

---

# **SENSORY AUDITORY PROCESSING AND INTUITIVE SOUND DETECTION**

An Investigation of Musical Experts and Nonexperts

---

Titia van Zuijen

Cognitive Brain Research Unit

University of Helsinki

Finland

*DOCTORAL DISSERTATION*

To be presented, by due permission of the Faculty of Behavioral Sciences of the University of Helsinki, for public examination, in auditorium XII, Unioninkatu 34, 11<sup>th</sup> of October, 2006, 10 a.m.



UNIVERSITY OF HELSINKI

CBRU

ISBN: 952-10-3328-2 (paperback)

ISBN: 952-10-3329-0 (PDF) (<http://ethesis.helsinki.fi>)

ISSN: 0781-8254 (Research Reports no. 36)

Yliopistopaino

Helsinki 2006

Supervisors

**Mari Tervaniemi, PhD**

Cognitive Brain Research Unit  
University of Helsinki, Finland  
& Helsinki Brain Research Centre

Academy Professor **Risto Näätänen**

Cognitive Brain Research Unit  
University of Helsinki, Finland  
& Helsinki Brain Research Centre

Reviewers

Professor **Terence W. Picton**

Rotman Research Institute  
Baycrest Centre for Geriatric Care  
Toronto, Canada

Professor **Petri Toiviainen**

Department of Music  
University of Jyväskylä, Finland

Opponent

Professor **Thomas F. Münte**

Department of Psychology  
Magdeburg, Germany

## LIST OF PUBLICATIONS

This thesis is based on the following publications:

**Study I:** van Zuijen, T.L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience*, **16**, 331–338.

**Study II:** Winkler, I, van Zuijen, T.L., Sussman, E, Horváth, J., & Näätänen, R. (2006). Object representation in the human auditory system. *European Journal of Neuroscience*, **24**, 625–634.

**Study III:** van Zuijen, T.L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity—a mismatch negativity study comparing musicians and nonmusicians. *Cognitive Brain Research*, **23**, 270–276.

**Study IV:** Vuust, P., Pallesen K.J., Bailey, C., van Zuijen T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter. Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, **24**, 560–564.

**Study V:** van Zuijen, T.L., Simoens, V.L., Paavilainen, P., Näätänen, R., & Tervaniemi, M. (2006). Implicit, intuitive and explicit knowledge of abstract regularities in a sound sequence: an event-related brain potential study. *Journal of Cognitive Neuroscience*, **18**, 1292–1303.

## **ABSTRACT**

The auditory system can detect occasional changes (deviants) in acoustic regularities without the need for subjects to focus their attention on the sound material. Deviant detection is reflected in the elicitation of the mismatch negativity component (MMN) of the event-related potentials. In the studies presented in this thesis, the MMN is used to investigate the auditory abilities for detecting similarities and regularities in sound streams. To investigate the limits of these processes, professional musicians have been tested in some of the studies. The results show that auditory grouping is already more advanced in musicians than in nonmusicians and that the auditory system of musicians can, unlike that of nonmusicians, detect a numerical regularity of always four tones in a series. These results suggest that sensory auditory processing in musicians is not only a fine tuning of universal abilities, but is also qualitatively more advanced than in nonmusicians. In addition, the relationship between the auditory change-detection function and perception is examined. It is shown that, contrary to the generally accepted view, MMN elicitation does not necessarily correlate with perception. The outcome of the auditory change-detection function can be implicit and the implicit knowledge of the sound structure can, after training, be utilized for behaviorally correct intuitive sound detection. These results illustrate the automatic character of the sensory change detection function.

## ACKNOWLEDGEMENTS

The work presented in this thesis was conducted at the Cognitive Brain Research Unit at the Psychology Department of the University of Helsinki. The financial support has been provided by the Academy of Finland and the Finnish Graduate School of Psychology. The thesis was reviewed by Professor Terence Picton and Professor Petri Toiviainen. I would like to thank you both, I highly valued your comments and suggestions. I also greatly appreciate that Professor Thomas Münte has agreed to be my opponent in the forthcoming public defense.

Many other people have also contributed to this work in very different ways. First of all, I would like to thank my supervisors, Mari Tervaniemi and Risto Näätänen for providing me with both guidance and the freedom to develop my own ideas and experiments. I would also like to thank my co-authors for the insightful scientific discussions that boosted my motivation when that was mostly needed, first of all Elyse Sussman and István Winkler and in a later stage, Veerle Simoens and Petri Paavilainen. I would also like to thank my fellow graduate students at the CBRU and all others colleagues including the supporting staff. The whole ‘living and working in Finland’ experience would not have been nearly as special without all of you. Also my ‘Finnish’ friends contributed to that: Arzu, Freya, Iris, Marieke and Satu. Arzu and Marieke, not only do I want to thank you both for your friendship but also for your practical help and for generously hosting me when I was flying from there to here. Tamara and Serge, thank you too for hosting me when I was flying from here to there (and certainly also for all those other things that are beyond the scope of this thesis). Also thanks to my other Amsterdam friends and Astrid ☺ for their support of my life up north, although they sometimes looked somewhat doubtful.

I want to acknowledge all my subjects for being my ‘guinea pigs’. I especially appreciated the musicians for being so eager, and for finding the time to come to the lab.

Thanks to my parents, and Aase and Knud-Erik for looking after Tigo while I was writing, measuring or traveling. Thanks to Klaus for his support, his patience, and not at least for his (to me) original point of views. Og søde Tigo...knus!

~ Kiitos kaikille! ~



## CONTENTS

<b>List of publications</b> .....	4
<b>Abstract</b> .....	5
<b>Acknowledgements</b> .....	6
<b>1 Investigating sound perception and sensory processing</b> .....	10
1.1 Outline .....	10
1.2 Auditory processing .....	11
1.2.1 Auditory pathways .....	11
1.2.2 Auditory cortex .....	13
1.3 Auditory event-related potentials/fields .....	13
1.3.1 From the cochlea to the auditory cortex .....	14
1.3.2 N1 .....	15
1.3.3 Mismatch negativity .....	16
1.3.3.1 Change detection .....	16
1.3.3.2 Sensory memory .....	17
1.3.3.3 Automatic? .....	18
1.3.3.4 MMN and perception .....	19
1.3.3.5 The ‘abstract’ MMN .....	20
1.3.3.6 Musical expertise and the MMN .....	20
1.3.3.7 Functional relevance .....	21
1.3.4 P3a .....	22
1.3.5 N2b and P3 .....	22
<b>2 Gestalt theory and auditory objects</b> .....	23
2.1.1 Automatic and universal .....	23
2.1.2.1 <i>Study I</i> aims and experiment .....	24
2.1.2.2 <i>Study I</i> results and conclusions .....	25
2.1.3 Why are musicians better at automatic processing? .....	26



2. 2.1	The principle of exclusive allocation .....	26
2. 2.2.1	<i>Study II</i> aims and experiment .....	27
2. 2.2.2	<i>Study II</i> results and conclusions .....	28
2. 2.3	The role of attention in an ambiguous situation .....	28
<b>3</b>	<b>Number processing</b> .....	30
3. 1	Sensory intelligence .....	30
3. 2	Counting the beat in music .....	30
3. 3.1	<i>Study III</i> aims and experiment .....	30
3. 3.2	<i>Study III</i> results and conclusions .....	31
3. 4	Auditory number detection in musicians .....	32
<b>4</b>	<b>Hemispheric asymmetry</b> .....	33
4. 1	Speech perception .....	33
4. 2.1	<i>Study IV</i> aims and experiment .....	33
4. 2.2	<i>Study IV</i> results and conclusions .....	34
4. 3	Left lateralized rhythm processing in jazz musicians .....	34
<b>5</b>	<b>Awareness and auditory change detection</b> .....	36
5. 1	MMN elicitation and perception .....	36
5. 2	Implicit knowledge and intuition .....	36
5. 3.1	<i>Study V</i> aims and experiment .....	38
5. 3.2	<i>Study V</i> results and conclusions .....	38
5. 4	Implicit change detection and intuition .....	40
<b>6</b>	<b>Conclusions</b> .....	41
	<b>References</b> .....	44
	<b>Appendix: Methods</b> .....	53
	<b>Original publications</b> .....	57

# 1 INVESTIGATING SOUND PERCEPTION AND SENSORY PROCESSING

## 1. 1 Outline

The auditory system analyses the acoustic input to make sense out of the sound surrounding us. The auditory processes in the brain help us to perceive meaning so that we can communicate through speech, move safely through the traffic, or enjoy music. Auditory perceptual capabilities are not static, they can be shaped by exposure and practice. The auditory skills of musicians are under extreme demands, a violinist can hear much smaller pitch differences than someone with an untrained ear. Musicians are, therefore, an interesting group to study when investigating the limits of auditory processing capabilities. An interesting aspect of the auditory functions is that part of the auditory processes take place automatically, we are not all the time aware of all the sounds surrounding us. Sometimes we do not deliberately attend to a sound source until it suddenly changes. At other times we might not be fully aware of a sound but does that mean that it cannot influence our actions?

