. Software simulating acoustics cannot make perfect calculations therefore each auralisation does differ slightly. The acoustic stimuli can be found at: http://www.hsu-hh.de/allgpsychologie/index_Ld3q1e6qG8cZO048.html.

\(^{1}\) The following Odeon configuration was used: 1. Room setup: Impulse Response Length 16000 ms, Number of late rays 20000, Max. reflection order 2000, Impulse response resolution 3.0 ms, Transition Order 2, Number of early scatter rays 100, Angular absorption “Soft materials only”, Surface scattering “Actual”, Oblique Lambert, Reflection based scatter “enabled”, Key diffraction frequency 707 Hz, Interior margin 0.10 m, Scatter coefficients > 0.50 handled as uniform scatter. 2. Auralisation setup: Apply dither and noise shaping, Wave result file 16 bit PCM, Create binaural impulse response file, HRTF “Subject_021Res10deg_M3_0_SR414100_Apass0,50_Astop40,00_BOvrLap100%_PPrHRTF256", Headphone “Sennheiser HD250 Linear", DC filter, Overall Recording level 40 dB, Phase approximation “phase shift at surfaces /filter phase, A(stop) 40,00 dB, A(pass) 0,50 dB, Band overlap 100%, Sample rate 44100 Hz, Encoding “1. Order ambisonics”
Deviant stimuli could be actively discriminated from standard stimuli with high accuracy. To test participants’ awareness of the deviant stimuli in the passive oddball protocol the fourteen participants of the EEG experiment were interviewed about their subjective impression regarding the auditory stimuli (approximately 5 minutes after completing the EEG experiment). Ten participants reported that they did not notice any changes to the auditory stimuli, two participants said they felt that the rhythm was sometimes different, one participant reported differences in the sound’s source location and only one participant had the feeling that something with the room’s acoustics changed, but could not specify this observation further.

---

2 To assess discriminability of the two stimuli, another group of ten participants (six male, mean age 28.9 years, range 20-50) who reported normal auditory acuity were asked to detect the one deviant stimulus in a sequence of 10 stimuli. The “right” stimulus was interspersed in a sequence of nine presentations of the “left” stimuli and vice versa. All ten participants completed 20 of these sequences with one deviant at a random position in each of the sequences (1800 standard; 200 deviant). In total, 185 deviants (92.5%) were detected with 4 false alarms (0.22%).
Figure 1:
A: The room on the left was used for the auralisation of Stimulus Left and the room on the right was used for the auralisation of Stimulus Right. The gray shaded walls represent the walls with 90% absorbing material; all other walls represent walls with 10% absorbing material. The receiver symbol represents a single point response receiver and the sound source symbol represents a point source. (B) Frequency spectrum of the auditory stimuli, right channel on top and left channel on the bottom. Stimulus Left on the left and Stimulus Right on the right. The five vertical lines indicating the onset of each piano chord. (C) Results of the intensity measurements with the artificial head HMS III.2. The values on the left are the values from Stimulus Left, with the right channel on top and the left channel on the bottom. The values on the right are the values from Stimulus Right, with the right channel on top and the left channel on the bottom. Values are measured in dB (SPL) and dB (A).
7.1.3 Experimental design and procedure

The participants were seated in an electrically and acoustically shielded experimental chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). 2000 acoustic stimuli were presented binaurally at approximately 52dB SPL\(^3\) (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany) via headphones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany) in two blocks (~19min each with a five minute break between the blocks). The participants were instructed to ignore the acoustic stimuli and concentrate on a movie with German subtitles. All participants reported that they could ignore the acoustic stimuli easily while watching the movie. The movie was shown on a standard 24” 16:10 LCD computer screen at an approximate distance of 130 cm. The participants performed no additional tasks and did not respond to the acoustic stimuli in any way.

There were two types of blocks. In block 1 the “right” stimulus was used as standard (90% of the stimuli; \(n = 900\)) and the “left” was used as deviant (10% of the stimuli; \(n = 100\)). In block 2 the “left” was used as standard and the “right” was used as deviant. The stimuli were presented in a passive oddball protocol in a pseudo-random fashion with the constraints of five consecutive standard stimuli in the beginning of a block, and no consecutive deviant stimuli. The sequence of the two blocks was counterbalanced across participants.

7.1.4 Electrophysiological recordings

The Electroencephalogram (EEG) (Ag/AgCl electrodes, Falk Minow Services, V-Amp EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from nine standard scalp locations according to the extended 10–20 system (American Encephalographic Society, 1994); F3, Fz, F4, C3, Cz, C4, P3, Pz, P4) and the left and right mastoids. Standard scalp electrodes were embedded in a prefabricated cap (Brain Products GmbH, Gilching, Germany). For the recording, a sampling rate of 500 Hz (resolution 16bit) was used. The ground electrode was placed at FCz and the reference on the nose. Electro-ocularg activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) from the left eye (for one participant the right eye was used) by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer

\(^3\) Configuration: equalisation (LIN), synchronisation (44.1 kHz), tool (SPL), without torso
canthi of both eyes. Impedances were kept below 5 kΩ. On-line filtering was carried out using a 0.011-Hz high-pass, a 100-Hz low pass and 50-Hz notch filter.

7.1.5 Data analysis

Off-line signal processing was carried out (EEProbe 3.0; Advanced Neuro Technology, Enschede, Netherlands) on a Linux computer. EEG-data was band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 1140 ms, time-locked to the onset of the stimuli, including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each condition (standard and deviant) and for each participant. The ERP responses to the first five stimuli of each block as well as to the first standard stimulus after each deviant were not included in the analyses. Epochs showing an amplitude change exceeding 120μV at any of the recording channels were rejected. Grand-averages were subsequently computed from the individual-subject averages. Thereby, deviant and standard grand-averages were aggregated using both the left and right stimuli, such that waveforms resulted from physically identical stimulation. Visual inspection of the ERPs derived separately for the two block types revealed no apparent substantial differences in the data patterns.

To quantify the full MMN amplitude, the scalp-recorded ERPs were re-referenced to the averaged signal recorded from the electrodes positioned over the left and right mastoids. This computation results in an integrated measure of the total neural activity underlying the auditory MMN (Schröger, 1998). Deviant-minus-standard difference waveforms were calculated by subtracting point by point the standard ERPs from the associated deviant ERPs. The deviance-related effects for the room acoustics change were quantified with a mean value for a fixed window of 40 ms at the center of the highest peak amplitude in the grand-average difference waves.

The MMN responses were analyzed by comparing the mean ERP amplitudes from the standard with those from the deviants using a three-way repeated-measures analysis of variance (ANOVA) with the factors Stimulus (standard, deviant), Position (F-, C-, P-line) and Laterality (3-, z-, 4-line). The first two negative deflections of the deviant-minus-standard difference waveform revealed by visual inspection were analyzed\textsuperscript{4}. Additionally, a two-way repeated-measure ANOVA with the factors Stimulus (standard, deviant) and Laterality

\textsuperscript{4} Visual inspection revealed a third negative deflection of the deviant-minus-standard difference waveform at about 900 ms after stimulus onset, which was not further evaluated, however.
Experiment 1: Automatic detection of unattended changes in room acoustics

(right-, left-mastoid) was carried out on non-re-referenced data to check for the MMN-typical polarity inversion at the mastoids.

Reduced degrees of freedom (Greenhouse-Geisser) were used where applicable to prevent violating the sphericity assumption underlying ANOVA with repeated measures. Uncorrected degrees of freedom are reported.
Figure 2:
Deviant and standard and difference ERP waveforms. Grand-averaged ERPs averaged separately for the deviant (dotted lines) and standard stimuli (dashed lines). Deviant-minus-standard difference waveforms are also presented (solid lines). The gray bars mark the statistically analyzed time windows of the first effect (160-200 ms; gray bar) and the second effect (586-626 ms; gray bar with a black border). Scales are in s and μV.
7.2 Results

On average, 7.66% of the epochs per participant were rejected prior to ERP computation (7.00% of the standard epochs and 7.94% of the deviant epochs; range 0.45% to 40.17%) due to amplitude changes exceeding 120μV during an epoch. Two consecutive effects were observed. For the first effect, peak latencies occurred at 180 ms after stimulus onset with a maximum peak amplitude at electrode F4 of -0.778μV (mean amplitude at Fz -0.745μV) in the deviant-minus-standard waveform. For the second effect, a peak latency at 606 ms with a peak maximum of -1.431μV at electrode Fz (mean amplitude at Fz -1.431μV) was observed. Correspondingly, the time windows for the ERP quantization were set from 160 to 200 ms for the first effect and from 586 to 626 ms for the second effect.

The three-way repeated-measures ANOVAs with data for the re-referenced ERPs revealed significant main effects for Stimulus in the first time window, $F_{1,13}=12.6$, $p=0.004$, and the second time window, $F_{1,13}=13.4$, $p=0.003$. These main effects resulted from more negative-going ERP amplitudes for deviants than for standards. The same ANOVAs revealed significant effects for the interaction of Stimulus and Position for the first time window, $F_{2,12}=10.4$, $p=0.002$, and the second time window, $F_{2,12}=10.4$, $p=0.002$, demonstrating that the effects were larger at more anterior electrode sites.

The two-way repeated-measures ANOVAs on the non-re-referenced data of the mastoid electrodes revealed a significant main effect of Stimulus for the second time window, $F_{1,13}=13.0$, $p=0.003$ (mean amplitude at the left mastoid 0.708μV; mean amplitude at the right mastoid 0.826μV) and a marginally significant effect for the first time window ($F_{1,13}=4.0$, $p=0.068$) (mean amplitude at the left mastoid 0.367μV; mean amplitude at the right mastoid 0.484μV), resulting from a mastoidal polarity inversion.
8 Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

8.1 Material and methods

8.1.1 Participants
Sixteen volunteers participated (three female, mean age 24.6 years, age range 22–28, one left-handed). Handedness was assessed using an inventory adopted from Oldfield (1971).

All participants reported normal auditory and normal visual acuity and no neurological, psychiatric, or other medical conditions. The experiment was carried out in line with ethical guidelines, specifically, The Code of Ethics of the World Medical Association Declaration of Helsinki (2008). Informed written consent was obtained from all participants prior to the experimental session.

8.1.2 Materials
For the Experiment a 500-ms-long sequence of three sinusoidal tones was generated. It consisted of 100 ms of tone, followed by 100 ms of silence, 100 ms of tone, 100 ms of silence, and 100 ms of tone. In addition, a constant interstimulus interval (ISI) of 100 ms was used between sequences (triplets).

The tones were then modified. First, the sinusoidal tones were converted into square waves in order to add spectral complexity. This addition of harmonics made the sounds more realistic than pure sinusoidal tones, thus adding external validity. Then, to generate different room acoustics, room simulations were used to reproduce the acoustics of a room. To this end, calculation of the stimulus parameters was based on a model of a fully symmetrical (i.e., spherical) room with no room modes. The direct audio signals without room impressions were mapped to certain room acoustics. Specifically, the acoustical room properties of the

---

5An idealized room was used that cannot be found in reality, but rooms resembling this room closely can be built. This perfect room without room modes was used to eliminate confounds due to resonances depending on the room and not the stimuli. This idealized room reflects all frequencies equally strong producing a linear spectrum without any peaks, resulting in an environment with reverberation with no resonances. The reflection properties of the simulated room resemble those of a diffuse environment like, for instance, a forest where the reflection from the trees is so diffuse that no frequencies are reflected stronger than others.
room to be simulated were transformed into a room impulse response (Zölzer, 2008). A fast convolution algorithm was used to generate the stimuli; the sounds presented to the participants were computed as the sum of the convolution (modeling the reverberation) and the original sound (square waves). The calculation (Matlab code) is provided in Appendix D.

For the experiment, four stimuli differing in reverberation times and convolution, corresponding to rooms of different sizes, were derived. The rooms were Alpha 1 (largest room), Alpha 3 (second largest room), Alpha 7 (third largest room), and Alpha 10 (smallest room). As a consequence of the manipulation between room sizes, the first two sound segments of the sequences (triplets) slightly differed in duration from the third. Because this difference was constant we considered it unlikely that it would affect the automatic detection of changes in room acoustics. At the end of the stimulus generation, a 5 ms fade in and fade out was added to all sounds in the sequences.

Finally, the total Root Mean Square (RMS) amplitude was equalized for all four stimuli. Time domain representations of the acoustic stimuli are presented in Appendix F. The stimuli can be found at: http://www.hsu-hh.de/allgpsychologie/index_Ld3q1e6qG8cZO048.html.

To test participants’ awareness of the deviant stimuli in the passive oddball protocol, the participants were interviewed post-experimentally about their subjective impressions regarding the acoustic stimuli. Thirteen participants reported not noticing changes in the acoustic stimuli. Three participants reported noticing differences in frequency, volume, duration, and other features. No participant reported noticing changes in perceived sound source location or differences in left-right symmetry. After being informed that changes had occurred in the acoustic stimuli and being asked to describe these changes, seven participants reported various differences (e.g., an echo, differing frequency, faster sound sequences, different sounds from time to time, and lower volume of the stimuli over the course of the experiment). Nine participants insisted that the acoustic stimuli remained the same throughout the entire experiment.

8.1.3 Experimental design and procedure
The participants were seated in an electrically and acoustically shielded experimental chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). Acoustic stimuli were presented binaurally at approximately 43.6 dB SPL (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany) via headphones (Sennheiser HD 25-1 II;
8 Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). The participants were instructed to ignore the acoustic stimuli and to concentrate on a movie that was presented with subtitles. All participants reported that they were able to ignore the acoustic stimuli while watching the movie. The movie was shown on a standard 24” 16:10 LCD computer screen at an approximate distance of 130 cm from the participants. The participants performed no additional tasks and were asked not to respond to the acoustic stimuli in any way.

In each experimental block, 900 standard and 100 deviant stimuli were presented. The blocks differed with respect to the specific standard-deviant combination that was administered. A larger contrast in room acoustics was created in Blocks 1 and 2, and a smaller one in Blocks 3 and 4. We refer to the former as high-contrast condition and to the latter as low-contrast condition. In Block 1, the Alpha 10 stimulus was used as the standard and the Alpha 1 stimulus was used as the deviant. In Block 2, the roles of standard and deviant were reversed, i.e., Alpha 1 was the standard and Alpha 10 was the deviant. In Block 3, Alpha 7 was used as the standard and Alpha 3 was used as the deviant. In Block 4, the roles were again reversed, with Alpha 3 as the standard and Alpha 7 as the deviant. In these four blocks, the stimuli were presented in a passive oddball protocol in a pseudo-random fashion with the constraints that two deviants could not be presented in direct succession and five consecutive standard stimuli would be presented at the beginning of each block. In four additional blocks, the four stimuli were each presented in isolation 220 times (Alpha 1 in Block 5, Alpha 3 in Block 6, Alpha 7 in Block 7, and Alpha 8 in Block 8). Block 5 to 8 were used to control for potential context effects of oddball conditions on the ERP to the standard (Jacobsen et al., 2004; Jacobsen, Schröger, Winkler and Horvath, 2005). As no such effects were observed, Block 5 to 8 were not considered further. The sequence of presentation of all eight blocks was counterbalanced across participants.