The Gestalt theory says that in order to create meaning from the world surrounding us, stimulus features that are similar to each other are perceptually grouped together while segregated from relatively dissimilar features. In this theory, it is assumed that one feature cannot belong to two different objects at the same time. The latter is called the principle of exclusive allocation. The theory also proposes that the forming of these perceptual organizations is automatic and universal. In *Study I*, the universality of automatic auditory grouping is addressed by comparing sensory grouping abilities of musicians and nonmusicians. In *Study II* the principle of exclusive allocation is studied in an ambiguous auditory scene.

How advanced are automatic auditory functions? Can something as conceptually abstract as detecting the number of tones in a series take place outside of the focus of attention? Since perceiving the number of beats in a measure is an important aspect of the perceptual organization of music, this question was addressed in *Study III* by examining the auditory processing skills of musicians.

The meaning or message conveyed in acoustic information depends as much on the nature of the sound as on the mindset of the perceiver. Music may sometimes convey language-like messages (although less concrete) that are, in particular, picked up by the musical listener. An example is the way rhythm is used by improvising jazz musicians. In *Study IV* it was addressed whether rhythm is processed in the language-dominant hemisphere in the brains of jazz musicians.

In the afore-mentioned studies, the aim was to study automatic auditory processes by utilizing the ability of the auditory system to detect occasional changes in a repetitive/regular sound stream without the requirement of focused attention. Based on a consensus in the literature, the assumption was that there is a direct correlation between this form of change detection and perception. However, the stimulus sequences that were used in the studies presented in this thesis became more complex and the question arose whether this assumption was still valid. In *Study V* it was investigated whether automatic change detection can be implicit and whether implicit knowledge can, after training, be utilized for intuitive sound detection.

## **1. 2 Auditory processing**

### **1. 2.1 Auditory pathways**

Sound is air set into vibrating motion. When someone hits a drum the membrane of the drum starts to vibrate. The movement of the membrane compresses the air next to it creating an area of increased pressure followed by an area of reduced pressure. This vibrating motion repeats itself in a regular period of time (hundreds or thousands times per second) creating a sound wave (Rossing, 1990).

When a sound wave reaches our ear, the eardrum starts vibrating in the same frequency as the sound. The eardrum is connected through three small bones to the inner ear. These three bones amplify and transmit the vibrations of the sound to the fluid in the cochlea of the inner ear, creating a compression wave in the fluid. The cochlea is a snail-shaped organ containing hair-like receptor cells lined along the basilar membrane. The basilar membrane moves when the compression wave travels through the fluid in the cochlea. The width and resiliency of the membrane vary along the cochlea so that each portion of the membrane moves in response to a characteristic frequency inducing an

electrical impulse in the hair cells. This tonotopic organization of the responses (for high frequencies) as well as the periodicity of the hair-cells discharge rate (for low frequencies) decomposes the sound signal into neural codes for separate frequencies. Intensity information is also converted into a neural code by the hair-cell discharge rate (for high frequencies) or by the number of successive responses of cells further away from the characteristic frequency (for low frequencies) (Biacabe, 2001; Kandel *et al.*, 2000).

The electrical signal travels from the hair cells along the 8<sup>th</sup> cranial nerve to the cochlear nuclei in the brain stem (see Figure 1). The cochlear nuclei contain functionally distinct cells types such as frequency-specific cells but also cells specifically responding to sound onsets. One level higher, in the superior olivary complex, left- and right-ear auditory fibers intersect. Sound onset times are then used for sound localization as certain types of neurons only respond to sounds with specific interaural timing differences (crucial for the localization of low sounds) and other neurons respond only to specific interaural intensity differences (for the localization of high sounds).

The auditory tracts continue via the inferior colliculus in the midbrain to the medial geniculate nucleus in the thalamus. The ascending auditory pathway terminates in the primary auditory cortex.

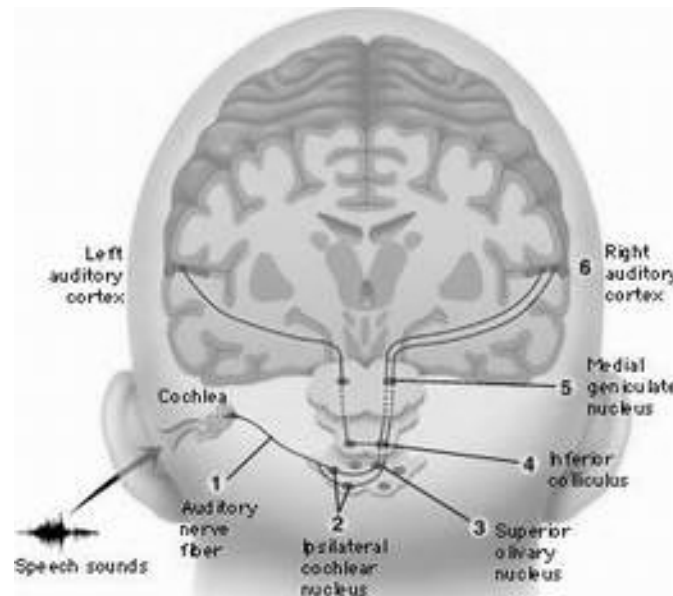


Figure 1. Auditory pathways

### **1. 2.2 Auditory cortex**

The auditory cortex is located at the dorsal surfaces of the temporal lobes within the lateral sulcus. It consists of functionally different areas. The primary auditory area is located in the medial two-thirds of the first transverse gyrus of Heschl (Hackett *et al.*, 2001). The secondary areas are more laterally positioned. The primary and secondary auditory areas are tonotopically organized but with the cells in the secondary areas tuned to spectrally more complex sounds (Palmer & Summerfield, 2002). There are connections from Heschl's gyrus to auditory association areas in the temporal lobes (planum temporale) and parietal lobes. These areas are involved in higher-order processes. A portion of the right parietal cortex is, for instance, activated by sound motion (Griffiths *et al.*, 1998), the superior temporal gyrus is involved in auditory short-term memory (Colombo *et al.*, 1990), and the most anterior part of the superior temporal gyrus is activated by melodies (Schmithorst & Holland, 2003). In musicians, certain areas of the auditory cortex are morphologically different: the anteriomedial portion of Heschl's gyrus is larger in musicians than in nonmusicians (Schneider *et al.*, 2002), and the planum temporale is asymmetric in musicians with perfect pitch ( Schlaug *et al.*, 1995; Keenan *et al.*, 2001).

### **1. 3 Auditory event-related potentials/fields**

A non-invasive technique for investigating cortical brain function is electroencephalography (EEG). EEG is a measure of the electric potential differences on the scalp as a function of time. It is generally assumed that the origin of the cortical EEG is synchronized post-synaptic activity of pyramidal apical dendrites. When enough (~10000) neurons are simultaneously active, the extra-cellular current flow can be measured on the scalp through volume conduction (Hämäläinen *et al.*, 1993). The neural signal is, however, small and contaminated with environmental and instrumental noise. The event-related potentials (ERPs) can be calculated by averaging periods of signal time-locked to the onset of a repetitive event, e.g., a stimulus. Noise is then averaged out and reflections of the neural processing triggered by the event under interest can become visible. Both amplitudes and latencies of the voltage fluctuations in the ERP can give information on the evoked neural processes.

A technique related to EEG is magnetoencephalography (MEG). Subjects are seated with their head in a helmet containing super-conducting sensors that measure the magnetic field (Hämäläinen *et al.*, 1993). The neural current sources generating magnetic fields that can be picked up by the sensors outside of the head are assumed to be from the intracellular-currents of apical dendrites. Magnetic fields generated by neural sources radial (the head is commonly approximated by a sphere model) to tissue boundaries (brain matter, skull, and scalp) are not visible. Only magnetic fields from tangential sources are picked up by the sensors and pass through the different tissues without hindrance from conduction boundaries. Event-related fields (ERFs) and ERPs show similarities in their evoked components partially depending on whether source orientations are tangential or not. The magnetic counterpart of an ERP is usually denoted with the same name but with an ‘m’ added. The cortical location of the neural sources generating ERP and ERF components can be estimated with iterative inverse dipole modeling.

### **1. 3.1 From the cochlea to the auditory cortex**

Electric reflections of auditory processing can also be measured before the neural signal reaches the cortex. Brainstem auditory evoked potentials (BAEPs) reflect synchronized cell activity in the brainstem (Biacabe, 2001) and are elicited within 1–10 ms after stimulus presentation. The origin of these components is somewhat uncertain but the first five of these responses are believed to be generated by ganglion cells in the cochlea (component I & II), the cochlear nucleus (component III), and the superior olivary complex (component IV & V) (Markand, 1994; Shaw, 1995). Already at this early level of auditory processing, sound qualities such as intensity (Markand, 1994), frequency (Stapells & Oates, 1997), and location (Riedel & Kollmeier, 2002) are reflected in the evoked potentials. Riedel and Kollmeier (2002) showed for instance that the amplitude of component V is larger for central sounds than for lateral sounds.