8.1.4 Electrophysiological recordings
An electroencephalogram (EEG) (Ag/AgCl electrodes, Falk Minow Services, Brain Amp, EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from twenty-six standard scalp locations according to the extended 10–20 system6. Standard scalp electrodes were embedded in a prefabricated cap (Brain Products GmbH, Gilching, 6Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FCz, FC2, FC6, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2.)
Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

Germany). For the recording, a sampling rate of 500 Hz (resolution: 16 bit) was used. The ground electrode was placed at FCz and the reference on the nose. Electrooculographic activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) from the left eye by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 5 kΩ. On-line filtering was carried out using a 0.011-Hz high-pass filter and a 100-Hz low-pass filter.

8.1.5 Data analysis

Off-line signal processing (EEProbe 3.0; Advanced Neuro Technology, Enschede, Netherlands) was carried out on a Linux computer. EEG data were band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 600 ms, time-locked to the onset of the stimuli, including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each condition (standard and deviant) and for each participant. The ERP responses to the first five stimuli of each block as well as to the first standard stimulus after each deviant were not included in the analyses. Epochs which showed an amplitude change exceeding 120 μV at any of the recorded channels were rejected. Visual inspection of the ERPs that were computed for the separate stimuli within the high-contrast as well as the low-contrast condition revealed no apparent substantial differences in the data patterns. Therefore, the ERPs to deviants and standards were aggregated over both individual stimuli of the high-contrast condition (Alpha 1 and Alpha 10 from Blocks 1 and 2) and the low-contrast condition (Alpha 3 and Alpha 7). Thus, the ERP waveforms resulted from physically identical stimulations (Jacobsen et al., 2004; Jacobsen et al., 2005). Grand averages were subsequently computed from the individual subject averages.

Deviant-minus-standard difference waveforms were calculated by subtracting, point by point, the standard ERPs from the associated deviant ERPs. The deviance-related ERP effects, elicited by room acoustics changes, were quantified by computing the mean amplitude across a 40 ms window centered around the peaks in the grand-mean difference waves. The MMN responses were analyzed by comparing the mean ERP amplitudes for the standards with those for the deviants through a three-way repeated-measures analysis of variance (ANOVA) with the factors Stimulus (standard, deviant), Position (F-, C-, P-line) and Laterality (3-, z-, 4-line). Three negative deflections of the deviant-minus-standard difference waveform, revealed by visual inspection, were analyzed. Additionally, two-way repeated-
Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

measures ANOVAs with the factors Stimulus (standard, deviant) and Laterality (right-mastoid, left-mastoid) were carried out to check for the MMN-typical polarity inversion at the mastoids.

Only ANOVA statistics directly relevant for our hypotheses are reported in the Results section. A full report of the statistical analyses (including topographical factors) is provided in Appendix F.

Figure 3:
Deviant and standard and difference ERP waveforms. Grand-averaged ERPs averaged separately for the deviant (grey lines) and standard stimuli (black lines) of the high contrast condition in Experiment 2. Deviant-minus-standard difference waveforms are also presented (solid small lines). The grey bars mark the statistically analyzed time windows of the first effect (208–248 ms; grey bar), the second effect (420–460 ms; dark grey bar with a black border), and the third effect (538–578 ms; light grey bar with a black border). Scales are in s and μV.
Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

8.2 Results

For the high contrast condition (“Alpha 1” and “Alpha 10”), 28.34% of the epochs were rejected prior to ERP computation (28.65% of the standard epochs and 25.86% of the deviant epochs; range 4.94% to 68.41%) due to amplitude changes exceeding 120μV during an epoch. Three consecutive effects were observed for the high contrast condition (Figure 3).

At electrode site Fz the peak latencies of the three effects were 228 ms, 440 ms, and 558 ms, with mean peak amplitude differences (deviant – standard) of -2.719μV, -2.407μV, and -2.555μV, respectively. Significant main effects of Stimulus were obtained in all three ANOVAs, $F_{1,15} = 51.84, p<0.000$; $F_{1,15} = 45.62, p<0.000$; and $F_{1,15} = 45.53, p<0.000$; for the first, second, and third time window, respectively.

Also the two-way repeated-measures ANOVAs on the non-re-referenced data of the mastoid electrodes revealed significant main effects of Stimulus, resulting from mastoidal polarity inversion, $F_{1,15} = 38.23, p<0.001$ (mean amplitude difference at the left mastoid 1.482μV; mean amplitude at the right mastoid 1.550μV), $F_{1,15} = 72.20, p<0.001$ (mean amplitude at the left mastoid 1.170μV; mean amplitude difference at the right mastoid 1.015μV), and $F_{1,15} = 107.23, p<0.001$ (mean amplitude difference at the left mastoid 1.360μV; mean amplitude at the right mastoid 1.313μV), for the first, the second, and the third time window, respectively.

For the low contrast condition (“Alpha 3” and “Alpha 7”) on average, 28.43% of the epochs were rejected prior to ERP computation (28.69% of the standard epochs and 26.36% of the deviant epochs; range 4.09% to 70.62%) due to amplitude changes exceeding 120μV during any epoch. Again, three consecutive effects were observed (Figure 4).
Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

Figure 4:
Deviant and standard and difference ERP waveforms. Grand-averaged ERPs averaged separately for the deviant (grey lines) and standard stimuli (black lines) of the high contrast condition in Experiment 2. Deviant-minus-standard difference waveforms are also presented (solid small lines). The grey bars mark the statistically analyzed time windows of the first effect (160–200 ms; grey bar), the second effect (352–392 ms; dark grey bar with a black border), and the third effect (542–582 ms; light grey bar with a black border). Scales are in s and μV.
8 Experiment 2: Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

At electrode site Fz the peak latencies of the three effects were 180 ms, 372 ms, and 562 ms, with mean peak amplitude differences (deviant – standard) of -3.360μV, -2.867μV, and -3.300μV, respectively. Significant main effects of Stimulus were obtained in all three ANOVAs, $F_{1,15}=62.07, p<0.001$, $F_{1,15}=30.22, p<0.001$; and $F_{1,15}=42.52, p<0.001$; for the first, second, and third time window, respectively.

Again, the two-way repeated-measures ANOVAs on the non-re-referenced data of the mastoid electrodes revealed significant main effects of Stimulus, resulting from mastoidal polarity inversions, $F_{1,15}=30.07, p<0.001$ (mean amplitude difference at the left mastoid 1.338μV; mean amplitude difference at the right mastoid 1.327μV), $F_{1,15}=25.89, p<0.001$ (mean amplitude difference at the left mastoid 0.814μV; mean amplitude difference at the right mastoid 1.146μV), and $F_{1,15}=70.05, p<0.001$ (mean amplitude difference at the left mastoid 1.157μV; mean amplitude difference at the right mastoid 1.179μV), for the first, the second, and the third time window, respectively.
9 Experiment 3: Replication of Experiment 2 with single tone stimuli

9.1 Material and methods

9.1.1 Participants
Sixteen volunteers participated (nine female, mean age 23.4 years, age range 21–26, one left-handed). Handedness was assessed using an inventory adopted from Oldfield (1971).

All participants reported normal auditory and normal visual acuity and no neurological, psychiatric, or other medical problems. The experiment was carried out in line with ethical guidelines, specifically, The Code of Ethics of the World Medical Association Declaration of Helsinki (2008). Informed written consent was obtained from all participants prior to the experimental session.

9.1.2 Materials
For the third Experiment, a sinusoidal tone of 500 ms duration was generated, matching the tone triplet of Experiment 2 in overall duration. A constant ISI of 100 ms like in Experiment 2 was also used in Experiment 3.

The tones were then modified. First, the sinusoidal tones were converted into square waves in order to add spectral complexity. This addition of harmonics made the sounds more realistic than pure sinusoidal tones, thus adding external validity. Then, to generate different room acoustics, room simulations were used to reproduce the acoustics of a room. To this end, calculation of the stimulus parameters was based on a model of a fully symmetrical (i.e., spherical) room with no room modes\(^7\). The direct audio signals without room impressions were mapped to certain room acoustics. Specifically, the acoustical room properties of the room to be simulated were transformed into a room impulse response Zölzer (2008). A fast convolution algorithm was used to generate the stimuli; The sounds presented to the

---

\(^7\)An idealized room was used that cannot be found in reality, but rooms resembling this room closely can be built. This perfect room without room modes was used to eliminate confounds due to resonances depending on the room and not the stimuli. This idealized room reflects all frequencies equally strong producing a linear spectrum without any peaks, resulting in an environment with reverberation with no resonances. The reflection properties of the simulated room resemble those of a diffuse environment like, for instance, a forest where the reflection from the trees is so diffuse that no frequencies are reflected stronger than others.
participants were computed as the sum of the convolution (modeling the reverberation) and the original sound (square waves). The calculations are the same as in Experiment 2 and are provided in Appendix D (Matlab code).

For each experiment, four stimuli differing in reverberation times and convolution, corresponding to rooms of different sizes, were derived. The rooms were Alpha 1 (largest room), Alpha 3 (second largest room), Alpha 7 (third largest room), and Alpha 10 (smallest room). In the third experiment, there are no slight differences in duration between the stimuli like in Experiment 2, because the sequence consisted only from one sound. At the end of the stimulus generation, a 5 ms fade in and fade out was added to all sounds in the sequences. Finally, the total Root Mean Square (RMS) amplitude was equalized for all four stimuli. Time domain representations of the acoustic stimuli are presented in Appendix F. The stimuli can be found at: http://www.hsu-hh.de/allpsychologie/index_Ld3q1e6qG8cZO048.html.

To test participants’ awareness of the deviant stimuli in the passive oddball protocol, the participants were interviewed post-experimentally about their subjective impressions regarding the acoustic stimuli. Twelve participants reported not noticing changes in the acoustic stimuli. Among the four participants who reported noticing changes, only one reported noticing changes in room acoustics. No participant reported noticing changes in the perceived sound source location or differences in left-right symmetry.

9.1.3 Experimental design and procedure

The participants were seated in an electrically and acoustically shielded experimental chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). Acoustic stimuli were presented binaurally at approximately 43.6 dB SPL (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany) via headphones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). The participants were instructed to ignore the acoustic stimuli and to concentrate on a movie that was presented with subtitles. All participants reported that they were able to ignore the acoustic stimuli while watching the movie. The movie was shown on a standard 24” 16:10 LCD computer screen at an approximate distance of 130 cm from the participants. The participants performed no additional tasks and were asked not to respond to the acoustic stimuli in any way.

In each experimental block, 900 standard and 100 deviant stimuli were presented. The blocks differed with respect to the specific standard-deviant combination that was administered. A larger contrast in room acoustics was created in Blocks 1 and 2, and a smaller
one in Blocks 3 and 4. We refer to the former as high-contrast condition and to the latter as low-contrast condition. In Block 1, the Alpha 10 stimulus was used as the standard and the Alpha 1 stimulus was used as the deviant. In Block 2, the roles of standard and deviant were reversed, i.e., Alpha 1 was the standard and Alpha 10 was the deviant. In Block 3, Alpha 7 was used as the standard and Alpha 3 was used as the deviant. In Block 4, the roles were again reversed, with Alpha 3 as the standard and Alpha 7 as the deviant. In these four blocks, the stimuli were presented in a passive oddball protocol in a pseudo-random fashion with the constraints that two deviants could not be presented in direct succession and five consecutive standard stimuli would be presented at the beginning of each block. In four additional blocks, the four stimuli were each presented in isolation 220 times (Alpha 1 in Block 5, Alpha 3 in Block 6, Alpha 7 in Block 7, and Alpha 8 in Block 8). Block 5 to 8 were used to control for potential context effects of oddball conditions on the ERP to the standard (Jacobsen et al., 2004; Jacobsen et al, 2005). As no such effects were observed, Block 5 to 8 were not considered further. The sequence of presentation of all eight blocks was counterbalanced across participants.

9.1.4 Electrophysiological recordings

An electroencephalogram (EEG) (Ag/AgCl electrodes, Falk Minow Services, Brain Amp, EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from twenty-six standard scalp locations according to the extended 10–20 system. Standard scalp electrodes were embedded in a prefabricated cap (Brain Products GmbH, Gilching, Germany). For the recording, a sampling rate of 500 Hz (resolution: 16 bit) was used. The ground electrode was placed at FCz and the reference on the nose. Electrooculographic activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) from the left eye by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 5 kΩ. On-line filtering was carried out using a 0.011-Hz high-pass filter and a 100-Hz low-pass filter.

---

8Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FCz, FC2, FC6, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2.
9.1.5 Data analysis

Off-line signal processing (EEProbe 3.0; Advanced Neuro Technology, Enschede, Netherlands) was carried out on a Linux computer. EEG data were band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 600 ms, time-locked to the onset of the stimuli, including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each condition (standard and deviant) and for each participant. The ERP responses to the first five stimuli of each block as well as to the first standard stimulus after each deviant were not included in the analyses. Epochs which showed an amplitude change exceeding 80 μV at any of the recorded channels were rejected. Visual inspection of the ERPs that were computed for the separate stimuli within the high-contrast as well as the low-contrast condition revealed no apparent substantial differences in the data patterns. Therefore, the ERPs to deviants and standards were aggregated over both individual stimuli of the high-contrast condition (Alpha 1 and Alpha 10 from Blocks 1 and 2) and the low-contrast condition (Alpha 3 and Alpha 7).

Thus, the ERP waveforms resulted from physically identical stimulations (Jacobsen et al., 2004; Jacobsen et al, 2005). Grand averages were subsequently computed from the individual subject averages.

Deviant-minus-standard difference waveforms were calculated by subtracting, point by point, the standard ERPs from the associated deviant ERPs. The deviance-related ERP effects, elicited by room acoustics changes, were quantified by computing the mean amplitude across a 40 ms window centered around the peaks in the grand-mean difference waves. The MMN responses were analyzed by comparing the mean ERP amplitudes for the standards with those for the deviants through a three-way repeated-measures analysis of variance (ANOVA) with the factors Stimulus (standard, deviant), Position (F-, C-, P-line) and Laterality (3-, z-, 4-line). Two negative deflections of the deviant-minus-standard difference waveform, revealed by visual inspection, were analyzed. Additionally, two-way repeated-measures ANOVAs with the factors Stimulus (standard, deviant) and Laterality (right-mastoid, left-mastoid) were carried out to check for the MMN-typical polarity inversion at the mastoids.

Only ANOVA statistics directly relevant for our hypotheses are reported in the Results section. A full report of the statistical analyses (including topographical factors) is provided in Appendix F.
9.2 Results

For the high-contrast condition (Alpha 1 and Alpha 10), 67.69% of the epochs were rejected prior to ERP computation (67.93% of the standard epochs and 65.75% of the deviant epochs; range 27.5% to 93.95%) due to amplitude changes exceeding 80 μV during an epoch. Two consecutive negative deflections were observed for the high-contrast condition.

Figure 5:
Deviant and standard and difference ERP waveforms. Grand-averaged ERPs averaged separately for the deviant (grey lines) and standard stimuli (black lines) of the high contrast condition in Experiment 3. Deviant-minus-standard difference waveforms are also presented (solid small lines). The grey bars mark the statistically analyzed time windows of the first effect (220–260 ms; grey bar), the second effect (518–558 ms; dark grey bar with a black border). Scales are in s and μV.

Excluding the data from the participant with the highest rejection rate from the analyses did not substantially change the pattern of results.
At electrode site F4, the peak latency of the first negative deflections was 240 ms and for the second negative deflection the peak latency was 538 ms, with mean peak amplitude differences (deviant – standard) of -3.221 µV and -2.607 µV, respectively. The ANOVAs of the ERP data revealed significant main effects for Stimulus in the first time window, $F_{1,15}=39.03$, $p<0.001$, and also in the second time window, $F_{1,15}=15.52$, $p<0.001$.

The ANOVAs of the data from the mastoids revealed a significant main effect of Stimulus in the first time window, resulting from mastoidal polarity inversion, $F_{1,15}=14.83$, $p=0.002$ (mean peak amplitude difference at the left mastoid 0.982 µV, mean peak amplitude difference at the right mastoid 0.968 µV), but not for the second time window, $F_{1,15}=2.46$, $p=0.137$ (mean peak amplitude difference at the left mastoid 0.430 µV, mean peak amplitude difference at the right mastoid 0.321 µV).

For the low-contrast condition (Alpha 3 and Alpha 7), 66.61% of the epochs were rejected prior to ERP computation (66.8% of the standard epochs and 65.04% of the deviant epochs; range 20.45% to 87.12%) due to amplitude changes exceeding 80 µV during an epoch. Again, two consecutive negative deflections were observed.
Figure 6:
Deviant and standard and difference ERP waveforms. Grand-averaged ERPs averaged separately for the deviant (grey lines) and standard stimuli (black lines) of the high contrast condition in Experiment 3. Deviant-minus-standard difference waveforms are also presented (solid small lines). The grey bars mark the statistically analyzed time windows of the first effect (238–278 ms; grey bar), and the second effect (538–578 ms; dark grey bar with a black border). Scales are in s and μV.

At electrode site Fz, the peak latencies of the two negative deflections were 258 ms and 558 ms, with mean peak amplitude differences (deviant – standard) of -2.162 μV and -2.329 μV, respectively. The ANOVAs of the ERP data revealed significant main effects for Stimulus in both time windows, $F_{1,15}=25.68$, $p<0.001$, and $F_{1,15}=21.80$, $p<0.001$ for the first and second time window, respectively.

The ANOVAs of data from the mastoids revealed significant main effects of Stimulus for both time windows, resulting from mastoidal polarity inversion, $F_{1,15}=6.87$, $p=0.019$ (mean peak amplitude difference at the left mastoid 0.422 μV, mean peak amplitude
difference at the right mastoid 0.603 μV), and $F_{1,15}=12.95$, $p=0.003$ (mean peak amplitude difference at the left mastoid 0.567 μV, mean peak amplitude at the right mastoid 1.003 μV), for the first and second time windows, respectively.
Part III

Discussion and Outlook:
10 Discussion and Outlook

10. Discussion

10.1 Discussion Experiment 1
This experiment investigated auditory monitoring outside the focus of attention. Stimuli with differing room acoustics were randomly presented in a passive oddball protocol. Although the majority of participants did not report perceiving any irregularity in the to-be-ignored acoustic stimulation, ERPs elicited by deviants consisting of five consecutive piano chords displayed a negative deflection at 180 ms, associated with the typical characteristics of the MMN. These results are consistent with the assumption of continuous, automatic monitoring regarding changes in room acoustics.

The current study reports the first MMN experiment using changing room acoustics as a violation of auditory regularities, aiming to achieve a high external validity with respect to pre-attentive environmental auditory monitoring. Even with the auditory scene presented via headphones, the experimental situation was modeled based on a scenario where an individual is sitting in a dark living room watching a movie. The stimulus manipulation used in this study resembles the room acoustical change resulting from a large open door, changing the reflection properties of the wall.

Because changes in room acoustics-related aspects are necessarily associated with changes in more basic auditory dimensions, we cannot dismiss the possibility that one (or a specific subset) of the latter is sufficient to trigger a MMN response. Most obviously, in this connection, standard and deviant stimuli in our experiment differed regarding intensity levels at the left and right ear (i.e., interaural level difference, ILD). Manipulation of ILD is often perceived as variation regarding the azimuth of the sound source. Paavilainen et al. (1989) found that irregularity regarding ILD elicited a MMN in a passive oddball protocol (although this effect did not reach statistical significance for intensity differences of 2 and 3 dB per ear, which most closely resembles the manipulation used in the experiment here). Noteworthily, whereas in the study of Paavilainen et al. (1989) participants reported difficulty ignoring the (apparently moving) auditory stimuli, the participants of our experiment did not experience such difficulty and only one out of fourteen reported perceiving changes in sound source location.

In this study, standard and deviant stimuli were generated by manipulating the reflecting properties of the virtual walls on the left and right side. By contrast, a left-right
symmetrical manipulation of room acoustics, such as changing the properties of the wall in front of or behind the perceiver, would prevent ILD and generate left-right symmetrical spectral and, potentially, level differences for standard and deviant stimuli. Such deviation, produced by manipulations other than changing room acoustics, has also been found to elicit a MMN in previous studies. For example, presenting repetitive sequences of the same tone from a loudspeaker aligned with the perceivers’ sagittal midline, Winkler, Tervaniemi, Schröger, Wolff and Näätänen (1998) observed a MMN when the identical tone, approximately matched for overall intensity, was presented simultaneously from a left-sided and a right-sided loudspeaker. In this condition, stimulus intensity was symmetrically reduced in each of the lateral loudspeakers and the tone was perceived to originate from a spacially extended central sound source rather than to come from a different location. This manipulation relates to a change in the size of the apparent sound source while room characteristics remained constant. Given the complexity of room acoustical-related changes in auditory stimulation, future research is needed to pinpoint the functional role of specific components, such as ILD, absolute level difference, or spectral distribution.

An aspect of our results that deserves discussion relates to the occurrence of more than one negative deflection during the course of deviant processing. Specifically, in contrast to only one MMN expected at 100-250 ms after stimulus onset, the violation of auditory regularities in this experiment elicited a first negative deflection after 180 ms and a second negative deflection after 606 ms. Although we did not explicitly expect the second effect, it seems plausible to assume that it constitutes an additional MMN, elicited by the second chord of our tone sequence (Jacobsen, Steinberg, Truckenbrodt and Jacobsen 2013). That is, the salience of regularity violation may vary during the course of the presentation of a multiple chord stimulus, possibly reaching maximum after the onset of the individual chords, turning them into successive perceived deviants. The timing of the second negative deflection (the maximum amplitude of the difference wave occurred 256 ms after the onset of the second chord) as well as the MMN-typical polarity inversion at the mastoids, found for the second time window, are in line with this interpretation. Assuming that the detection of a change in room acoustics necessitates a comparably long period of stimulation it seems an intriguing possibility that the MMN found 180 ms after stimulus onset reflects the detection of low-level acoustic changes whereas the MMN after 606 ms constitutes a reflection of the processing of scene difference.
That said, it cannot be dismissed that the negative deflection observed during the second time window might represent a later ERP component elicited during processing of the first chord rather than a MMN locked to the second chord. A possible candidate for such a component is the Late Discriminative Negativity (LDN, a.k.a. late MMN) (Korpilahti et al., 1995; Ceponiene, Cheour, and Näätänen, 1998).

In summary, the results of the first experiment indicate that auditory irregularities resulting from changing room acoustics elicit an MMN in the absence of a corresponding intention and little awareness of the acoustic stimulation. Inasmuch as such automatic detection provides cues of behavioral relevance, these findings can be considered a first step concerning a thorough understanding of the usage of auditory monitoring for adapting to room acoustics-mediated environmental changes.

10.2 Discussion Experiment 2
The results of the second experiment replicated and extended findings of MMN elicitation by room-acoustics-related changes in unattended stimuli from the first experiment. Deviant stimuli were consistently associated with a pronounced negative deflection at frontal electrode sites, at 228 ms and 180 ms after stimulus onset. To match standards and deviants in terms of their left-right intensity distribution, thus ruling out an account of the results from the first experiment in terms of overall interaural intensity differences. In the second experiment a symmetrical room was used for the simulation of the room acoustical changes. MMNs occurred not only in the high-contrast condition (i.e., Alpha 1 vs. Alpha 10) but also in the subtler low-contrast condition (i.e., Alpha 3 vs. Alpha 7).

Different rooms have differing reflection characteristics and therefore produce differing reverberations. Variations of point-by-point intensity between standards and deviants, each within a single auditory object, may indeed lead to a MMN in oddball protocols if the intensity variation over time is detectable by the auditory system. It seems unlikely, however, that the MMN observed in the current study was driven by the variation of point-by-point intensity over time resulting from the room acoustics manipulation. If this was the case, a larger amplitude in the first negative deflection for the high-contrast condition would be expected compared, than in the first negative deflection for the low-contrast condition, because a larger and thus more salient intensity difference around stimulus onset should affect both the MMN and the N1 (Tervaniemi et al., 1994; Näätänen et al., 2007).
Another potential caveat hinges on the duration differences between the standard and the deviant. These could, in principle, elicit MMNs. As has already been mentioned, the first and second sounds of the sequences (triplets) differ in duration between Alpha 1 and Alpha 10 and between Alpha 3 and Alpha 7. It is, however, highly unlikely that the MMNs observed in Experiment 2 resulted from differences in duration, because the MMN latencies do not fit this assumption. The MMN peak latencies were too early to be compatible with elicitation by duration differences.

Replicating previous findings from Experiment 1, additional negative deflections occurred subsequently to the initial MMN triggered by the deviant stimulus. Contrasting with the first experiment, in which sequences of heterogeneous piano chords were used as stimuli, the second experiment used triplets of sounds as standards and deviants. These modifications did not prevent the occurrence of later negative deflections following the first MMN.

In the second experiment, two additional negative deflections were observed, peaking at 372 ms and 562 ms and at 440 ms and 558 ms in the low-contrast condition and the high-contrast condition, respectively. Each of these negative deflections was accompanied by a significant polarity inversion at the mastoid electrodes, which is typically observed for MMNs. These multiple MMN responses could simply reflect separate detection of deviance for the three sound segments of the triplets. The temporal correspondence of the negative deflections and the onsets of the constituent sounds of the triplets are in line with the assumption that the deflections reflect consecutive instances of mismatch detection, related to each individual sound of the deviant triplet. On the other hand, neither complexity nor “novelty” (i.e., deviance from the directly preceding sound segment) constitute necessary conditions for multiple MMNs to occur. Given that the acoustic characteristics of the second and third sound segment were completely predictable due to the identity relation with the first sound of the triplet, the occurrence of sound-related multiple MMNs suggests that separable expectations were formed for the constituent elements of the repeatedly presented sequence and were not adjusted after detection of the first sound of a deviant sequence.

Korpilahti et al. (1995) described a late-latency waveform following an MMN, arguing that it reflects the automatic processing of complex auditory stimuli. This late-latency Mismatch Negativity, sometimes called a late differentiating negativity (LDN), peaks at about 400 to 450 ms (Korpilahti, Krause, Holopainen and Lang, 2001). Most LDN research has focused on speech stimuli (Kuuluvainen et al., 2016), but there have also been LDN findings in non-speech studies (for an overview of the LDN research, see Cheour et al., 2001).
findings regarding LDN responses are less robust compared to those for MMN responses. Developmentally oriented LDN research has also revealed much greater maturation effects. In our view, taking these findings together, no clear pattern for interpretation of the present findings with respect to the LDN emerges. It remains an open question why two further negative deflections were found in the second experiment, rather than only a single MMN as found in other comparable studies (Vaz Pato, Jones, Perez and Sprague, 2002). Explanations of the LDN representing further processing of the detected change in the acoustic input come close to our hypotheses (Cheour et al., 1998). All in all, the functional significance of the additional negative deflections remains to be determined.

10.3 Discussion Experiment 3
In the third experiment, only one single sound was presented in each sequence. Using only these single sounds and no triplets like in Experiment 2, or a sequence of piano chords like in Experiment 1, while leaving everything else constant makes it possibility to test the hypothesis that all the additional negative deflections in Experiment 1 and 2 are simply due to the multiple sound onsets.

The results of Experiment 3 demonstrate, that the occurrence of a second negative deflection does not depend on the presentation of successive sounds. That is, even though only a single sound was used, an additional negative deflection occurred at frontal and central electrode sites, peaking at 538 ms and 558 ms after stimulus onset in the high-contrast and low-contrast condition, respectively. Again, these negative deflections were accompanied by MMN-typical polarity inversions at the mastoid electrodes (albeit only significant in the low-contrast condition). These findings are in line with the assumption that acoustic irregularities associated with a complex pattern of changes in the stimulus may elicit two successive processes of deviance detection. Potentially, this involves a fast process of detecting one or more low-level featural changes and a slower process of identifying the meaning of the specific irregular pattern of featural changes. In the case of the deviant stimuli made up of successive separated sounds, the occurrence of a sound-onset-related additional MMN may be superimposed on the second detection response.

Also, the potential duration confound due to the successive sounds in Experiment 2 can be ruled out because the confound was removed in Experiment 3, which revealed comparable initial MMNs. Therefore, room acoustics again stand as the elicitor.
As noted previously in the discussion of Experiment 2 the variations of point-by-point intensity between standards and deviants, each within a single auditory object, may indeed lead to an MMN in oddball protocols if the intensity variation over time is detectable by the auditory system. The results of Experiment 3 extend the findings of hypothesis that the MMN is driven by the variation of point-by-point intensity over time. Because a larger amplitude in the first negative deflection would be expected for the high-contrast condition than in the first negative deflection for the low-contrast condition. In the third experiment (in contrast to the second experiment), however, the amplitude of the first negative deflection is higher in the high-contrast condition than in the low-contrast condition. It appears that despite the more salient difference in terms of sound amplitude in the high-contrast condition, the room acoustics were more salient in the low-contrast condition. This can be due to more external validity of the low-contrast condition compared to the high-contrast condition. Overall, there is no clear pattern of a larger MMN in the high-contrast condition. Therefore, it seems reasonable to conclude that the MMNs in the present study did not hinge upon interaural intensity differences, neither overall nor point by point. Instead, differing room acoustics stand as the elicitor.
11 General discussion and Outlook

11.1 General discussion

The first experiment investigated auditory monitoring outside the focus of attention. Stimuli with differing room acoustics were randomly presented in a passive oddball protocol. In this experiment, auditory irregularities brought about by a change in room acoustics elicit a MMN in the absence of a corresponding intention and little awareness of the acoustic stimulation. Inasmuch as such automatic detection provides cues of behavioral relevance, these findings can be considered a first step concerning a thorough understanding of the usage of auditory monitoring for adapting to room acoustics-mediated environmental changes. In the first experiment, it was successfully observed for the first time that room acoustical changes can be detected outside the focus of attention. This was confirmed by eliciting a MMN following the deviant stimuli.