Middle-latency auditory evoked responses (MAEPs) are elicited between 10 and 80 ms. The MAEP consist of several components including the Na, Pa, Nb, and Pb or P50. The Na is elicited at around 19 ms and there is some evidence that it is the first evoked response from the primary auditory cortex (Rupp *et al.*, 2002; Shaw, 1995). The

Na is followed by the Pa (30 ms) and it originates more certainly from the primary auditory cortex (Liégeois-Chauvel *et al.*, 1994). The Pb/P50 and later components (elicited at around 60–75 ms) might come from a more lateral area in the secondary auditory cortex (Liégeois-Chauvel *et al.*, 1994; Yvert *et al.*, 2001). Stimulus properties are reflected in the MAEPs (e.g., Pantev *et al.*, 1995), but also sleep-related arousal changes can affect MAEPs (Erwin & Buchwald, 1986). In addition, there is a positive correlation between tone-discrimination aptitude in musicians, the size of the primary auditory cortex and amplitude of the Na-Pa complex (Schneider *et al.*, 2002) indicating that the functional significance of sound can be reflected in the neural processing after approximately 20 ms.

### **1.3.2 N1**

ERP components following the MAEP are called long-latency components. The first most prominent long-latency component is the N1 elicited around 100 ms. There are several sources in the supratemporal cortex that contribute to the N1 (Liégeois-Chauvel *et al.*, 1994; Näätänen & Picton, 1987) including tonotopically organized cortical areas (Cansino *et al.*, 1994). The N1 amplitude attenuates after repeated stimulation with the same stimulus features (Näätänen & Picton, 1987). The N1 amplitude also reflects feature specificity: the N1 amplitude increases when the intensity of a stimulus increases (Bak *et al.*, 1985) and decreases with an increasing tonal frequency (Näätänen & Picton, 1987). The N1 amplitude is, furthermore, modulated by attention (Woldorff *et al.*, 1993) and affected by task relevance (Jerger *et al.*, 1992).

The N1 amplitude can reflect musical expertise; it is larger to musical tones in musicians compared to nonmusicians (Pantev *et al.*, 1998; Shahin *et al.*, 2003; see however, Lütkenhöner *et al.*, 2005). Moreover, the N1 reflects functional changes (Pantev & Lütkenhöner, 2000); the N1 amplitude becomes larger after sounds have been used in a discrimination training (Menning *et al.*, 2000; Tremblay *et al.*, 2001) even though mere exposure reduces the N1 amplitude (Brattico *et al.*, 2003).

### 1. 3.3 Mismatch negativity

The ERP component central in the research presented in this thesis is the mismatch negativity (MMN) (for reviews, see e.g., Näätänen, 1992; Ritter *et al.*, 1995; Picton *et al.*, 2000). The MMN is measured on the scalp in the range of 120 to 300 ms post-stimulus and is, depending on stimulus parameters, sometimes partially overlapping and extending the N1 and sometimes following the N1. The MMN was found to be a separate ERP component by Näätänen and colleagues in 1978 (Näätänen *et al.*, 1978; see also, Näätänen *et al.*, 2005). The MMN is frontally negative, with declining amplitudes towards posterior sites and sometimes accompanied by a positive deflection on the mastoid electrodes (see, Figure 2). EEG and MEG studies have localized the main MMN sources in the supratemporal cortex about 1 cm anterior of the N1m generator (Alho, 1995; Hari *et al.*, 1984). Combined EEG/fMRI studies show that activity from the inferior frontal cortex can also contribute to the MMN (Doeller *et al.*, 2003; Rinne *et al.*, 2005). Intracranial recordings (in presurgical patients) have found MMN generators in secondary auditory areas (Halgren *et al.*, 1995), in the auditory association cortex (Kropotov *et al.*, 2000), and in the inferior frontal cortex (Rosburg *et al.*, 2005). Moreover, the exact location of the MMN generator varies depending on stimulus features and on stimulus complexity (Alho, 1995; Giard *et al.*, 1995; Molholm *et al.*, 2005).

#### 1. 3.3.1 Change detection

Unlike the N1 (and other earlier ERP components), the MMN is not elicited by sound-onset *per se* but by deviant sounds occasionally replacing frequently presented standard sounds in an oddball paradigm. The MMN amplitude is determined by the magnitude of the stimulus change: larger differences between standards and deviants give rise to larger MMN amplitudes (e.g., Novitski *et al.*, 2004; Tiitinen *et al.*, 1994). Stimulus energy or stimulus features *per se* are not reflected in the MMN as exemplified by the following situations. Irrespective of whether a repetitive sound is occasionally changed from loud to soft, or from soft to loud the MMN amplitude reflects the magnitude of change (Näätänen *et al.*, 1987). Also an occasionally early onset in a regularly paced sound stream can elicit an MMN indicating the magnitude of change (the earlier the onset, the larger the violation of the regular interval and the larger the MMN). Remarkably, even the absence



of a sound, at least in fast paced regular sequences, can elicit an MMN (Yabe *et al.*, 1997).

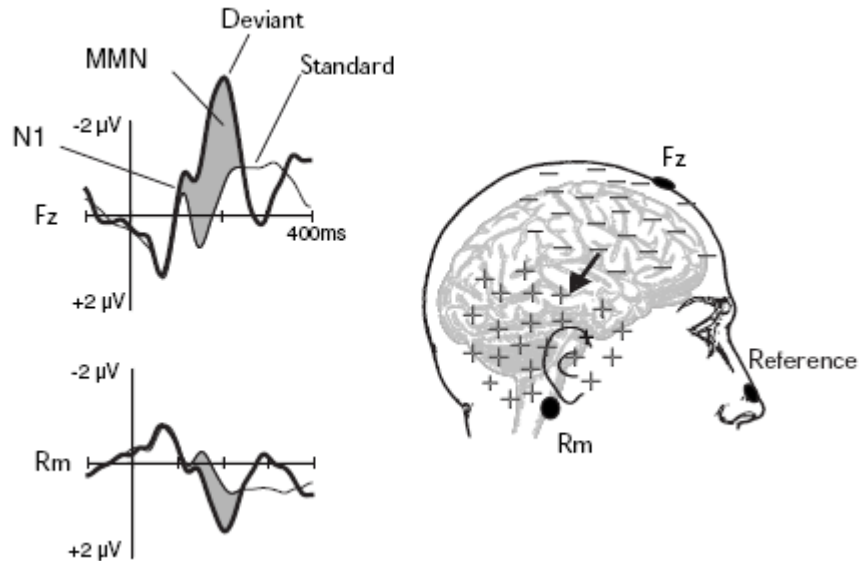


Figure 2. MMN generation. The MMN measured from a frontal electrode (Fz), and from an electrode at the right mastoid (Rm) referenced to a noise electrode. Note that negativity is plotted upwards. The grey area in the waveforms is the difference between the deviant wave and the standard wave. The MMN is generated in the auditory cortex below the arrow.

### 1. 3.3.2 Sensory memory

In order to detect change, the auditory system must encode and maintain a model of the regular features in the acoustic environment and compare the present with the past. The auditory change-detection mechanism underlying the MMN elicitation should thus involve a form of sensory memory (Näätänen & Winkler, 1999). Accordingly, the MMN exhibits behavior that is expected from a memory system. It takes a few presentations (~3) of the standard sound before a deviant will elicit an MMN indicating that the memory representation has to build up (Cowan *et al.*, 1993). The more presentations precede a deviant the stronger the memory becomes and the larger the MMN amplitude will be (Imada *et al.*, 1993; Sams *et al.*, 1983). The memory representation decays if it is not reinforced with standard-sound presentations. The memory-decay function has been

estimated to last up to 10 seconds, after which a deviant does not elicit an MMN anymore (Sams *et al.*, 1993). The next question then is: what can be stored in the sensory memory underlying the MMN elicitation? It has been shown that it can contain any type of audible acoustic repetition/regularity (within certain duration limits, see below). First of all, single features (e.g., intensity, location or pitch) can be stored but also feature conjunctions (Takegata *et al.*, 1999), temporal structure, speech (phonemes, syllables, and words; Näätänen *et al.*, 1997; Näätänen, 2001; Shtyrov & Pulvermüller, 2002), and short melodies (Tervaniemi *et al.*, 2001). To take the melody example, if five tones are repeatedly presented with the same melodic contour, an MMN can be elicited by a tone changing the melodic contour. The regular melody contour is encoded in an auditory memory template and when a deviation from the melody contour not matching the template is detected, an MMN is elicited (Tervaniemi *et al.*, 2001).

A major constraint in what can, and cannot elicit an MMN, is the duration of the standard sound or regularity that can be presented. Estimating from unpublished observations and the chunk durations generally reported in the literature, the maximum is about 1 s in the average person (see also, Grimm *et al.*, 2004; Näätänen *et al.*, 2004).