The results of the second and third experiment replicate and extend these findings of MMN elicitation by room acoustics-related changes outside the focus of attention in Experiment 1. Room acoustical changes are much more complex than changes of primary auditory dimensions. Therefore, it is important to control for possible confounds with other auditory dimensions and the replication of the results from Experiment 1 is essential to get further insight on how room acoustics are perceived in the brain.

In all experiments, deviant stimuli were consistently associated with a pronounced negative deflection at frontal electrode sites, starting as early as 180 ms after stimulus onset. In contrast to Experiment 1 in the second and third experiment a symmetrical room, with no room modes, was used for the simulation of the room acoustical changes. It was important to use this adaption in room acoustical simulation to remove possible confounding of overall right and left ear intensity differences between deviant and standard using a simulation technique without room modes. This is done to eliminate this possible confound completely. Standards and deviants were matched regarding their left-right overall intensity differences, dismissing an account of our results in terms of inter-aural intensity differences. MMNs occurred not only in the high contrast condition (i.e., Alpha 1 vs. Alpha 10) but also in the subtler low contrast condition (i.e., Alpha 3 vs. Alpha 7). This provides more evidence how well the auditory system can detect even subtle changes in room acoustics. This seems to happen effortless in the absence of attention like in an automatic process. This corresponds
well to the findings of behavioral studies pointing out the relevance of room acoustics to human.

Replicating previous findings from the first experiment additional negative deflections occurred subsequently to the initial MMN triggered by the deviant stimulus. Contrasting with the first experiment, in which sequences of heterogeneous piano chords were used as stimuli, in the experiments, triplets of homogenous room acoustical modified sinusoidal tones (Experiment 2) or a constant room acoustical modified sinusoidal tone (Experiment 3) as standards and deviants, were applied. These modifications did not abolish the occurrence of later negative deflections, following the first MMN. The temporal correspondence of the negative-going peaks and the onsets of the constituent tones of a tone triplet is in line with the assumption that they reflect consecutive instances of mismatch detection, related to each individual tone of the deviant triplet, suggesting that neither complexity nor “novelty” (i.e., deviance from the directly preceding tone) constitute necessary conditions for multiple MMNs to occur. Given that the acoustic characteristics of the second and third tone were completely predictable due to the identical relation with the first tone of a triplet, the occurrence of tone-related multiple MMNs suggests that separable expectations were formed for the constituent elements of the repeatedly presented sequence, which were not adjusted after identification of the first tone of a deviant sequence (see Jacobsen et al., 2013; for a discussion of successive MMNs to predictable consecutive deviations).

The results of Experiment 3 demonstrate, however, that the occurrence of a second negative deflection does not depend on the presentation of successive tones. An additional negative deflection occurred, in the high-contrast and the low-contrast condition respectively. These findings are in line with the assumption that acoustic irregularities associated with a complex pattern of changes in the stimulus may elicit two successive processes of deviance detection, that is, a fast process involving the detection of one or more low-level feature changes, and a slower process of identifying the meaning of the specific irregular pattern of feature changes. In the case of deviant stimuli made up of successive separated tones, the occurrence of a tone-onset-related additional MMN may be superimposed on the second detection response.

These results of the three Experiments provide support for the hypothesis of automatic detection of unattended changes in room acoustics, demonstrating the elicitation of a MMN by changes in room acoustics that cannot be attributed to overall inter-aural intensity differences between standards and deviants. Further, support for the hypothesis that the
auditory complex room acoustical changes may elicit two successive processes of deviance detection, was found.

11.2 Outlook:
Clearly, more research is needed to achieve a full-fledged account of the aspects of room acoustics-mediated irregularity that are critical for eliciting a MMN response as well as the boundary conditions thereof. Given the diversity and complexity of room acoustics-related manipulations, this should turn out to be a tedious but important endeavor. The results of the current experiments show that a difference in primary auditory features, such as loudness or left-right intensity distribution of the stimulus is not necessary for the detection of unattended room acoustics-mediated auditory irregularity. Therefore, it can be assumed that room acoustical changes alone can be detected automatically.

Further research is needed to better understand the later deflections and to test different hypothesis about the additional deflections. This research is needed to answer the question if room acoustical changes are processed like other sound dimensions, for example pitch. On the other hand, this thesis provides significant evidence that auditory changes in room acoustics are processed pre-attentively.
References:


References


References


# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>dB</td>
<td>Decibel</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalogram</td>
</tr>
<tr>
<td>ERP</td>
<td>Event-related potential</td>
</tr>
<tr>
<td>EOG</td>
<td>Electrooculogram</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>ILD</td>
<td>Interaural level differences</td>
</tr>
<tr>
<td>kHZ</td>
<td>Kilo Hertz</td>
</tr>
<tr>
<td>LDN</td>
<td>Late Discriminative Negativity</td>
</tr>
<tr>
<td>ML</td>
<td>Left mastoid</td>
</tr>
<tr>
<td>MMN</td>
<td>Mismatch Negativity</td>
</tr>
<tr>
<td>MR</td>
<td>Right mastoid</td>
</tr>
<tr>
<td>ms</td>
<td>Milisecond</td>
</tr>
<tr>
<td>Ω</td>
<td>Ohm</td>
</tr>
<tr>
<td>s</td>
<td>Seconds</td>
</tr>
</tbody>
</table>
Part IV

Appendix:
A Informed written consent

**Vertraulich**

Experiment:
Name der Versuchsperson: ________________________________
Name des Versuchsleiters: Johannes Frey

**Erklärung**

Hiermit erkläre ich, an den/m Experiment(en) am Institut für Allgemeine Psychologie der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg freiwillig teilgenommen zu haben. Ich wurde vor Beginn des Experiments vollständig über die Natur des Experiments aufgeklärt.

Mir wurde mitgeteilt, dass meine zur Identifizierung notwendigen persönlichen Daten vom Versuchsleiter vertraulich behandelt werden und dass er/sie keinen anderen Personen als den mit der Durchführung des Experiments Betrauten Zugang zu diesen Daten gewähren wird. Ich bin damit einverstanden, dass die erhobenen Daten anderen Forschern zur Verfügung gestellt werden, soweit diese nicht zur Identifizierung meiner Person ausreichen.

Ich bewahre mir das Recht, das Experiment jederzeit abzubrechen. In diesem Fall sollen alle erhobenen Daten gelöscht werden. Ich bin mir bewusst, dass ich in diesem Fall keinen vollen Anspruch auf Bezahlung/Gutschrift von Versuchspersonenstunden habe.

Ich erkläre, meinem Wissen nach nicht an neurologischen Störungen oder Erkrankungen zu leiden. Ich stehe nicht unter dem Einfluss von Beruhigungsmitteln oder Medikamenten die auf das zentrale Nervensystem wirken.

Ich weiß, dass die Daten in diesem Experiment ausschließlich für Forschungszwecke und nicht zur Diagnostik erhoben werden, und ich werde keine Auskunft oder Expertenmeinung dieser Art fordern.

Hamburg, __________________________
Datum __________________________
Unterschrift
B Questionnaire after the experiments

Fragebogen nach dem Experiment (____)  VP-Code:

Bitte beantworten Sie die folgenden Fragen der Reihe nach!

1. Welchen Film haben Sie geguckt?

2. Fassen Sie den Inhalt der Handlung, soweit sie den Film gesehen haben, so komprimiert wie möglich zusammen.

3. Was haben Sie gehört?

4. Beschreiben Sie bitte die Töne, die sie gehört haben.

5. Was ist Ihnen an den Tönen aufgefallen?
6. Wie störend haben Sie die Töne empfunden?

kaum störend  1  2  3  4  5  stark störend

7. Wie gut konnten Sie die Töne ignorieren? Ist es Ihnen schwergefallen, die Töne zu ignorieren? Wenn ja, wann?

Bitte erst wenden, wenn Sie alle Fragen beantwortet haben.
8. Können Sie sagen, ob Sie das Gefühl hatten die Töne haben sich im Verlauf eines Blockes in irgendeiner Weise verändert? Wenn nein beschreiben sie warum nicht und wenn ja, versuchen sie so genau wie möglich zu beschreiben welche Veränderung Sie wahrgenommen haben.

Vielen Dank, dass Sie an dem Experiment teilgenommen haben.
### EEG-Recording report

#### Experiment _____

<table>
<thead>
<tr>
<th>Vp-Code:</th>
<th>Datum:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Zeit:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geschlecht</th>
<th>Alter</th>
<th>Beruf</th>
<th>Hörvermögen</th>
<th>Sichtigkeit</th>
<th>Händigkeit</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>männlich</td>
<td></td>
<td></td>
<td>normal</td>
<td>normal</td>
<td>rechts</td>
<td>gesund</td>
</tr>
<tr>
<td>weiblich</td>
<td></td>
<td></td>
<td>korrigiert</td>
<td>korrigiert</td>
<td>links</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>beeinträchtigt</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oldfield</th>
<th>Rechte Hand</th>
<th>Teilweise linke Hand</th>
<th>Linke Hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Händigkeit der Eltern/Geschwister</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schreiben / Zeichnen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schneiden</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zähne putzen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Werfen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Löflel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behälter öffnen</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Allgemeine Fragen

<table>
<thead>
<tr>
<th>Zigaretten</th>
<th>Ja</th>
<th>Nein</th>
<th>wenn ja, wieviel tgl.</th>
<th>wann zuletzt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alkohol</td>
<td>Ja</td>
<td>Nein</td>
<td>wenn ja, wieviel tgl.</td>
<td>wann zuletzt</td>
</tr>
<tr>
<td>Medikamente</td>
<td>Ja</td>
<td>Nein</td>
<td>wenn ja, wieviel tgl.</td>
<td>wann zuletzt</td>
</tr>
<tr>
<td>Kaffee/Koffein</td>
<td>Ja</td>
<td>Nein</td>
<td>Wenn ja, wieviel tgl.</td>
<td>wann zuletzt</td>
</tr>
<tr>
<td>Schlaf der letzten Nacht</td>
<td>gut</td>
<td>mäßig</td>
<td>schlecht</td>
<td>Wie viel Stunden?</td>
</tr>
<tr>
<td>Schon vorher EEG</td>
<td>Ja</td>
<td>Nein</td>
<td>Wenn ja, welche Art von Experiment?</td>
<td>Sprache</td>
</tr>
<tr>
<td>Konzentration</td>
<td>gut</td>
<td>mäßig</td>
<td>schlecht</td>
<td></td>
</tr>
</tbody>
</table>

Rückseite füllt der Versuchsleiter aus…
### Impedanzen/Dauer (füllt Versuchsleiter aus):

<table>
<thead>
<tr>
<th>Impedanzen (alle kleiner als 10kΩ):</th>
<th></th>
<th>Dauer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ja</td>
<td></td>
<td>Vorbereitung</td>
</tr>
<tr>
<td>Nein (wenn nein, welche nicht?)</td>
<td></td>
<td>Experiment</td>
</tr>
<tr>
<td>&gt; 10 kΩ Beginn</td>
<td></td>
<td>Gesamt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Startzeit Experiment</td>
</tr>
<tr>
<td>&gt; 10 kΩ Ende</td>
<td></td>
<td>Nachbereitung</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gesamt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gesamt</td>
</tr>
</tbody>
</table>

### Film:

Probleme mit bestimmten Elektroden während des Experimentes?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Probleme mit den Kopfhörern (Funktion in allen Blöcken gewährleistet)
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Auffälligkeiten bei der VP
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Probleme beim Anbringen der Elektroden
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Sonstige Bemerkungen und Notizen:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Bitte alle Ereignisse während des Experimentes genau dokumentieren!

Versuchsleiter: ..........................................................
D MATLAB code for stimulus generation

%MATLAB script for stimulus generation (see main text for details)
%requires external functions:
% rms.m (https://de.mathworks.com/matlabcentral/fileexchange/11871-signal-rms)
% powernoise.m (http://freesourcecode.net/matlabprojects/5649/sourcecode/powernoise.m)

clear;clc;
fs=48e3;

tx=0.1;
Delay=0.01;
Alpha=1;
% Comment: Alpha was adjusted to generate 4 different stimuli (Alpha = 1,3,7 or 10)
n=tx/(1/fs);

F=0.2;
HL1=powernoise(F,tx*fs);
HL2=powernoise(F,tx*fs);

signHL1=HL1./abs(HL1);
HL2=abs(HL2).*signHL1;

Dn=floor(Delay/(1/fs));
HL1(1:Dn)=0;
HL2(1:Dn)=0;

A=gausswin(2*n,Alpha)';
A=A(n+1:2*n)';

HL1a=HL1.*A;
HL2a=HL2.*A;
fx=18000;
nx=1;
[numd,dend]=butter(nx,fx/(fs/2),'low');
HL2a=filter(numd,dend,HL2a);
HL1a=filter(numd,dend,HL1a);

fx=50;
nx=2;
[numd,dend]=butter(nx,fx/(fs/2),'high');
HL2a=filter(numd,dend,HL2a);
HL1a=filter(numd,dend,HL1a);

RMS1=rms(HL1a,length(HL1a),length(HL1a)-1,0);
RMS2=rms(HL2a,length(HL2a),length(HL2a)-1,0);

HL1a=HL1a/RMS1*1;
HL2a=HL2a/RMS2*1;

frequency=500;
t1=0.1;  \textbf{%Comment: for Experiment 3: t1=0.5}
S2=sin(2*pi*frequency*linspace(0,t1,t1*fs))';
S2=sign(S2);

tnull=0.1;
Null=zeros(tnull*fs,1);

L=S2;
R=S2;

timeL=linspace(0,length(L)*1/fs,length(L));
rise=5e-3;
IV Appendix

Flanke1=linspace(0,1,rise/(1/fs));
Flanke2=linspace(1,0,rise/(1/fs));

Window=ones(1,max(timeL)/(1/fs));
Window(1:rise/(1/fs))=Flanke1;
Window((max(timeL)/(1/fs)-rise/(1/fs))+1:max(timeL)/(1/fs))=Flanke2;

L=L.*Window';
R=R.*Window';

L=[L;Null;L;Null;L]; %Comment: for Experiment 3: L=[L];
R=[R;Null;R;Null;R]; %Comment: for Experiment 3: R=[R];
Ende=3*t1+2*null; %Comment: for Experiment 3: Ende=1*t1;

YL=conv(HL1a,L);
YR=conv(HL2a,R);

Hdelta=zeros(length(HL1a),1);
Hdelta(1)=1;

XL=conv(Hdelta,L);
XR=conv(Hdelta,R);

RMSYL=rms(YL,length(YL),length(YL)-1,0);
RMSYR=rms(YR,length(YR),length(YR)-1,0);
YL=YL/RMSYL*1;
YR=YR/RMSYR*1;

ZL=1*XL+1*YL;
ZR=1*XR+1*YR;

ZL=ZL(1:Ende/(1/fs));
ZR=ZR(1:Ende/(1/fs));
time=linspace(0,length(ZL)*1/fs,length(ZL));

rise=5e-3;
Flanke1=linspace(0,1,rise/(1/fs));
Flanke2=linspace(1,0,rise/(1/fs));

Window=ones(1,max(time)/(1/fs));
Window(1:rise/(1/fs))=Flanke1;
Window((max(time)/(1/fs)-rise/(1/fs))+1:max(time)/(1/fs))=Flanke2;
ZL=ZL.*Window';
ZR=ZR.*Window';

RMSZL=rms(ZL,length(ZL),length(ZL)-1,0);
RMSZR=rms(ZR,length(ZR),length(ZR)-1,0);
ZL=ZL/RMSZL*1;
ZR=ZR/RMSZR*1;

ZL=ZL/5;
ZR=ZR/5;

zf=[ZL ZR]*0.08;
sound(zf,fs);
wavwrite((zf),fs,16,'room_alpha1')
E Statistical analyses

Complete results of the repeated-measures ANOVAs including the factors Stimulus (standard, deviant), Position (F-, C-, P-line) and Laterality (3-, z-, 4-line) are presented in Tables 1, 2, 3, 7, 8, 9, 13, 14, 17, and 18. Additional analyses were carried out at the mastoids (Tables 4, 5, 6, 10, 11, 12, 15, 16, 19, 20) and included the factors Stimulus (standard, deviant) and Laterality (left, right). The Greenhouse-Geisser-correction was employed when applicable. Uncorrected degrees of freedom and corrected p-values are reported.

Table 1

*Experiment 2: High contrast condition. Time window 208 ms to 248 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>51.84*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>33.20*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>3.37</td>
<td>.073</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>89.12*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>1.88</td>
<td>.184</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>2.86</td>
<td>.074</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>3.46</td>
<td>.033</td>
</tr>
</tbody>
</table>

*Notes.* *p < .05.

Table 2

*Experiment 2: High contrast condition. Time window 420 ms to 460 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>45.62*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>50.51*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>12.50*</td>
<td>.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>84.09*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>10.06*</td>
<td>.003</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>2.78</td>
<td>.057</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>1.84</td>
<td>.157</td>
</tr>
</tbody>
</table>

*Notes.* *p < .05.
Table 3

*Experiment 2: High contrast condition. Time window 538 ms to 578 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>45.53*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>66.38*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>17.93*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>119.80*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>11.42*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>4.23*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>2.26</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.

Table 4

*Experiment 2: High contrast condition. Time window 208 ms to 248 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>38.23*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>2.40</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>0.51</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.

Table 5

*Experiment 2: High contrast condition. Time window 420 ms to 460 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>72.20*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>5.31*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>2.01</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.
### Table 6

**Experiment 2: high contrast condition. Time window 538 ms to 578 ms**

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>107.23*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>2.05</td>
<td>.173</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>0.42</td>
<td>.526</td>
</tr>
</tbody>
</table>

*Notes. *p < .05.*

### Table 7

**Experiment 2: Low contrast condition. Time window 160 ms to 200 ms**

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>62.07*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>21.75*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>13.40*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>27.94*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>12.18*</td>
<td>.001</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>3.00</td>
<td>.081</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>2.97</td>
<td>.074</td>
</tr>
</tbody>
</table>

*Notes. *p < .05.*
Table 8

*Experiment 2: Low contrast condition. Time window 352 ms to 392 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>30.22*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>49.29*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>25.70*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>55.15*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>16.52*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>5.11*</td>
<td>.004</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>3.55*</td>
<td>.027</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.

Table 9

*Experiment 2: Low contrast condition. Time window 542 ms to 582 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>42.52*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>85.73*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>7.16*</td>
<td>&lt;.011</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>116.90*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>4.49*</td>
<td>.039</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>2.11</td>
<td>.120</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>2.20</td>
<td>.099</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.
Table 10

*Experiment 2: Low contrast condition. Time window 160 ms to 200 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>30.07*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>0.43</td>
<td>.521</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>0.01</td>
<td>.914</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.

Table 11

*Experiment 2: Low contrast condition. Time window 352 ms to 392 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>25.89*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>12.69*</td>
<td>.003</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>6.10*</td>
<td>.026</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.

Table 12

*Experiment 2: Low contrast condition. Time window 542 ms to 582 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>70.05*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>1.93</td>
<td>.184</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>0.12</td>
<td>.739</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.
### IV Appendix

Table 13  
*Experiment 3: High contrast condition. Time window 220 ms to 260 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>39.03*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>0.10</td>
<td>.801</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>0.56</td>
<td>.538</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>25.24*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>6.18*</td>
<td>.012</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>0.75</td>
<td>.490</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>3.61*</td>
<td>.022</td>
</tr>
</tbody>
</table>

*Notes. * p < .05.*

Table 14  
*Experiment 3: High contrast condition. Time window 518 ms to 558 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>15.52*</td>
<td>.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>4.01</td>
<td>.056</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>2.02</td>
<td>.154</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>26.57*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>2.55</td>
<td>.120</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>0.28</td>
<td>.812</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>2.29</td>
<td>.098</td>
</tr>
</tbody>
</table>

*Notes. * p < .05.*

Table 15  
*Experiment 3: High contrast condition. Time window 220 ms to 260 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>14.83*</td>
<td>.002</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>0.01</td>
<td>.925</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>0.01</td>
<td>.943</td>
</tr>
</tbody>
</table>

*Notes. * p < .05.*
Table 16

*Experiment 3: High contrast condition. Time window 518 ms to 558 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>2.46</td>
<td>.137</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>1.79</td>
<td>.201</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>0.26</td>
<td>.617</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.

Table 17

*Experiment 3: Low contrast condition. Time window 238 ms to 278 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>25.68*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>4.03</td>
<td>.059</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>2.58</td>
<td>.101</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>19.86*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>4.58*</td>
<td>.034</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>1.85</td>
<td>.178</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>0.95</td>
<td>.402</td>
</tr>
</tbody>
</table>

*Notes.* *p* < .05.
Table 18

*Experiment 3: Low contrast condition. Time window 538 ms to 578 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>21.80*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Position</td>
<td>2,30</td>
<td>3.20</td>
<td>.086</td>
</tr>
<tr>
<td>Laterality</td>
<td>2,30</td>
<td>8.05*</td>
<td>.004</td>
</tr>
<tr>
<td>Interaction Stimulus*Position</td>
<td>2,30</td>
<td>23.26*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality</td>
<td>2,30</td>
<td>7.75*</td>
<td>.005</td>
</tr>
<tr>
<td>Interaction Position*Laterality</td>
<td>4,60</td>
<td>2.08</td>
<td>.144</td>
</tr>
<tr>
<td>Interaction Stimulus<em>Position</em>Laterality</td>
<td>4,60</td>
<td>1.18</td>
<td>.325</td>
</tr>
</tbody>
</table>

*Notes.* *p < .05.

Table 19

*Experiment 3: Low contrast condition. Time window 238 ms to 278 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>6.87*</td>
<td>.019</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>0.79</td>
<td>.389</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>2.32</td>
<td>.149</td>
</tr>
</tbody>
</table>

*Notes.* *p < .05.

Table 20

*Experiment 3: Low contrast condition. Time window 538 ms to 578 ms*

<table>
<thead>
<tr>
<th>Variables</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus</td>
<td>1,15</td>
<td>12.95*</td>
<td>.003</td>
</tr>
<tr>
<td>Laterality Left/Right</td>
<td>1,15</td>
<td>33.64*</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction Stimulus*Laterality Left/Right</td>
<td>1,15</td>
<td>13.41*</td>
<td>.002</td>
</tr>
</tbody>
</table>

*Notes.* *p < .05.
F Time domain representations of the acoustic stimuli from Experiment 2 and 3
G Peer-reviewed papers
Automatic detection of unattended changes in room acoustics

Johannes Daniel Frey*, Mike Wendt, Thomas Jacobsen

Experimental Psychology Unit, Helmut Schmidt University/University of the Federal Armed Forces, Hamburg, Germany

HIGHLIGHTS

• An auditory automatic detection of changing room acoustics is proposed.
• A passive oddball protocol including auditory stimuli with deviating room acoustics was used.
• Violation of auditory regularities with respect to room acoustics resulted in a mismatch negativity.
• The mismatch negativity reflects automatic detection of violations of auditory regularities.
• Violation of auditory regularities due to changed room acoustics are detected automatically.

ARTICLE INFO

Article history:
Received 11 July 2014
Received in revised form 12 September 2014
Accepted 29 September 2014
Available online 6 October 2014

Keywords:
Event-related potentials (ERP)
Mismatch negativity (MMN)
Pre-attentive auditory processing
Auditory space perception
Virtual acoustics
Auditory room effects

ABSTRACT

Previous research has shown that the human auditory system continuously monitors its acoustic environment, detecting a variety of irregularities (e.g., deviation from prior stimulation regularity in pitch, loudness, duration, and (perceived) sound source location). Detection of irregularities can be inferred from a component of the event-related brain potential (ERP), referred to as the mismatch negativity (MMN), even in conditions in which participants are instructed to ignore the auditory stimulation. The current study extends previous findings by demonstrating that auditory irregularities brought about by a change in room acoustics elicit a MMN in a passive oddball protocol (acoustic stimuli with differing room acoustics, that were otherwise identical, were employed as standard and deviant stimuli), in which participants watched a fiction movie (with subtitles). While the majority of participants reported no awareness for any changes in the auditory stimulation, only one out of 14 participants reported to have become aware of changing room acoustics or sound source location. Together, these findings suggest automatic monitoring of room acoustics.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The mismatch negativity (MMN), component of the human event-related brain potential (ERP), is a well-established measure for the investigation of pre-attentive auditory processing. The MMN is widely considered to reflect detection of violations of regularities extracted from the acoustic environment, that occurs even when the acoustic stimulation is not in the focus of attention [1]. In the simplest form, it is observed when a repeated sound (standard) is followed by a differing sound (deviant) at an unpredictable time. A mechanism that constantly monitors the acoustic environment and detects changes is likely to govern this process [2,3]. Different auditory dimensions have been shown to elicit MMN. So far, the four first-order auditory regularity violation dimensions that have been found to elicit a MMN are pitch [4–6], duration [7–9], loudness [10,11] and sound source location [12–14]. Higher-order auditory regularity violation dimensions eliciting a MMN have also been reported, for instance the omission of a tone in a recurrent pattern [15,16], or by speech stimuli violating abstract phonological rules followed by a sequence of standard stimuli [17].

Another auditory dimension that bears importance for perception and behavior, particularly in real life contexts, relates to sound properties arising from the reflecting characteristics of objects that make up the environment of the sound source-perceiver system. In a built-up environment as well as in a considerable portion of the natural environment, virtually all sound is affected by this phenomena, referred to as room acoustics. Previous research has shown that room acoustics impact both perceptual quality and behavioral performance. For instance, the reverberation time of a room influences the ability to localize the sounds, especially for the localization of continuous broadband noise [18]. Behavioral relevance of room acoustics has been demonstrated for (musical) sound production (i.e., professional piano players adapted their
playing style to varying room acoustics [19]). Additionally, variations in room acoustics created by virtual rooms differing in size, influence the emotional valence of sounds [20]. Consequently a metric to assess the mental and neural mechanisms underlying the processing of room acoustics could improve research in this field.

In the current study, we investigated whether the human auditory system monitors the acoustic environment regarding this particular dimension. Given the sophisticated ability in detecting acoustic irregularities, on the one hand, and the perceptual–behavioral–emotional importance of room acoustics on the other, it seems likely that room acoustics-based sound changes are detected automatically (i.e., in the absence of a corresponding goal and possibly without awareness). Humans are not necessarily aware of these changes, which renders measuring brain waves a good means of observing reaction to changing room acoustics [21]. To examine this issue, we applied ERP recording in a passive oddball protocol involving standard and deviant stimuli that differed regarding a room-acoustical aspect.

2. Materials and methods

2.1. Participants

Fourteen volunteers participated in the experiment (five male, mean age 25.5 years, range 19–60, one left handed). Handedness was assessed using an inventory adopted from Oldfield [22]. All participants were native German speakers, reported normal auditory and normal visual acuity and no neurological, psychiatric, or other medical problems. The experiment was carried out in line with ethical guidelines, in particular The Code of Ethics of the World Medical Association (Declaration of Helsinki) [23]. Informed written consent was obtained from all participants prior to the experimental session.