### **1. 3.3.3 Automatic?**

The MMN is elicited whether or not subjects are focusing their attention on the stimuli. Deviants do not need to be detected or to be task relevant for the MMN to be elicited. It has even been proposed that the MMN is elicited fully independent of attention (Näätänen, 1990). This would implicate that the MMN amplitude should not under any circumstances be modulated by attention; not enhanced by strong attention and not diminished/abolished by a complete absence of attention. Woldorff *et al.*, (1991) presented evidence against this claim. A small intensity deviant presented in one ear disappears when the attention is strongly focused on detecting small intensity deviations in the other ear. More recently, this has also been found for small frequency deviants (Sussman *et al.*, 2003a). Sussman and colleagues explained these effects as deviance-competition effects. When in each ear deviants are presented that violate the same feature, and attention is directed to one of the ears, deviants delivered to that ear elicit an MMN while the MMN to the deviants in the unattended ear is diminished or even

abolished. The MMN elicitation is thus not completely independent of attention. In addition, the MMN amplitude can be enhanced when attention is focused on the deviants (Woldorff *et al.*, 1998). Note, however, that in attended conditions the MMN cannot always be dissociated from the N2b, an ERP component overlapping the MMN (Novak *et al.*, 1990; Näätänen *et al.*, 1982).

The reason that the MMN is still used to probe automatic aspects of auditory processing is because when competition is not so strong, such as in the case of a slow stimulus pace (Näätänen *et al.*, 1978), or when deviants violate different acoustic features in different ears, attention modulation is not seen. It is, furthermore, elicited in a wide range of situations in which other ERP components requiring focused attention are not elicited (e.g., P3, see below). The MMN is elicited in coma patients (Fischer *et al.*, 1999), certain sleep stages (Loewy *et al.*, 1996; Sallinen *et al.*, 1994), and under many different attentional conditions; subjects can be passively distracted from the auditory stimuli by watching a movie or they can simultaneously perform a visual or an unrelated auditory task (Alho *et al.*, 1992; Alho *et al.*, 1994; Alho & Sinervo, 1997; Näätänen *et al.*, 1978; Paavilainen *et al.*, 1993). Most importantly, deviants do not need to be task relevant, so that the ability/willingness to respond (e.g., in children or certain patient groups), cognitive factors (e.g., motivation or performance strategies), do not need to be taken into account in the experimental setup.

### **1. 3.3.4 MMN and perception**

It is well-known that MMN elicitation correlates with the perception of stimulus deviance. This has been shown in several studies testing discriminative abilities in a behavioral session subsequent to the unattended MMN recording (Amenedo & Escera, 2000; Jaramillo *et al.*, 2000; Tervaniemi *et al.*, 2001; Tiitinen *et al.*, 1994). A larger discrepancy between deviants and standards is reflected in larger MMN amplitudes, a shorter MMN latency, as well as higher hit rates and shorter RTs. On the contrary, when the difference between deviants and standards becomes too small for perceptual discrimination the MMN disappears (Sams *et al.*, 1985; see however, Allen *et al.*, 2000). From these studies, it became evident that even though the MMN can be elicited when stimuli are not deliberately attended to, if subjects do subsequently attend to the sound

material, deviants are detected. In other words, the outcome of the auditory change-detection process reaches awareness. In chapter 5 this view is challenged.

### 1. 3.3.5 The ‘abstract’ MMN

An important aspect of the sensory change detection function is the extraction of regularity from the acoustic environment. Relatively advanced is the ability to detect so-called ‘abstract’ regularities (for a review see, Näätänen *et al.*, 2001). An abstract regularity is a relative invariance between sounds/sound features (see, Figure 3). Saarinen *et al.*, (1992) showed that an MMN could be elicited by violating a regular interval within tone pairs that were, as couples, jumping around over a wide range of frequencies. The standard tone-pairs were ascending and the deviants were descending. The fact that the tone pairs were roving and that there were many physically different tone pairs, made the regularity violation ‘abstract’ not the relationship violation as such (see also, Paavilainen *et al.*, 1999; 2003). Also third-order violations of abstract feature conjunctions can elicit an MMN (Paavilainen *et al.*, 2001). In the Paavilainen *et al.* (2001) study, the standard stimuli varied randomly over a large range, both in frequency and intensity, but followed the rule “the higher the frequency, the higher the intensity”. Occasional deviant stimuli following the opposite rule elicited an MMN. These types of advanced abstract regularity processing have been termed ‘sensory intelligence’ (see also, 3. .1).

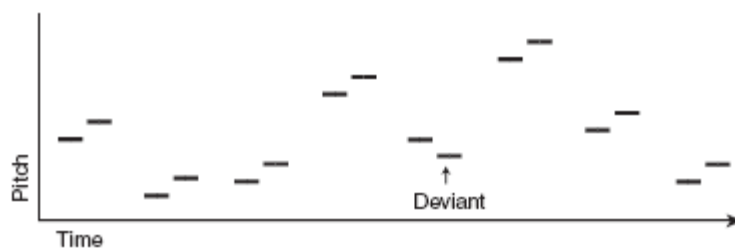


Figure 3. Illustration of an abstract oddball paradigm.

### 1. 3.3.6 Musical expertise and the MMN

Musical expertise in auditory function is reflected in the MMN (Fujioka *et al.*, 2004; 2005; Koelsch *et al.*, 1999; Rüsseler *et al.*, 2001; Tervaniemi *et al.*, 1997; Tervaniemi *et al.*, 2001). The excellent sound-discrimination skills of musicians correlate with the

auditory change-detection function. The MMN is elicited in professional violinists by deviant chords containing a slightly mistuned tone whereas musical novices show an MMN to much larger pitch deviations, only (Koelsch *et al.*, 1999). In addition, the MMN to a musically relevant interval violations is larger in musicians than in nonmusicians (Fujioka *et al.*, 2004). This effect seems specific to musical stimuli because the MMN in a control condition with a simple pure-tone frequency deviant did not differ between the two groups. In chapter 2, 3 and 4 the auditory functions of musicians and nonmusicians are also compared on the basis of their MMN or MMNm responses.

### **1.3.3.7 Functional relevance**

What is the function of an automatic auditory change-detection mechanism in our daily life? A brain mechanism specifically detecting change might have certain advantages; repeated sounds do not contain new information and do not need to be fully processed over and over again, whereas new information might requires more extensive processing (Sinkkonen *et al.*, 1996). That regularity detection in general is beneficial, is supported by psychophysical studies showing benefits of regular contexts compared to irregular contexts in the detection of just noticeable differences (Drake *et al.*, 1993), or short silent gaps (Minzuno *et al.*, 1994).

The automatic change-detection mechanism reflected in the MMN elicitation is not an isolated event. Subsequent to eliciting an MMN, deviants can trigger an involuntary attention switch to draw processing resources to the deviating sound. This has been shown by a delay in the RT, and decrease in response accuracy to target tones presented in one ear, while in the other ear MMN eliciting deviants were presented (Schröger, 1996). Similarly, deviants and novels (large deviations with alarming characteristics, e.g., dog barks or the sound of breaking glass) can delay the behavioral performance in a simultaneous visual task (Escera *et al.* 1998).

It is also interesting to mention that the MMN is elicited by ecological valid sound material. It has been shown that the MMN can be elicited in very natural situations: Occasionally transposed chords, replacing the regular chords endings in naturally expressive music elicit the MMN (Koelsch & Mulder, 2002), but also changes in the sound of a series of footsteps of a person walking (Winkler *et al.*, 2002).

### **1. 3.4 P3a**

The involuntary capturing of attention by salient deviants is associated with a positive ERP component elicited subsequent to the MMN: the P3a (for reviews, see Escera *et al.*, 2000; Friedman *et al.*, 2001) peaking around 230–300ms. The P3a amplitude is sensitive to novelty and it decreases when stimuli become less novel (i.e., the P3a habituates, see Friedman & Simpson, 1994). Whereas the MMN reflects a process that is relatively automatic the P3a is more distinctly affected by attention (Friedman *et al.*, 2001; Sussman *et al.*, 2003b). The MMN and the P3a can thus be viewed to reflect two different processes that operate serially to analyze the auditory input for salient information.

### **1. 3.5 N2b and P3**

In case deviants are attended to or task relevant, additional attention and detection related ERP components are elicited. The first one is the N2b (Näätänen & Picton, 1986; Perrault & Picton, 1984) The N2b overlaps in time with the MMN and these components are sometimes difficult to dissociate. The scalp distribution of the N2b is somewhat more centrally negative than the MMN and the N2b does not reverse polarity at the mastoids (Novak *et al.*, 1990). N2b elicitation might reflect stimulus awareness and/or identification.

The P3 is also elicited by deviants when stimuli are attended (Hermann & Knight, 2001; Picton, 1992; Sutton *et al.*, 1965). The P3 has a centrally positive scalp distribution peaking around 300 ms post stimulus and is not specific to the auditory modality. The P3 amplitude is affected by many parameters related to target detection and evaluation such as task difficulty (Kok, 2001), stimulus expectancy (Squires *et al.*, 1976), and the informational content of stimuli (Johnson, 1986).

# 2 GESTALT THEORY AND AUDITORY OBJECTS

A challenge of the auditory system is to form meaningful percepts from the sensory input. From the acoustic information impinging on the ear, it must reconstruct the sound-emitting sources by combining sounds that come from one source and separating them from sounds coming from different sources (Bregman, 1990). Perceptual grouping is governed by principles from the Gestalt theory (Koffka, 1935) such as ‘similarity’ or ‘good continuity’ (Bregman, 1990; Deutsch, 1999). In audition, grouping especially means integrating sounds over time, in speech or music, current sounds should be combined with past sounds.