2.2. Materials

By playing a sequence of five piano chords in two variations of a simulated room, two stimuli were constructed. A room acoustics software program (Odeon 11.00 Combined Demo Version; Odeon AS, Kgs. Lyngby, Denmark) was used for auralisation. Odeon was developed for simulating the interior acoustics of buildings and uses the image-source method combined with ray tracing. An ambiophonic recorded stimulus (derived from “Piano Over the rainbow Mic2 SHORT.wav” from the Odeon package) consisting of a sequence of five piano chords was used: F7 (2793 Hz), E7 (2637 Hz), D7 (2349 Hz), C7 (2093 Hz), & G6 (1567 Hz; base frequencies given in the parentheses). To avoid difficulties in perceiving room acoustics based on a single tone, we chose a stimulus of considerable complexity and duration. The chord sequence had an overall duration of 1040 ms, including 5 ms rise and 5 ms fall times. The onset times of the chords were at approximately 5 ms, 350 ms, 510 ms, 660 ms, and 837 ms after stimulus onset, with no silent periods between consecutive chords. For the auralisation a virtual room (“example Par” from the Odeon package) was used. The simulated room’s acoustic properties were altered to generate two auditory stimuli with different room acoustics but otherwise retained identical properties. The sound source was centered in front of the perceivers (point source; (x, y, z) = (0, 0, 2); see Fig. 1A), and the connection between the sound source and the receiver formed an imagined line dividing the room into two symmetrical parts. The perceivers was seated in a central position with respect to the right and left walls (single point response receiver; (x, y, z) = (20, 0, 5); see Fig. 1A). The surface area of the room was 1268.23 m², room temperature 20°C, relative humidity 50%. The reflective properties of walls were altered to generate two stimuli with different room acoustics (Fig. 1A). For one stimulus (“right”) a 90% absorbing material (equally absorbing all frequencies) was applied to the walls to the right of the receiver and for all other walls a 10% absorbing material (equally absorbing all frequencies) was used. This room setup produced the impression that the room was open to the right. For the second stimulus (“left”) a 90% absorbing material was applied to the walls to the left of the receiver and all other walls were covered with 10% absorbing material, giving the impression that the room was open to the left, creating a fully symmetric counterpart. As a consequence, the total Root Mean Square (RMS) of “left” and “right” was the same. Remaining intensity differences due to stochastic aspects of the re-synthesis procedure were equalized using Adobe Audition CS5.5 Demo Version (Adobe Systems GmbH, München, Germany); mean RMS amplitude for “right” (left channel: −18.33 dB; right channel: −21.91 dB) and “left” (left channel: −22.07 dB; right channel: −18.31 dB). Different mean RMS amplitudes between the channels were essential in order to maintain the different acoustic properties of the two rooms (see Fig. 1C). As a consequence, there are intensity differences between the two channels. The frequency spectrum (see Fig. 1B) reveals that channel right of stimulus “left” is not identical as channel left of stimulus “right”. Software simulating acoustics cannot make perfect calculations therefore each auralisation does differ slightly. The acoustic stimuli can be found at: http://www.hsu-hh.de/allgpsychologie/index_1d3q1elqG8cZ0048.html (Fig. 2).

Deviant stimuli could be actively discriminated from standard stimuli with high accuracy. To test participants’ awareness of the deviant stimuli in the passive oddball protocol the fourteen participants of the EEG experiment were interviewed about their subjective impression regarding the auditory stimuli (approximately 5 min after completing the EEG experiment). Ten participants reported that they did not notice any changes to the auditory stimuli, two participants said they felt that the rhythm was sometimes different, one participant reported differences in the sound’s source location and only one participant had the feeling that something with the room’s acoustics changed, but could not specify this observation further.

2.3. Experimental design and procedure

The participants were seated in an electrically and acoustically shielded experimental chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). 2000 acoustic stimuli were presented binaurally at approximately 52 dB SPL [2] (artificial head HMS

---

1 The following Odeon configuration was used: (1) Room setup: Impulse Response Length 16,000 ms, Number of late rays 20,000, Max. reflection order 2000, Impulse response resolution 3.0 ms, Transition Order 2, Number of early scatter rays 100, Angular absorption “Soft materials only”, Surface scattering “Actual”, Oblique Lambert, Reflection based scatter “enabled”, Key diffusion frequency 707 Hz, Interior margin 0.10 ms, Scatter coefficients > 0.5 handled as uniform scatter. (2) Auralisation setup: Apply dither and noise shaping, Wave result file 16 bit PCM, Create binaural impulse response file, HRIR “Subject_021-Res10deg_M3_0_5Rate44100_Apass50.5_Astop40.00_BovrLap1000_PPHPRT256”, Headphone “Sennheiser HD250 Linear LL.44100.ee.hph”, DC filter, Overall Recording level 40 dB, Phase approximation “phase shift at surfaces/filter phase, A/stop) 40,000 dB, A/Pass) 0.50 dB, Band overlap 100%, Sample rate 44100 Hz, Encoding “1. Order ambisonics”.

2 To assess discriminability of the two stimuli, another group of ten participants (six male, mean age 28.9 years, range 20–50) who reported normal auditory acuity were asked to detect the one deviant stimulus in a sequence of 10 stimuli. The “right” stimulus was interspersed in a sequence of nine presentations of the “left” stimulus and vice versa. All ten participants completed 20 of these sequences with one deviant at a random position in each of the sequences (1800 standard; 200 deviant). In total, 185 deviants (92.5%) were detected with 4 false alarms (0.22%).

3 Configuration: equalization (LIN), synchronization (44.1 kHz), tool (SPL), without torso.
The room on the left was used for the auralisation of Stimulus Left and the room on the right was used for the auralisation of Stimulus Right. The gray shaded walls represent the walls with 90% absorbing material; all other walls represent walls with 10% absorbing material. The receiver symbol represents a single point response receiver and the sound source symbol represents a point source. (B) Frequency spectrum of the auditory stimuli, right channel on top and left channel on the bottom. Stimulus Left on the left and Stimulus Right on the right. The five vertical lines indicating the onset of each piano chord. (C) Results of the intensity measurements with the artificial head HMS III.2. The values on the left are the values from Stimulus Left, with the right channel on top and the left channel on the bottom. The values on the right are the values from Stimulus Right, with the right channel on top and the left channel on the bottom. Values are measured in dB (SPL) and dB (A).

### III.2: HEAD acoustics GmbH, Herzogenrath, Germany
Headphones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany) in two blocks (~19 min each with a 5 min break between the blocks). The participants were instructed to ignore the acoustic stimuli and concentrate on a movie with German subtitles. All participants reported that they could ignore the acoustic stimuli easily while watching the movie. The movie was shown on a standard 24" 16:10 LCD computer screen at an approximate distance of 130 cm. The participants performed no additional tasks and did not respond to the acoustic stimuli in any way.

There were two types of blocks. In block 1 the “right” stimulus was used as standard (90% of the stimuli; n = 900) and the “left” was used as deviant (10% of the stimuli; n = 100). In block 2 the “left” was used as standard and the “right” was used as deviant. The stimuli were presented in a passive oddball protocol in a pseudo-random fashion with the constraints of five consecutive standard stimuli in the beginning of a block, and no consecutive deviant stimuli. The sequence of the two blocks was counterbalanced across participants.

### 2.4. Electrophysiological recordings

The Electroencephalogram (EEG) (Ag/AgCl electrodes, Falk Minow Services, V-Amp EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from nine standard scalp locations according to the extended 10–20 system [24]: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4 and the left and right mastoids. Standard scalp electrodes were embedded in a prefabricated cap (Brain Products GmbH, Gilching, Germany). For the recording, a sampling rate of 500 Hz (resolution 16 bit) was used. The ground electrode was placed at FCz and the reference on the nose. Electro-ocular activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) from the left eye (for one participant the right eye was used) by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 5 kΩ. On-line filtering was carried out using a 0.011-Hz high-pass, a 100-Hz low-pass and 50-Hz notch filter.

### 2.5. Data analysis

Off-line signal processing was carried out (EEProbe 3.0; Advanced Neuro Technology, Enschede, Netherlands) on a Linux computer. EEG-data was band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 1140 ms, time-locked to the onset of the stimuli, including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each condition (standard and deviant) and for each participant. The ERP responses to the first five stimuli of each block as well as to the first standard stimulus after each deviant were not included in the analyses. Epochs showing an amplitude change exceeding 120 µV at any of the recording channels were rejected. Grand-averages were subsequently computed from the

<table>
<thead>
<tr>
<th></th>
<th>Left (dB) SPL</th>
<th>Right (dB) SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left (dB A)</td>
<td>Right (dB A)</td>
</tr>
<tr>
<td></td>
<td>Left (dB A)</td>
<td>Right (dB A)</td>
</tr>
</tbody>
</table>

**Fig. 1.** (A) The room on the left was used for the auralisation of Stimulus Left and the room on the right was used for the auralisation of Stimulus Right. The gray shaded walls represent the walls with 90% absorbing material; all other walls represent walls with 10% absorbing material. The receiver symbol represents a single point response receiver and the sound source symbol represents a point source. (B) Frequency spectrum of the auditory stimuli, right channel on top and left channel on the bottom. Stimulus Left on the left and Stimulus Right on the right. The five vertical lines indicating the onset of each piano chord. (C) Results of the intensity measurements with the artificial head HMS III.2. The values on the left are the values from Stimulus Left, with the right channel on top and the left channel on the bottom. The values on the right are the values from Stimulus Right, with the right channel on top and the left channel on the bottom. Values are measured in dB (SPL) and dB (A).
individual-subject averages. Thereby, deviant and standard grand-averages were aggregated using both the left and right stimuli, such that waveforms resulted from physically identical stimulation. Visual inspection of the ERPs derived separately for the two block types revealed no apparent substantial differences in the data patterns.

To quantify the full MMN amplitude, the scalp-recorded ERPs were re-referenced to the averaged signal recorded from the electrodes positioned over the left and right mastoids. This computation results in an integrated measure of the total neural activity underlying the auditory MMN [25]. Deviant-minus-standard difference waveforms were calculated by subtracting point by point the standard ERPs from the associated deviant ERPs. The deviance-related effects for the room acoustics change were quantified with a mean value for a fixed window of 40 ms at the center of the highest peak amplitude in the grand-average difference waves.

The MMN responses were analyzed by comparing the mean ERP amplitudes from the standard with those from the deviants using a three-way repeated-measures analysis of variance (ANOVA) with the factors Stimulus (standard, deviant), Position (F-, C-, P-line) and Laterality (3-, z-, 4-line). The first two negative deflections of the deviant-minus-standard difference waveform revealed by visual inspection were analyzed. Additionally, a two-way repeated-measure ANOVA with the factors Stimulus (standard, deviant) and Laterality (right-, left-mastoid) was carried out on non-referenced data to check for the MMN-typical polarity inversion at the mastoids.

Reduced degrees of freedom (Greenhouse–Geisser) were used where applicable to prevent violating the sphericity assumption underlying ANOVA with repeated measures. Uncorrected degrees of freedom are reported.

3. Results

On average, 7.66% of the epochs per participant were rejected prior to ERP computation (7.00% of the standard epochs and 7.94% of the deviant epochs; range 0.45–40.17%) due to amplitude changes exceeding 120 μV during an epoch. Two consecutive effects were observed. For the first effect, peak latencies occurred at 180 ms after stimulus onset with a maximum peak amplitude at electrode F4 of −0.778 μV (mean amplitude at Fz = −0.745 μV) in the deviant-minus-standard waveform. For the second effect, a peak latency at 606 ms with a peak maximum of −1.431 μV at electrode Fz (mean amplitude at Fz = −1.431 μV) was observed. Correspondingly, the time windows for the ERP quantization were set from 160 to 200 ms for the first effect and from 586 to 626 ms for the second effect.

The three-way repeated-measures ANOVAs with data for the re-referenced ERPs revealed significant main effects for Stimulus in the first time window, $F_{1,12} = 12.6$, $p = 0.004$, and the second time window, $F_{1,12} = 13.4$, $p = 0.003$. These main effects resulted from more negative-going ERP amplitudes for deviants than for standards. The same ANOVAs revealed significant effects for the interaction of Stimulus and Position for the first time window, $F_{2,12} = 10.4$, $p = 0.002$, and the second time window, $F_{2,12} = 10.4$, $p = 0.002$, demonstrating that the effects were larger at more anterior electrode sites.

The two-way repeated-measures ANOVAs on the non-referenced data of the mastoid electrodes revealed a significant main effect of Stimulus for the second time window, $F_{1,12} = 13.0$, $p = 0.003$ (mean amplitude at the left mastoid 0.708 μV; mean amplitude at the right mastoid 0.826 μV) and a marginally significant effect for the first time window ($F_{1,12} = 4.0$, $p = 0.068$) (mean amplitude at the left mastoid 0.367 μV; mean amplitude

---

4 Visual inspection revealed a third negative deflection of the deviant-minus-standard difference waveform at about 900 ms after stimulus onset, which was not further evaluated, however.
4. Discussion

This study investigated auditory monitoring outside the focus of attention. Stimuli with differing room acoustics were randomly presented in a passive oddball protocol. Although the majority of participants did not report perceiving any irregularity in the tone-ignored acoustic stimulation, event-related potentials elicited by deviants consisting of five consecutive piano chords displayed a negative deflection at 180 ms, associated with the typical characteristics of the MMN. These results are consistent with the assumption of continuous, automatic monitoring regarding changes in room acoustics.

The current study reports the first MMN experiment using changing room acoustics as a violation of auditory regularities, aiming to achieve a high external validity with respect to pre-attentive environmental auditory monitoring. Even with the auditory scene presented via headphones, the experimental situation was modeled based on a scenario where an individual is sitting in a dark living room watching a movie. The stimulus manipulation used in this study resembles the room acoustical change resulting from a large open door, changing the reflection properties of the wall.

Because changes in room acoustics-related aspects are necessarily associated with changes in more basic auditory dimensions, we cannot dismiss the possibility that one (or a specific subset) of the latter is sufficient to trigger a MMN response. Most obviously, in this connection, standard and deviant stimuli in our experiment differed regarding intensity levels at the left and right ear (i.e., interaural level difference, ILD). Manipulation of ILD is often perceived as variation regarding the azimuth of the sound source. Paavilainen et al. [13] found that irregularity regarding ILD elicited a MMN in a passive oddball protocol (although this effect did not reach statistical significance for intensity differences of 2 and 3 dB per ear, which closely resembles the manipulation used in the experiment here). Noteworthy, whereas in the study of [13] participants reported difficulty ignoring the (apparently moving) auditory stimuli, the participants of our experiment did not experience such difficulty and only one out of fourteen reported perceiving changes in sound source location.

In our study, standard and deviant stimuli were generated by manipulating the reflecting properties of the virtual walls on the left and right. By contrast, a left–right symmetrical manipulation of room acoustics, such as changing the properties of the wall in front of or behind the perceiver, would prevent ILD and generate left–right symmetrical spectral and, potentially, level differences for standard and deviant stimuli. Such deviation, produced by manipulations other than changing room acoustics, has also been found to elicit a MMN in previous studies. For example, presenting repetitive sequences of the same tone from a loudspeaker aligned with the perceivers’ sagittal midline, Winkler et al. [26] observed a MMN when the identical tone, approximately matched for overall intensity, was presented simultaneously from a left-sided and a right-sided loudspeaker. In this condition, stimulus intensity was symmetrically reduced in each of the lateral loudspeakers and the tone was perceived to originate from a sparsely extended central sound source rather than to come from a different location. This manipulation relates to a change in the size of the apparent sound source while room characteristics remained constant. Given the complexity of room acoustics-related changes in auditory stimulation, future research is needed to pinpoint the functional role of specific components, such as ILD, absolute level difference, or spectral distribution.

An aspect of our results that deserves discussion relates to the occurrence of more than one negative deflection during the course of deviant processing. While a first negative deflection peaked after 180 ms and a second negative deflection peaked after 606 ms. Although we did not explicitly expect the second effect, it seems plausible to assume that it constitutes an additional MMN, elicited by the second chord of our tone sequence [27]. That is, the salience of regularity violation may vary during the course of the presentation of a multiple chord stimulus, possibly reaching maximum after the onset of the individual chords, turning them into successive perceived deviants. The timing of the second negative deflection (the maximum amplitude of the difference wave occurred 256 ms after the onset of the second chord) as well as the MMN-typical polarity inversion at the mastoids, found for the second time window, are in line with this interpretation. Assuming that the detection of a change in room acoustics necessitates a comparatively long period of stimulation it seems an intriguing possibility that the MMN found 180 ms after stimulus onset reflects the detection of low-level acoustic changes whereas the MMN after 606 ms constitutes a reflection of the processing of scene difference.

That said, it cannot be dismissed that the negative deflection observed during the second time window might represent a later ERP component elicited during processing of the first chord rather than a MMN locked to the second chord. A possible candidate for such a component is the late discriminative negativity (LDN, a.k.a. late MMN) [28,29].

In summary, the results of the present study indicate that auditory irregularities resulting from changing room acoustics elicit an MMN in the absence of a corresponding intention and little awareness of the acoustic stimulation. Inasmuch as such automatic detection provides cues of behavioral relevance, our findings can be considered a first step concerning a thorough understanding of the usage of auditory monitoring for adapting to room acoustics-mediated environmental changes.

Acknowledgements

The valuable comments and technical help from Thomas Konstantin Jacobsen and Susan Beudt are gratefully acknowledged. The authors thank Svantje Kähler and Jonathan Manske for their help in data acquisition, Pierce Henebry for proof reading.

Appendix A. Supplementary data

Supplementary material related to this article can be found in the online version, at http://dx.doi.org/10.1016/j.neulet.2014.09.050.

References


The authors are grateful to an anonymous reviewer for suggesting this possibility.


Research article

Changes in room acoustics elicit a Mismatch Negativity in the absence of overall interaural intensity differences

Johannes Daniel Frey, Mike Wendt, Andreas Löw, Stephan Möller, Udo Zölzer, Thomas Jacobsen

A Experimental Psychology Unit, Helmut Schmidt University/University of the Federal Armed Forces Hamburg, Germany
b Central Electronics Services, Helmut Schmidt University/University of the Federal Armed Forces Hamburg, Germany
c Department of Signal Processing and Communications, Helmut Schmidt University/University of the Federal Armed Forces Hamburg, Germany
d Faculty of Human Sciences, Medical School Hamburg/University of Applied Sciences and Medical University, Germany

HIGHLIGHTS

• An auditory pre-attentive detection of changing room acoustics is proposed.
• Violation of auditory regularities with respect to room acoustics resulted in a mismatch negativity.
• The mismatch negativity reflects pre-attentive detection of violations of auditory regularities.
• Violations of auditory regularities due to changed room acoustics are detected pre-attentively.
• Additional negative deflections follow after a mismatch negativity.

ARTICLE INFO

Article history:
Received 8 April 2016
Received in revised form
24 November 2016
Accepted 29 November 2016
Available online 30 December 2016

Keywords:
Event-related potentials (ERP)
Mismatch Negativity (MMN)
Pre-attentive auditory processing
Successive MMNs
Auditory room effects
Room acoustics

ABSTRACT

Changes in room acoustics provide important clues about the environment of sound source-perceiver systems, for example, by indicating changes in the reflecting characteristics of surrounding objects. To study the detection of auditory irregularities brought about by a change in room acoustics, a passive oddball protocol with participants watching a movie was applied in this study. Acoustic stimuli were presented via headphones. Standards and deviants were created by modelling rooms of different sizes, keeping the values of the basic acoustic dimensions (e.g., frequency, duration, sound pressure, and sound source location) as constant as possible. In the first experiment, each standard and deviant stimulus consisted of sequences of three short sounds derived from sinusoidal tones, resulting in three onsets during each stimulus. Deviant stimuli elicited a Mismatch Negativity (MMN) as well as two additional negative deflections corresponding to the three onset peaks. In the second experiment, only one sound was used; the stimuli were otherwise identical to the ones used in the first experiment. Again, an MMN was observed, followed by an additional negative deflection. These results provide further support for the hypothesis of automatic detection of unattended changes in room acoustics, extending previous work by demonstrating the elicitation of an MMN by changes in room acoustics.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

A Mismatch Negativity (MMN) is a component of the event-related potential [1]. This component has consistently been observed when an infrequently presented sound, referred to as the deviant, occurs in a sequence of repetitions of a different sound, referred to as the standard, even when participants are attending to another (e.g., visual) source of stimulation and report no awareness of the occurrence of the deviant. MMN responses have been observed in connection with changes along basic auditory dimensions, such as pitch [2] or loudness [3], as well as more abstract irregularities [4]. An MMN reflects detection of a violation of an implicit prediction of impending auditory events, based on previously experienced regularity [1,5]. Given the particular sensitivity of the auditory system for perceiving distant events in the entire surrounding environment, such a monitoring mechanism seems particularly well suited for early identification of unpredictable, potentially important changes in the environment.

http://dx.doi.org/10.1016/j.neulet.2016.11.063
0304-3940/© 2016 Elsevier Ireland Ltd. All rights reserved.
Although the behavioural relevance of the perception of room acoustics has long been established [6,7], research on automatic detection of room-acoustics-related changes in the environment has only recently begun. Specifically, Frey et al. [8] observed an MMN elicitation in an oddball protocol with a sequence of piano chords that differed in room acoustics. An MMN was reliably elicited, peaking about 180 ms after the onset of the first chord of the deviant chord sequence. This supports the notion of automatic detection of unattended changes in room acoustics.

Some peculiarities of the stimuli used in the study by Frey et al. [8] merit consideration. Most importantly, due to the manipulation of lateralized reflection properties, standards and deviants in that study were not symmetrical in terms of the overall sound intensity levels presented to the left and right ear. Given previous findings of MMN elicitation by interaural intensity differences [9], one cannot dismiss the possibility that the MMN observed by Frey et al. [8] was driven by this asymmetry. Interaural intensity differences can influence the perceived spatial location of sound sources. Within a certain range, they lead to perception of a sound as coming from a lateral source [9]. In addition, room acoustics also influence the reliability of judgements of sound source locations [10]. Moreover, assuming a difficulty of detecting room-acoustics-related changes with short, homogeneous stimuli, Frey et al. [8] repeatedly presented sequences made up of five different piano chords. An analysis of the time interval subsequent to the MMN revealed a second negative deflection in the ERP, about 260 ms after the onset of the second chord of the tone sequence. Although it appears straightforward to assume multiple MMNs in this case, each reflecting the detection of a room-acoustics-related change for a particular chord, it is also conceivable that the two negative deflections reflect qualitatively different processes. Specifically, because alterations of room acoustics are characterized by a complex pattern of changes affecting various acoustic parameters, some minimal period of time may be necessary to detect the room-acoustical deviation from this pattern. Viewed from this perspective, it might be conjectured that only the second negative deflection indicated such a detection process, whereas the first MMN might have been elicited by low-level constituents of the pattern, that is, by one or more featural deviations which, in isolation, would not be indicative of a change in room acoustics. The subtle interaural intensity difference in the study by Frey et al. [8] is an obvious candidate for such a feature.

Based on these considerations, the current study sought additional evidence for the notion of automatic detection of unattended changes in room acoustics, controlling for differences in overall interaural intensity. To this end, changes in room acoustics which were left-right-symmetrical were simulated. Furthermore, we aimed to shed light on the role of multiple sounds in sequences by contrasting the presentation of three identical sounds (Experiment 1A) and a single uninterrupted sound (Experiment 1B) with the same overall duration. The assumption that multiple MMNs are elicited by the onsets of the constituent sounds of the deviant sequence predicts the occurrence of additional negative deflections only in the three-sound-sequence condition of Experiment 1A. Conversely, the assumption of consecutive detection of featural and room-acoustics–specific patterns of deviance also predicts a second negative deflection in the single-sound condition of Experiment 1B.

2. Material and methods

2.1. Participants

Sixteen volunteers participated in Experiment 1A (three female, mean age 24.6 years, age range 22–28, one left-handed). Sixteen additional volunteers participated in Experiment 1B (nine female, mean age 23.4 years, age range 21–26, one left-handed). Handedness was assessed using an inventory adopted from Oldfield [11]. All participants reported normal auditory and normal visual acuity and no neurological, psychiatric, or other medical conditions. The experiment was carried out in line with ethical guidelines, specifically, The Code of Ethics of the World Medical Association Declaration of Helsinki [12]. Informed written consent was obtained from all participants prior to the experimental session.

2.2. Materials

For Experiment 1A, a 500-ms-long sequence of three sinusoidal tones was generated. It consisted of 100 ms of tone, followed by 100 ms of silence, 100 ms of tone, 100 ms of silence, and 100 ms of tone. In addition, a constant interstimulus interval (ISI) of 100 ms was used between sequences (triplets). For Experiment 1B, a sinusoidal tone of 500 ms duration was generated, matching the tone triplet of Experiment 1A in overall duration. A constant ISI of 100 ms was also used in Experiment 1B.

The tones were then modified. First, the sinusoidal tones were converted into square waves in order to add spectral complexity. This addition of harmonics made the sounds more realistic than pure sinusoidal tones, thus adding external validity. Then, to generate different room acoustics, room simulations were used to reproduce the acoustics of a room. To this end, calculation of the stimulus parameters was based on a model of a fully symmetrical (i.e., spherical) room with no room modes. The direct audio signals without room impressions were mapped to certain room acoustics. Specifically, the acoustical room properties of the room to be simulated were transformed into a room impulse response [13]. A fast convolution algorithm was used to generate the stimuli; the sounds presented to the participants were computed as the sum of the convolution (modelling the reverberation) and the original sound (square waves). The calculation (Matlab code) is provided in the Appendix.

For each experiment, four stimuli differing in reverberation times and convolution, corresponding to rooms of different sizes, were derived. The rooms were Alpha 1 (largest room), Alpha 3 (second largest room), Alpha 7 (third largest room), and Alpha 10 (smallest room). As a consequence of the manipulation between room sizes, the first two sound segments of the sequences (triplets) in Experiment 1A slightly differed in duration from the third. Because this difference was constant we considered it unlikely that it would affect the automatic detection of changes in room acoustics. At the end of the stimulus generation, a 5 ms fade in and fade out was added to all sounds in the sequences.

Finally, the total Root Mean Square (RMS) amplitude was equalized for all four stimuli. Time domain representations of the acoustic stimuli are presented in the Appendix.

---

1 Visual inspection revealed a third negative deflection, which was, however, not analysed further.

2 An idealized room was used that cannot be found in reality, but rooms resembling this room closely can be built. This perfect room without room modes was used to eliminate confounds due to resonances depending on the room and not the stimuli. This idealized room reflects all frequencies with equal strength producing a linear spectrum without any peaks, resulting in an environment with reverberation with no resonances. The reflection properties of the simulated room resemble those of a diffuse environment like, for instance, a forest where the reflection from the trees is so diffuse that no frequencies are reflected stronger than others.
To test participants’ awareness of the deviant stimuli in the passive oddball protocol, the participants were interviewed post-experimentally about their subjective impressions regarding the acoustic stimuli. In Experiment 1A, thirteen participants reported not noticing changes in the acoustic stimuli. Three participants reported noticing differences in frequency, volume, duration, and other features. No participant reported noticing changes in perceived sound source location or differences in left-right symmetry. After being informed that changes had occurred in the acoustic stimuli and being asked to describe these changes, seven participants reported various differences (e.g., an echo, differing frequency, faster sound sequences, different sounds from time to time, and lower volume of the stimuli over the course of the experiment). Nine participants insisted that the acoustic stimuli remained the same throughout the entire experiment.

In Experiment 1B, twelve participants reported not noticing changes in the acoustic stimuli. Among the four participants who reported noticing changes, only one reported noticing changes in room acoustics. No participant reported noticing changes in the perceived sound source location or differences in left-right symmetry.

2.3. Experimental design and procedure

The participants were seated in an electrically and acoustically shielded experimental chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). Acoustic stimuli were presented binaurally at approximately 43.6 dB SPL (artificial head HMS III.2; HEAD acoustics GmbH, Herzogenrath, Germany) via head-phones (Sennheiser HD 25-1 II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). The participants were instructed to ignore the acoustic stimuli and to concentrate on a movie that was presented with subtitles. All participants reported that they were able to ignore the acoustic stimuli while watching the movie. The movie was shown on a standard 24” 16:10 LCD computer screen at an approximate distance of 130 cm from the participants. The participants performed no additional tasks and were asked not to respond to the acoustic stimuli in any way.

In each experimental block, 900 standard and 100 deviant stimuli were presented. The blocks differed with respect to the specific standard-deviant combination that was administered. A larger contrast in room acoustics was created in Blocks 1 and 2, and a smaller one in Blocks 3 and 4. We refer to the former as high-contrast condition and to the latter as low-contrast condition. In Block 1, the Alpha 10 stimulus was used as the standard and the Alpha 1 stimulus was used as the deviant. In Block 2, the roles of standard and deviant were reversed, i.e., Alpha 1 was the standard and Alpha 10 was the deviant. In Block 3, Alpha 7 was used as the standard and Alpha 3 was used as the deviant. In Block 4, the roles were again reversed, with Alpha 3 as the standard and Alpha 7 as the deviant. In these four blocks, the stimuli were presented in a pseudo-random fashion with the constraints that two deviants could not be presented in direct succession and five consecutive standard stimuli would be presented at the beginning of each block. In four additional blocks, the four stimuli were each presented in isolation 220 times (Alpha 1 in Block 5, Alpha 3 in Block 6, Alpha 7 in Block 7, and Alpha 10 in Block 8). Blocks 5–8 were used to control for potential context effects of oddball conditions on the ERP to the standard [14,15]. As no such effects were observed, Blocks 5–8 were not considered further. The sequence of presentation of all eight blocks was counterbalanced across participants.

2.4. Electrophysiological recordings

An electroencephalogram (EEG) (Ag/AgCl electrodes, Falk Minow Services, Brain Amp, EEG amplifier; Brain Products GmbH, Gilching, Germany) was recorded continuously from twenty-six standard scalp locations according to the extended 10–20 system. Standard scalp electrodes were embedded in a prefabricated cap (Brain Products GmbH, Gilching, Germany). For the recording, a sampling rate of 500 Hz (resolution: 16 bit) was used. The ground electrode was placed at Fpz and the reference on the nose. Electro-oculogram activity was recorded with two bipolar electrode pairs, the vertical electrooculogram (EOG) from the left eye by one supraorbital and one infraorbital electrode and the horizontal EOG from electrodes placed lateral to the outer canthi of both eyes. Impedances were kept below 5 kΩ. On-line filtering was carried out using a 0.011-Hz high-pass filter and a 100-Hz low-pass filter.