## 2.1.1 Automatic and universal

The Gestalt theory postulated that perceptual grouping processes are automatic and universal. Grouping is assumed to be such a basic and general function that it does not require processing resources, and that it functions the same in every one of us, irrespective of age, culture, or (musical) skill (Koffka, 1935; see also, Imberty, 2000; Trehub, 2000). Jackendoff & Lerdahl (1983) proposed, along the same line, that perceptual groupings in music are intuitively formed and that the listener does not depend on musical knowledge to perceptually structure music. This idea has been empirically tested in behavioral studies (Deliège, 1987; Peretz, 1989). Musicians and nonmusicians scored, however, somewhat different in their segmentation of classical (Deliège, 1987) or folk music (Peretz, 1989) excerpts, musicians used certain grouping rules more often than nonmusicians. In these studies, subjects were attending to the sound material and because attention can modify groupings it might be that the differences between musicians and nonmusicians were caused by attention-driven grouping (Sussman *et al.*, 1998; 2002a).

### 2. 1.2.1 Study I aims and experiment

In this study the MMN was used to investigate whether automatic grouping principles operate similarly in musicians and nonmusicians. The hypothesis was that musicians might have more advanced abilities to group sequential tones compared to nonmusicians. This would indicate that—despite its fundamental character—auditory grouping is not independent of musical skill. Two grouping rules were tested in a group of professional musicians and a group of subjects without formal musical training. In the ‘pitch-similarity’ condition (see Figure 4, top) an isochronous sequence was presented in which four tones of one pitch were followed by four tones of another pitch, etc. The deviants were occasional fifth tones violating the length of the regular tone groups. In the ‘good-continuation-of-pitch’ condition (see Figure 4, bottom) an isochronous sequence was presented in which four consecutive tones were rising in pitch and then falling once (indicating the group-boundary) followed by again four rising tones etc. The deviant was a fifth tone continuing the rising pitch and thus violating the length of the standard tone groups. The four-tone standard groups started randomly on two different pitch levels. The deviant was always a fifth tone continuing the tone group starting on the lowest pitch level so that the fifth deviant tone did not introduce a new pitch. For an MMN to be

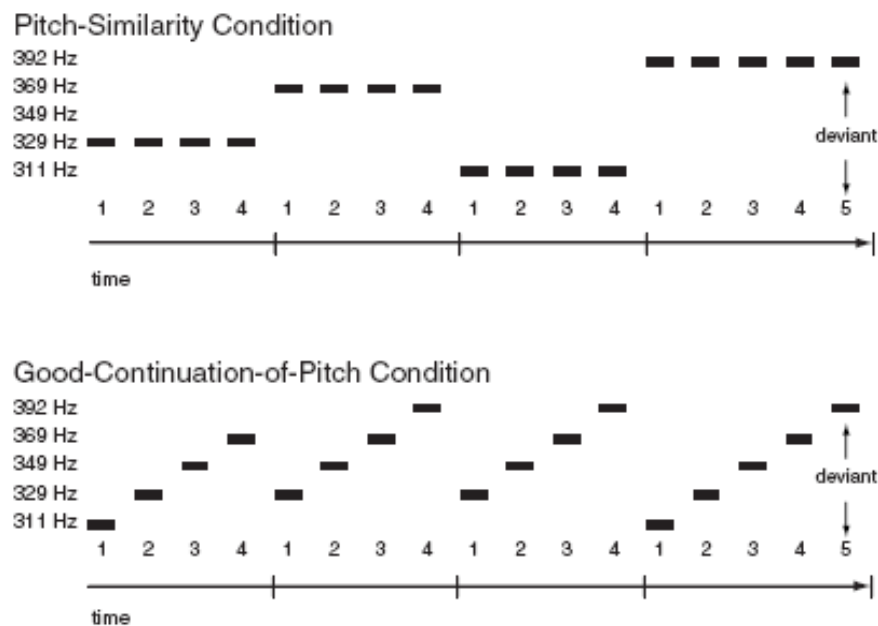


Figure 4. Illustration of the paradigm of Study I



elicited, the auditory system should group the four tones, encode the length of the standard tone groups, and detect the violation.

### 2. 1.2.2 *Study I* results and conclusions

An MMN was found to the deviants in the pitch-similarity condition, in both the musicians and nonmusicians (Figure 5, top) but in the good-continuation-of-pitch condition an MMN was found for the musicians, only (Figure 5, bottom). The auditory system of nonmusicians could thus group an isochronous sequence into tone groups of four, encode the tone groups, and detect the length violation. However, this depended on the difficulty of the grouping indicator because this was the only difference between the two conditions. Marking a group boundary by detecting a change from ‘same to different’ was a viable indicator for tone-group onset but a change in the direction of pitch change (from rise to fall) was too difficult for the nonmusicians. These results show that musicians and nonmusicians differ in their sensory auditory grouping skills. This difference between musicians and nonmusicians can be interpreted in two ways. Musical expertise might give rise to quantitative grouping advantages, i.e., musical experts might have relatively more fine-tuned grouping abilities than nonexperts, the formed group

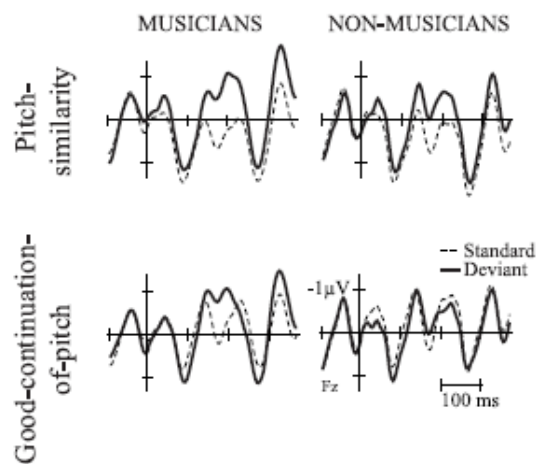


Figure 5. The ERPs obtained in *Study I*. Shown are the waveforms at Fz re-referenced with the average of the mastoid waves. The vertical bar at 0 ms indicates the onset of the fourth tones for the standard curve and the fifth tone for the deviant curve.

associations are stronger and deviants form more salient violations and thus elicit larger MMNs in musicians than in nonmusicians. The data support this view if assuming that no MMN is seen in the good-continuation-of-pitch condition for the nonmusicians, because it was too small to exceed the noise level. On the other hand, grouping processes might operate qualitatively different between musicians and nonmusicians so that musicians can make group associations that nonmusicians cannot make. If that is the case, an MMN is not observed for the nonmusicians in the good-continuation-of-pitch condition because it is simply not elicited. Adhering to the later view could lead to the conclusion that automatic grouping is not universal because musical experts use different grouping rules in organizing the auditory signal than nonmusicians.

### **2. 1.3 Why are musicians better at automatic processing?**

More fine-tuned or extended automatic grouping in musicians might be a result of a shift from controlled to automatic processing due to training (Jansma *et al.*, 2001; Shiffrin & Schneider, 1977). That musical training can cause plastic changes in the auditory system in general has been shown (Pantev *et al.*, 1998) but also the change detection mechanism underlying the MMN can improve due to training (even within one experimental training session) (e.g., Näätänen *et al.*, 1993; Tervaniemi *et al.*, 2001). The advanced grouping skills of musicians may have functional importance because it might leave limited attentional resources available for higher-order processes required for performing music at a professional level.

### **2. 2.1 The principle of exclusive allocation**

Another important principle of the Gestalt theory is that of exclusive allocation, stating that an element (e.g., a sound) cannot belong to two perceptual objects at the same time. It will be grouped to the object it is most similar or close to. In case an element is equally close to two objects an ambiguous situation emerges and the object perception becomes unstable, i.e., it can flip back and forth between the two alternatives. This flipping process can be modulated by conscious control (by attention-driven or schema-based grouping; see Bregman, 1995, Chapter 4) as for instance is the case in Escher's drawings of impossible staircases (Locher, 1992).

### 2. 2.2.1 Study II aims and experiment

The aim in this study was to investigate whether one tone in an ambiguous situation, is only part of one auditory object at a time while object perception is kept stable by attention. The stimulus sequence presented to the subjects contained two low (L) tones, two high (H) tones, and an intermediated tone (M) in the following order:  $\begin{matrix} H & H & H & H \\ L & LM & L & LM \end{matrix}$  etc. The pitch difference between the high and low tones and the high stimulus pace induced the streaming effect (Bregman, 1995). Streaming causes an alternating tone sequence to be perceived as two segregated streams of tones, one high and one low stream, i.e., the alternation is perceptually lost. In the paradigm used in this study, the intermediate tone could equally well belong to the high stream or the low stream, creating an ambiguous situation. Subjects in this experiment were instructed to group the intermediate tone to one of the patterns (which was controlled by a task, see appendix). Subjects were instructed to either perceive the repeating pattern MHH in one condition (selected-pattern-deviant condition) and the pattern HHM (alternative-pattern-deviant condition) in the other condition (high-stream subject group, for the low-stream subject group, see below). The inter-tone intervals were chosen to support the intended pattern formations. When the intermediate tone was occasionally presented slightly too early it violated the unity of the MHH pattern but not of the HHM pattern. In the MHH pattern, a too early intermediate tone made the tone not belong to the two high tones anymore but to the previous pattern (giving rise to the perception of a 'MHHM' and a 'HH' pattern instead). In the HHM pattern, a too early intermediate tone did not disturb the unity of the pattern. All tones were presented with a latency jitter so that the early presentation of the intermediate tone was not a temporal deviation *per se*. The patterns that subjects were *not* instructed to perceive, i.e., the alternative patterns, were the MLL pattern as an alternative to the perceived HHM pattern and the LLM pattern as an alternative to the perceived MHH pattern. If the intermediate tone is treated by the auditory system to belong to the perceived stream only, the early-onset violation of the intermediate tone in the selected-pattern-deviant condition (MHH pattern is perceived, LLM is not) should elicit an MMN but not the temporal violation of the intermediate tone in the alternative-pattern-deviant condition (HHM is perceived, MLL is not). Since only the early onset of the intermediate tone can violate the forming of the MLL grouping this condition can only give an MMN

if the intermediate tone belongs also to the not selected alternative pattern. In addition, there was a low-stream subject group receiving the mirroring instructions to perceive the MLL pattern (select-pattern-deviant condition) and the LLM pattern (alternative-pattern-deviant condition). There were no other differences. As a control condition served an unambiguous sound stream in which the MHH pattern was presented without the low tones.

### 2. 2.2.2 Study II results and conclusions

An MMN was found in the selected-pattern-deviant condition (see, Figure 6) and in the control condition indicating that the early onset of the intermediate tone indeed violated the auditory object perception. An MMN was not found in the alternative-pattern-deviant condition, indicating that the early onset of the intermediate tone in the alternative, not-perceived MLL pattern was not noticed by the auditory system. This shows that the intermediate tone cannot belong to two auditory objects at the same time and, thus, that the principle of exclusive allocation operates in auditory object perception.

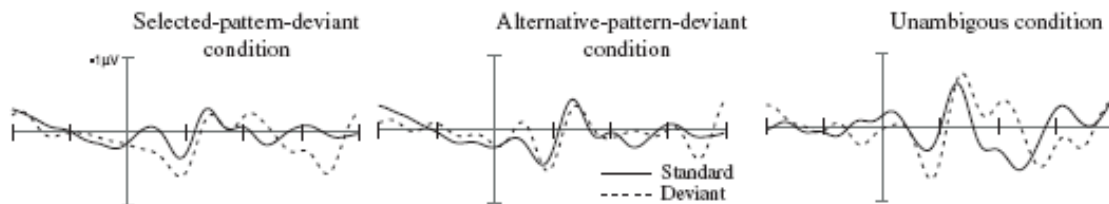


Figure 6. The ERPs obtained in *Study II*. Shown are the frontal electrodes for all three conditions (averaged over both low- and high- stream subject groups).

### 2. 2.3 The role of attention in an ambiguous situation

If strong acoustic grouping cues govern perception, a lot of conscious effort is required to change the immediate percept, if possible at all. The streaming effect can be so strong that it is not possible to hear the physical alternation of the high and low tones. When acoustic cues give rise to less strong groupings, attention can modulate whether tones are perceived grouped or segregated (van Noorden, 1975). These types of top-down influences can also modulate the input to the automatic auditory change-detection system

(Sussman *et al.*, 2002a). Sequential tones of the format SSSSDSSSSD etc. can, by will, be perceived as either individual tones (an MMN is elicited to the 'D') or as a repeating pattern (no MMN is elicited). In *Study II* also an interaction between automatic grouping and attended grouping was seen since attention could overcome the ambiguous acoustic cues and create a stable percept.

# 3 NUMBER PROCESSING

## 3.1 Sensory intelligence

Auditory detection of abstract regularities (see also 1. 3.3.5) has been termed ‘sensory intelligence’ because of the cognitive connotation (Näätänen *et al.*, 2001) of these processes that, nevertheless, still originate from the sensory cortex (Korzuykov *et al.*, 2003). An interesting question is how ‘intelligent’ auditory sensory processing can be. By studying experts the limits of auditory processing capabilities can be investigated.

## 3.2 Counting the beat in music

Counting the number of beats in a measure is part of the perceptual organization of music, and an important aspect of music performance (e.g., Palmer & Krumhansl, 1990). This could be reflected at a sensory level of auditory organization in musicians, and to some degree also be automatic. If detecting numerosity can take place outside of the focus of attention then this would exemplify rather advanced auditory processing functions.

### 3.3.1 Study III aims and experiment

In *Study III* it was tested in professional classical musicians and in nonmusicians whether the auditory system can extract and encode a numerical regularity. In addition, a temporal regularity was tested. In both conditions (see, Figure 7), the stimuli consisted of segments containing a certain number of tones of one pitch, followed by a segment of another pitch. In the ‘number’ condition, the number of tones in a segment was always four while the segment onset-to-onset time varied between 610 and 890 ms. The deviant violated the number of tones in a segment by adding one tone but did not violate the segment duration. In the ‘time’ condition, the onset-to-onset time of the segments was 750 ms while the number of tones in a segment varied between two and six. The deviant tone extended the segment duration without violating the number of tones in the standard segment.

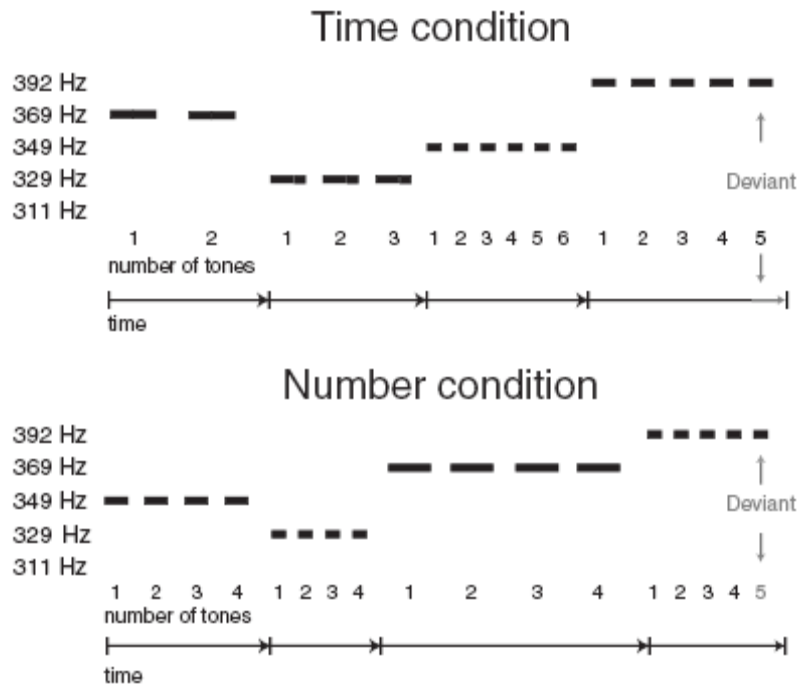


Figure 7. Illustration of the stimuli used in *Study III*.

### 3.3.2 *Study III* results and conclusions

An MMN was found in the time condition irrespective of musical skill, however, in the number condition, an MMN was found for the musicians, only (see, Figure 8). Thus, the auditory system of both subject groups could encode the regular duration of the segments but only the auditory system of musicians could discriminate the five-tone segments from the four-tone segments. Hence, musicians seem to have advanced and highly specialized auditory processing skills. It can, however, not be ruled out that nonmusicians can detect the number of tones in a series in an easier paradigm, containing fewer tones or shorter segments.

In the stimuli that were constructed for this study, time and number varied independently of each other. This was needed for the purpose of this study but it is an artificial situation. In music, counting the beats in a measure is not independent of the temporal organization; on the contrary, counting the beats helps musicians to stay synchronized with the rhythm. Even though the stimuli used in this study do not occur in natural situations, the results show that the sensory ability of musicians to detect

numerical regularities is neurophysiologically independent from the ability to detect temporal regularities.

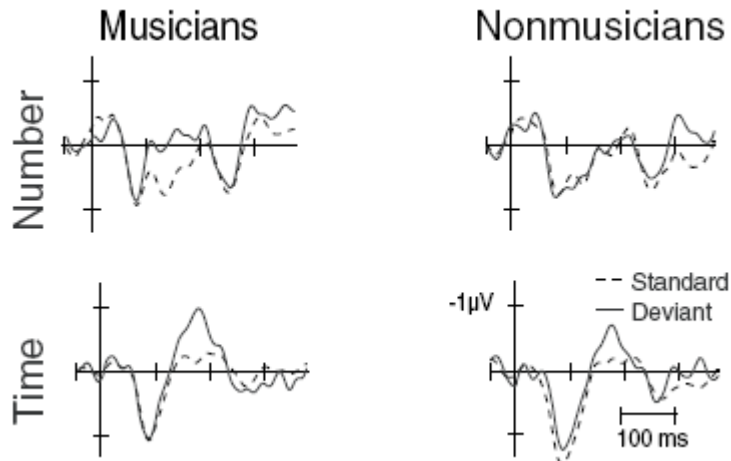


Figure 8. ERPs obtained in *Study III*. Shown are the ERPs at Fz.