2.5. Data analysis

Off-line signal processing (EEProbe 3.0: Advanced Neuro Technology, Enschede, Netherlands) was carried out on a Linux computer. EEG data were band-pass filtered with a finite impulse response filter: 4001 points, critical frequencies of 0.5 Hz (high-pass) and 15 Hz (low-pass). EEG epochs with a length of 600 ms, time-locked to the onset of the stimuli, including a 100 ms pre-stimulus baseline, were extracted and averaged separately for each condition (standard and deviant) and for each participant. The ERP responses to the first five stimuli of each block as well as to the first standard stimulus after each deviant were not included in the analyses. Epochs which showed an amplitude change exceeding 120 μV (Experiment 1A) and 80 μV (Experiment 1B) at any of the recorded channels were rejected. Visual inspection of the ERPs that were computed for the separate stimuli within the high-contrast as well as the low-contrast condition revealed no apparent substantial differences in the data patterns. Therefore, the ERPs to deviants and standards were aggregated over both individual stimuli of the high-contrast condition (Alpha 1 and Alpha 10 from Blocks 1 and 2) and the low-contrast condition (Alpha 3 and Alpha 7 from Blocks 3 and 4). Thus, the ERP waveforms resulted from physically identical stimulations [14,15]. Grand averages were subsequently computed from the individual subject averages.

Deviant-minus-standard difference waveforms were calculated by subtracting, point by point, the standard ERPs from the associated deviant ERPs. The deviance-related ERP effects, elicited by room acoustics changes, were quantified by computing the mean amplitude across a 40 ms window centered around the peaks in the grand-mean difference waves. The MMN responses were analyzed by comparing the mean ERP amplitudes for the standards with those for the deviants through a three-way repeated-measures analysis of variance (ANOVA) with the factors Stimulus (standard, deviant), Position (F-, C-, P-line) and Laterality (3-, 4-line).

Three negative deflections of the deviant-minus-standard difference waveform, revealed by visual inspection, were analyzed for Experiment 1A, and two negative deflections of the deviant-minus-standard difference waveform, also revealed by visual inspection, were analyzed for Experiment 1B. Additionally, two-way repeated-measures ANOVAs with the factors Stimulus (standard, deviant) and Laterality (right-mastoid, left-mastoid) were carried out to check for the MMN–typical polarity inversion at the mastoids.

Only ANOVA statistics directly relevant for our hypotheses are reported in the Results section. A full report of the statistical analyses (including topographical factors) is provided in the Appendix.
3. Results

3.1. Experiment 1A

For the high-contrast condition (Alpha 1 and Alpha 10), 28.34% of the epochs were rejected prior to ERP computation (28.65% of the standard epochs and 25.86% of the deviant epochs; range 4.94% to 68.41%), due to amplitude changes exceeding 120 μV during an epoch. Three consecutive negative deflections were observed for the high-contrast condition (Fig. 1).

At electrode site Fz, the peak latencies of the three negative deflections were 228 ms, 440 ms, and 558 ms, with mean peak amplitude differences (deviant - standard) of -2.719 μV, -2.407 μV, and -2.555 μV, respectively. Significant main effects of Stimulus were obtained in all three ANOVAs, F(1,15) = 5.84, p < 0.001; F(1,15) = 45.62, p < 0.001; and F(1,15) = 45.53, p < 0.001, for the first, second, and third time window, respectively.

The ANOVAs of the data from the mastoids also revealed significant effects of Stimulus, resulting from mastoidal polarity inversion, F(1,15) = 38.23, p < 0.001 (mean peak amplitude difference at the left mastoid 1.482 μV, mean peak amplitude difference at the right mastoid 1.550 μV); F(1,15) = 72.20, p < 0.001 (mean peak amplitude difference at the left mastoid 1.170 μV, mean peak amplitude difference at the right mastoid 1.015 μV); and F(1,15) = 107.23, p < 0.001 (mean peak amplitude difference at the left mastoid 1.360 μV, mean peak amplitude difference at the right mastoid 1.313 μV), for the first, second, and third time window, respectively.

For the low-contrast condition (Alpha 3 and Alpha 7), on average 28.43% of the epochs were rejected prior to ERP computation (28.69% of the standard epochs and 26.36% of the deviant epochs; range 4.09% to 70.62%) due to amplitude changes exceeding 120 μV during any epoch. Again, three consecutive negative deflections were observed (Fig. 2).

At electrode site Fz, the peak latencies of the three negative deflections were 180 ms, 372 ms, and 562 ms, with mean peak amplitude differences (deviant - standard) of -3.360 μV, -2.867 μV, and -3.300 μV, respectively. Significant main effects of Stimulus were obtained in all three ANOVAs, F(1,15) = 62.07, p < 0.001; F(1,15) = 30.22, p < 0.001; and F(1,15) = 42.52, p < 0.001, for the first, second, and third time window, respectively.

Again, the ANOVAs of the data from the mastoids revealed significant effects of Stimulus, resulting from mastoidal polarity inversions, F(1,15) = 30.07, p < 0.001 (mean peak amplitude difference at the left mastoid 1.338 μV, mean peak amplitude difference at the right mastoid 1.327 μV); F(1,15) = 25.89, p < 0.001 (mean peak amplitude difference at the left mastoid 0.814 μV, mean peak amplitude difference at the right mastoid 1.146 μV); and F(1,15) = 70.05, p < 0.001 (mean peak amplitude difference at the left mastoid 1.157 μV, mean peak amplitude difference at the right mastoid 1.179 μV), for the first, second, and third time window, respectively.
Fig. 2. Deviant and standard and difference ERP waveforms. Grand-averaged ERPs averaged separately for the deviant (grey lines) and standard stimuli (black) of the low contrast condition in Experiment 1A. Deviant-minus-standard difference waveforms are also presented (solid small lines). The grey bars mark the statistically analyzed time windows of the first effect (160–200 ms; grey bar), the second effect (352–392 ms; dark grey bar with a black border), and the third effect (542–582 ms; light grey bar with a black border). Scales are in s and μV.

3.2. Experiment 1B

For the high-contrast condition (Alpha 1 and Alpha 10), 67.69% of the epochs were rejected prior to ERP computation (67.93% of the standard epochs and 65.75% of the deviant epochs; range 27.5% to 93.95%) due to amplitude changes exceeding 80 μV during an epoch. Two consecutive negative deflections were observed for the high-contrast condition (Fig. 3).

At electrode site F4, the peak latency of the first negative deflections was 240 ms, and for the second negative deflection the peak latency was 538 ms, with mean peak amplitude differences (deviant – standard) of −3.221 μV and −2.607 μV, respectively. The ANOVAs of the ERP data revealed significant main effects for Stimulus in the first time window, F(1,15) = 39.03, p < 0.001, and also in the second time window, F(1,15) = 15.52, p = 0.001.

The ANOVAs of the data from the mastoids revealed a significant main effect of Stimulus for the first time window, resulting from mastoidal polarity inversion, F(1,15) = 14.83, p = 0.002 (mean peak amplitude difference at the left mastoid 0.982 μV, mean peak amplitude difference at the right mastoid 0.968 μV), but not for the second time window, F(1,15) = 2.46, p = 0.137 (mean peak amplitude difference at the left mastoid 0.430 μV, mean peak amplitude difference at the right mastoid 0.321 μV).

For the low-contrast condition (Alpha 3 and Alpha 7), 66.61% of the epochs were rejected prior to ERP computation (66.8% of the standard epochs and 65.04% of the deviant epochs; range 20.45% to 87.12%) due to amplitude changes exceeding 80 μV during an epoch. Again, two consecutive negative deflections were observed (Fig. 4).

At electrode site Fz, the peak latencies of the two negative deflections were 258 ms and 558 ms, with mean peak amplitude differences (deviant – standard) of −2.162 μV and −2.329 μV, respectively. The ANOVAs of the ERP data revealed significant main effects for Stimulus in both time windows, F(1,15) = 25.68, p < 0.001, and F(1,15) = 21.80, p < 0.001 for the first and second time window, respectively.

The ANOVAs of data from the mastoids revealed significant main effects of Stimulus for both time windows, resulting from mastoidal polarity inversion, F(1,15) = 6.87, p = 0.019 (mean peak amplitude difference at the left mastoid 0.422 μV, mean peak amplitude difference at the right mastoid 0.603 μV), and F(1,15) = 12.95, p = 0.003 (mean peak amplitude difference at the left mastoid 0.567 μV, mean peak amplitude at the right mastoid 1.003 μV), for the first and second time windows, respectively.

4 Excluding the data from the participant with the highest rejection rate from the analyses did not substantially change the pattern of results.
4. Discussion

The results of the current experiments replicated and extended previous findings of MMN elicitation by room-acoustics-related changes in unattended stimuli [8]. Deviant stimuli were consistently associated with a pronounced negative deflection at frontal electrode sites, starting earlier than 150 ms after stimulus onset. To match standards and deviants in terms of their left-right intensity distribution, thus ruling out an account of our results in terms of overall interaural intensity differences, we used a symmetrical room for the simulation of the room acoustical changes. MMNs occurred not only in the high-contrast condition (i.e., Alpha 1 vs. Alpha 10) but also in the more subtle low-contrast condition (i.e., Alpha 3 vs. Alpha 7).

Different rooms have differing reflection characteristics and therefore produce differing reverberations. Variations of point-by-point intensity between standards and deviants, each within a single auditory object, may indeed lead to an MMN in oddball protocols if the intensity variation over time is detectable by the auditory system. It seems unlikely, however, that the MMN observed in the current study was driven by the variation of point-by-point intensity over time resulting from the room acoustics manipulation. If this were the case, we would expect a larger amplitude in the first negative deflection for the high-contrast condition than in the first negative deflection for the low-contrast condition, because a larger and thus more salient intensity difference around stimulus onset should affect both the MMN and the N1 [16,17]. At least in Experiment 1A (but not in Experiment 1B), however, the opposite pattern was observed, which argues against the notion of detection of point-by-point intensity differences. Rather, it appears that despite the more salient difference in terms of sound amplitude in the high-contrast condition, the room acoustics were more salient in the low-contrast condition. Overall, there is no clear pattern of a larger MMN in the high-contrast condition. Therefore, it seems reasonable to conclude that the MMNs in the present study did not hinge upon interaural intensity differences, neither overall, nor point by point. Instead, differing room acoustics stand as the elicitor. For Experiment 1A, another potential caveat hinges on the duration differences between the standard and the deviant. These could, in principle, elicit MMNs. As has already been mentioned, the first and second sounds of the sequences (triplets) differ in duration between Alpha 1 and Alpha 10 and between Alpha 3 and Alpha 7. In our view, however, it is highly unlikely that the MMNs observed in Experiment 1A resulted from differences in duration, because the MMN latencies do not fit. The MMN peak latencies were too early to be compatible with elicitation by duration differences. Please note that the potential duration confound was removed in Experiment 1B, which revealed comparable initial MMNs. Therefore, room acoustics again stand as the elicitor.

Replicating previous findings of Frey et al. [8], additional negative deflections occurred subsequently to the initial MMN triggered by the deviant stimulus. Contrasting with the study of Frey et al. [8], in which sequences of heterogeneous piano chords were used as stimuli, the current study used triplets of sounds (Experiment 1A) or a single sound (Experiment 1B) as standards and deviants. These modifications did not prevent the occurrence of later nega-
tive deflections following the first MMN. Specifically, in Experiment 1A, two additional negative deflections were observed, peaking at 372 ms and 562 ms and at 440 ms and 558 ms in the low-contrast condition and the high-contrast condition, respectively. Each of these negative deflections was accompanied by a significant polarity inversion at the mastoid electrodes, which is typically observed for MMNs. These multiple MMN responses could simply reflect separate detection of deviance for the three sound segments of the triplets. The temporal correspondence of the negative deflections and the onsets of the constituent sounds of the triplets are in line with the assumption that the deflections reflect consecutive instances of mismatch detection, related to each individual sound of the deviant triplet. On the other hand, neither complexity nor “novelty” (i.e., deviance from the directly preceding sound segment) constitute necessary conditions for multiple MMNs to occur. Given that the acoustic characteristics of the second and third sound segment were completely predictable due to the identity relation with the first sound of the triplet, the occurrence of sound-related multiple MMNs suggests that separable expectations were formed for the constituent elements of the repeatedly presented sequence and were not adjusted after detection of the first sound of a deviant sequence (for a discussion of the relation of successive MMNs to predictable consecutive deviations see [18]).

The results of Experiment 1B demonstrate, however, that the occurrence of a second negative deflection does not depend on the presentation of successive sounds. That is, even though only a single sound was used, an additional negative deflection occurred at frontal and central electrode sites, peaking at 538 ms and 558 ms after stimulus onset in the high-contrast and low-contrast condition, respectively. Again, these negative deflections were accompanied by MMN-typical polarity inversions at the mastoid electrodes (albeit only significant in the low-contrast condition). These findings are in line with the assumption that acoustic irregularities associated with a complex pattern of changes in the stimulus may elicit two successive processes of deviance detection. Potentially, this involves a fast process of detecting one or more low-level featural changes and a slower process of identifying the meaning of the specific irregular pattern of featural changes. In the case of the deviant stimuli made up of successive separated sounds, the occurrence of a sound-onset-related additional MMN may be superimposed on the second detection response.

Korpilahti et al. [19] described a late-latency waveform following an MMN, arguing that it reflects the automatic processing of complex auditory stimuli. This late-latency Mismatch Negativity, sometimes called a late differentiating negativity (LDN), peaks at about 400–450 ms [20]. Most LDN research has focussed on speech stimuli [21], but there have also been LDN findings in non-speech studies (for an overview of LDN, see [22]). The findings regarding LDN responses are less robust compared to those for MMN responses. Developmentally oriented LDN research has also revealed much greater maturation effects. In our view, taking these findings together, no clear pattern for interpreting the present findings with respect to the LDN emerges. It remains an open question why two further negative deflections were found in Experiment 1A and one in Experiment 1B, rather than only a single MMN as found in other comparable studies [23].
senting further processing of the detected change in the acoustic input come close to our hypotheses [24]. All in all, the functional significance of the additional negative deflections remains to be determined.

For a full-fledged account of the aspects of room-acoustics-mediated irregularity that are critical for eliciting an MMN response, as well as the boundary conditions thereof, more research is needed. Given the diversity and complexity of room-acoustics-related manipulations, this should turn out to be a tedious but important endeavour. The results of the current study show that even in the absence of differences in primary auditory features such as overall loudness or the left-right intensity distribution of the stimulus, unattended acoustic irregularities mediated by room acoustics are detected.

Acknowledgements

The authors thank the students in the lab for their help with data acquisition and Merrie Bergman for proof reading. The valuable comments of two anonymous reviewers are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neulet.2016.11.063.

References

Eidesstattliche Versicherung


Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum 15.1.2017

Unterschrift des Doktoranden