### 3. 4 Auditory number detection in musicians

The musicians in this study were not actively counting to four or five but the auditory system could nevertheless encode the number of tones in the segments as it could distinguish the segments containing four tones from those containing five. Perfect instant and effortless detection of the number of items on a visual display has been called subitizing and it works for small numbers only, up to four or five (Jevons, 1871; Trick & Pylyshyn, 1994, see also, Piazza *et al.*, 2002). For larger numbers people start to make errors but also slow down in their response because they start to count attentively. In audition and especially in auditory grouping, a maximum of four might also play an important role. Music is often counted into two or four or grouped into units of two, three, or four (Abecasis *et al.*, 2005; Fraise, 1982). In addition, the results of *Study III* show, that at least in professional musicians, the processing of number, up to four/five, can take place outside of the focus of attention. Because of the abstract nature of numerosity processing (Miller *et al.*, 2003) this skill could be considered as a type of ‘sensory intelligence’.



# 4 HEMISPHERIC ASYMMETRY

## 4. 1 Speech perception

The two hemispheres are anatomically and functionally asymmetric (Zatorre *et al.*, 2002, Tervaniemi & Hugdahl, 2003). The most remarkable difference is that the left hemisphere is most commonly specialized in language processing. This asymmetry is already seen in the neural processing at the level of Heschl's gyrus (Liégeois-Chauvel *et al.*, 1999) and is also reflected in the left lateralization of the MMNm elicited by speech-sound deviants, at least in the majority of subjects (for an overview, see Pulvermüller *et al.*, 2001; Näätänen, 2001; Näätänen *et al.*, 1997). Other MMNs, elicited by non-lingual deviations are right lateralized. Whether sounds are perceived as language, and not just their physical features, determines where they are processed. Vowel contrasts from the own native language are processed predominantly in the left hemisphere whereas vowel contrasts from a foreign language are not (Näätänen *et al.*, 1997). The same lateralization effects have been found for words compared to pseudo-words (Pulvermüller *et al.*, 2001). Also interesting to mention are the MMNm responses to deviations in Morse code syllables that changed from being right lateralized in untrained subjects, to left lateralized in the same subjects after a very intensive Morse code training (Kujala *et al.*, 2003). These results reinforce the notion that the brain is functionally specialized already at a sensory cortical level and at that level not only processing physical sound features (Pantev *et al.*, 1996).

Music performance also includes aspects of communication and music can, although in a non-referential way, convey meaning. Musical communication often takes place by violating musical expectancies. In improvised jazz music, especially subtle variations in rhythm are used as means of communication.

### 4. 2.1 Study IV aims and experiment

*Study IV* is an MEG study testing whether the auditory system of jazz musicians is more sensitive to rhythm violations compared to that of nonmusicians and, further, whether

these deviations are predominantly processed in the (language dominant) left hemisphere in jazz musicians. Professional jazz musicians with a high score on a rhythm-aptitude test and nonmusicians with a low rhythm proficiency were presented with three stimuli, presented with equal probability (see, Figure 9). Each stimulus consisted of four measures of a four-beat rock rhythm. Two stimuli contained a rhythm deviation, either a syncopation replacing a weak beat with a strong one, or a beat coming too early and being incongruent with the temporal grid. The first violation is a smaller deviation than the second one and a known style figure in jazz music, whereas the second violation makes the music stumble. The last beats of the fourth measures were sometimes tuned lower or higher and were targets: subjects were instructed to press a button when they heard one.

#### **4.2.2 Study IV results and conclusions**

The incongruent beat gave larger response amplitudes in the ERFs in the MMNm latency range (100–150 ms) than the syncopation (see, Figure 10). This was the case for both subject groups and in both hemispheres. The responses of the musicians were overall larger than for the nonmusicians. This confirms that the incongruent beat violated the rhythm more than the syncopation and that musicians are more sensitive to rhythmic violations than nonmusicians. In addition, the estimated dipole amplitude for the incongruent beat was larger in the left than the right hemisphere of the musicians. In the nonmusicians it was the other way around, the dipole amplitude was larger in the right hemisphere than in the left hemisphere. In two subjects who underwent anatomical imaging, the estimated dipolar sources of the MMNm component were located in the auditory cortex of the temporal lobe. This shows that the level of rhythmic aptitude can determine the predominant hemisphere for processing rhythmic violations.

#### **4.3 Left lateralized rhythm processing in jazz musicians**

The auditory processing of an incongruent rhythm is left lateralized in musicians but not in nonmusicians indicating that the neural processing of rhythm in jazz musicians is special. Like speech processing is left lateralized in native listeners but not in foreign listeners, rhythm too can in certain experts, be predominantly processed in the language

dominant hemisphere. This is possibly the case because rhythm variations are used, especially in improvising jazz musicians, as a means of interaction and communication.



Figure 9. Illustration of the stimuli of *Study IV*: I) contains three standard measures without a violation, II) contains a syncopation and III) an incongruent beat. The arrows indicate the expected beat in stimulus I and the location of the violations in II and III.

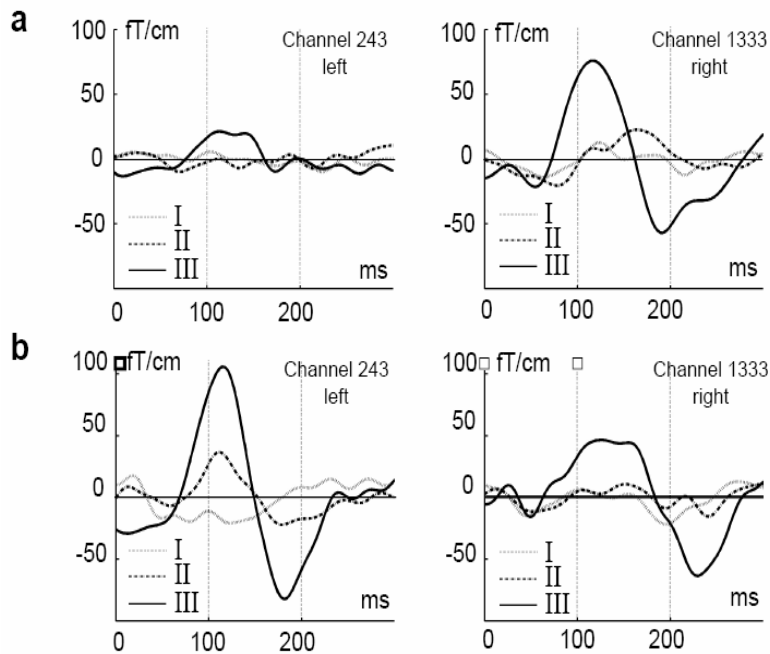


Figure 10. The ERFs obtained in *Study IV*. Shown are a left and right MEG channel from (a) one nonmusician and (b) one musician. I is the response elicited by the beat as expected, II is the response to the syncopation and III is the response to the incongruent beat.

# 5 AWARENESS AND AUDITORY CHANGE DETECTION

## 5. 1 MMN elicitation and perception

As mentioned in the introduction (1.3.3.4), MMN elicitation is in general believed to correlate with perception. The stimuli used in MMN experiments are, however, becoming more and more complex, and this might change the way stimuli are perceived. The stimulus sequences used in *Study III* were especially complex and an informal test after the experiment indicated that it was not very obvious that all subjects could detect the deviants, even though an MMN was elicited. It might thus be that in certain situations the outcome of the change-detection mechanism eliciting the MMN does not reach awareness. There is some supporting evidence for this view (Allen *et al.*, 2000; Tremblay *et al.*, 1998). In Tremblay *et al.* (1998) subjects were trained for several days to learn to discriminate a difficult speech contrast. The MMN as well as the ability to behaviorally discriminate the speech contrast were measured at various phases in the experiment. In about half of the subjects who learned to discriminate the speech contrast, the MMN appeared a day before the behavioral discrimination ability developed, indicating that MMN elicitation can, in certain cases, precede the ability to perceive the deviants. In another study, presenting an ‘abstract’ deviant (Paavilainen *et al.*, 2001), the auditory system could detect the deviants as indicated by MMN elicitation. The ‘knowledge’ of the auditory system was, however, not necessarily consciously available to all subjects. In a behavioral detection task and interview subsequent to the MMN measurement, three out of seven subjects did not express knowledge of the deviants. However, the ERPs of these three subjects were not analyzed separately from the other subjects and it cannot yet be ruled out that these subjects did not show an MMN.

## 5. 2 Implicit knowledge and intuition

Evidence that subjects can have, or even acquire, knowledge without being aware of it originates from implicit-learning studies, such as sequence learning (Buchner & Steffens, 2001) and artificial-grammar learning studies (Altmann *et al.*, 1995; Howard & Ballas,

1980). Subjects perform a task while a task-irrelevant aspect of the stimulus sequence is ordered. Subjects are not aware of this sequential order but behavioral indices show that they, after a while, perform faster to the ordered sequence than to a random control sequence, indicating that they have implicitly learned the stimulus order.

Implicit cognitive processes are assumed to be automatic (Nisbett & DeCamp Wilson, 1977) whereas explicit processes require consciousness and effort. Implicit knowledge has, furthermore, been proposed to be the substrate of intuition since intuitive judgments resemble implicit processes (Lieberman, 2000; Reber, 1989). It is difficult to verbalize all the information involved in an intuitive decision and one is not fully aware of all the processes contributing to it. This does, however, not mean that intuitive decisions cannot lead to satisfying outcomes (Dijksterhuis *et al.*, 2006). If implicit knowledge is the substrate of intuition this would mean that implicit knowledge could be used to make *correct* intuitive decisions.

Experimental Phase	Subject classification	Names of subject groups (number of subjects)
<u>EEG recording</u> <u>Ignore I</u>		<u>Ignore I Explicit (4)</u>
		<u>Ignore I No knowledge (19)</u>
<i>Interview I</i>	<i>4 subjects express explicit knowledge</i> <i>19 subjects do not express knowledge</i>	
Training		
<i>Interview II</i>	<i>9 subjects express explicit knowledge</i>	
	<i>12 subjects express intuitive knowledge</i>	
	<i>2 subjects are not able to detect the deviants</i>	
<u>EEG recording</u> <u>Attend</u>		<u>Attend Explicit (9)</u>
		<u>Attend Intuitive (12)</u>
<i>Interview III</i>	<i>None of the subjects has changed knowledge type</i>	
<u>EEG recording</u> <u>Ignore II</u>		<u>Ignore II Explicit (9)</u>
		<u>Ignore II Intuitive (12)</u>

Table 1. Chronology and subject classification of *Study IV*.

### **5. 3.1 *Study V* aims and experiment**

The aim of *Study V* was to determine whether MMN elicitation can be based on implicit knowledge and whether this implicit knowledge can be utilized for correct intuitive sound detection. The stimuli that were used contained an abstract regularity (like in Figure 3) and consisted of ascending tone pairs that were pair-wise roving within a range of frequencies. Occasionally, deviating descending tone pairs were presented. First an ignore condition ('Ignore I' condition) was recorded (see, Table 1) in which subjects were instructed to ignore the stimuli and watch a video. This was followed by a listening test and an interview (Interview I) to determine whether subjects could describe the stimulus structure and discriminate the deviants from the standards. Subjects performed subsequently an associative training task, in which each deviant coincided with a visual cue. The subjects were instructed to learn to detect the sounds coinciding with the visual stimulus. They did not at any time, receive information on the stimulus structure. The training was followed by a second interview (Interview II) to determine what they had learned about the stimuli during the training. Then an attended condition ('Attend' condition) was recorded in which the deviants were presented without the visual cue and subjects were instructed to press a button when they heard the sound that in the training had coincided with the flash, followed by a third interview (Interview III), and a second ignore condition ('Ignore II' condition). Subjects were classified, and their ERPs group-averaged, according to the knowledge of the stimulus sequence that they expressed in each of the three interviews.

### **5. 3.2 *Study V* results and conclusions**

In Interview I, four out of the 23 subjects (the 'Ignore I Explicit' group) were able to describe the stimuli exactly: they mentioned both the ascending frequency relationship in the standard tone pairs and the descending frequency relationship in the deviant tone pairs. An MMN was observed in these subjects (see, Figure 11). The remaining 19 subjects did not describe the stimuli accurately and none of these subjects (the 'Ignore I No-knowledge' group) noticed sounds that were sometimes different, or noticed sounds standing out among the others. Nevertheless, the auditory system was able to discriminate

the deviants from the standards because an MMN was elicited indicating that these subjects had implicit knowledge.

During the training session, 19 subjects learned to detect the deviants. In Interview II, five additional subjects expressed explicit knowledge of the sound structure making a total of nine subjects in this group (the ‘Attend Explicit’ group). In these subjects, an MMN and an N2b/P3 were elicited to the detected deviants (hits). Interestingly, the missed deviants also elicited an MMN, indicating that even though the subjects did not give a target response to these deviants, the auditory system did detect them. This shows that not all the knowledge available in the auditory system could be utilized by these subjects. The 12 subjects who did not express explicit knowledge in Interview II (the ‘Attend Intuitive’ group), gave either a wrong (e.g., ‘I pressed when a sound came quicker’) or a subjective explanation (e.g., ‘I pressed when it sounded darker’). In these subjects using an intuitive target-detection strategy, an MMN was elicited for the hits, but this MMN was not followed by an N2b/P3 complex. There was no MMN for the misses indicating that the knowledge available to the auditory system in

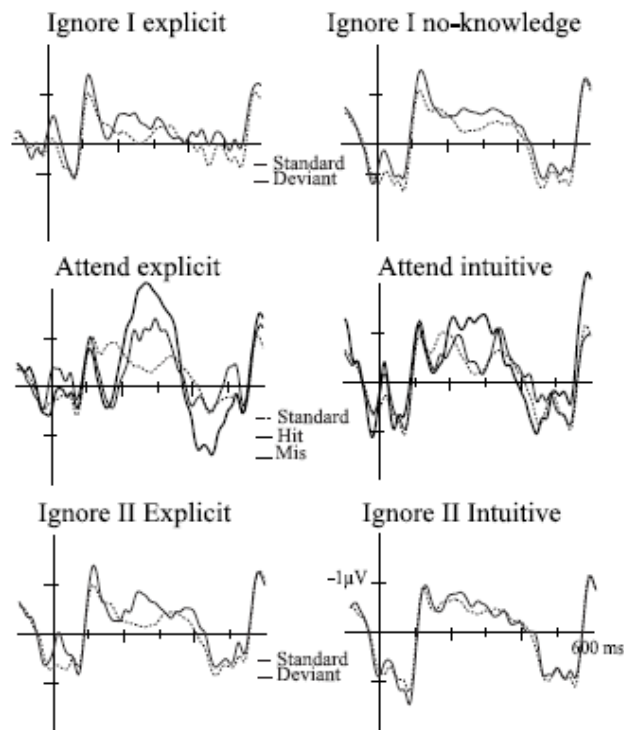


Figure 11. ERPs obtained *Study V*. Shown is Fz.

this group was optimally used for target detection. The overall behavioral performance was, however, lower in Attend Intuitive group than in the Attend Explicit group. The reaction time was respectively, 784 versus 553 ms and the deviant detection 38% versus 53% while the false alarm rates were similar (1.4% versus 0.4%).

In Interview III, no subjects expressed explicit knowledge, indicating that no further learning had occurred during the Attend condition. In the Ignore II condition an MMN was elicited in the subjects with explicit knowledge (the 'Ignore II explicit' group) as well as for the subjects who had expressed intuitive knowledge (the 'Ignore II intuitive' group).

#### **5. 4 Implicit change detection and intuition**

The results from *Study V* show that MMN elicitation does not necessarily correlate with perception. The auditory change-detection mechanism underlying the MMN elicitation can thus in certain cases operate implicitly. During a short associative training session, subjects could learn to access the implicit sensory knowledge and become aware of the deviants either intuitively or explicitly. The subjects who learned to detect the deviants intuitively were aware of them but not to such a degree that they could explain why a deviant was a deviant. This means that the detection of MMN-eliciting deviants does not necessarily require explicit knowledge of the stimuli. Subjects can with an intuitive detection strategy use the knowledge contained in the memory underlying the sensory change detection function for correct sound detection.



## 6 CONCLUSIONS

From behavioral research it is known that Gestalt principles govern auditory perception (Bregman, 1990). In *Study I* of this thesis, evidence was shown for grouping in the auditory system. The auditory system could group an isochronous sound stream into groups consisting of multiple-tones and detect a grouping deviant. This was indicated by MMN elicitation. Professional musicians showed grouping abilities according to both the principle of similarity and the principle of good-continuation. In nonmusicians, evidence was found for grouping according to the similarity rule, only. This shows that even basic auditory functions, are more advanced in musical experts. This is in line with previous research (e.g., Koelsch *et al.*, 1999). The current results can also be interpreted to suggest that sensory processes in musicians are not only a fine-tuning of universally available processes but that they are qualitatively different because the auditory system of musicians could make group associations based on a more advanced grouping rule than that of nonmusicians.

In *Study II* another Gestalt principle was investigated and it was demonstrated that the auditory system uses the principle of exclusive allocation in object perception. Subjects were asked to attentively maintain a stable pattern perception in an ambiguous situation. The pattern was formed by an intermediate tone that had to be grouped with either a high or a low stream of tones. Since the intermediate tone had an equal distance to each of the streams the two alternative percepts were equally likely. The auditory brain responses to pattern violations showed that the intermediate tone was grouped to either the low or the high stream but not to both at the same time, thus illustrating the operation of the principle of exclusive allocation.

How advanced can the sensory auditory processing be in experts? In *Study III* it was shown that the auditory sensory system of musicians can detect a numerical regularity of always four tones in a series. This was indicated by the MMN elicitation to occasional fifth tones. Nonmusicians did not show this ability even though they could detect a similarly complex temporal regularity. The ability to detect the number of tones

in a series without the need for focused attention might be an advantage for musicians in perceptually structuring music. In addition to the results of *Study I*, this difference between experts and nonexperts could be a candidate for a fundamental difference in sensory auditory function between these two subject groups. It can, however, not yet be ruled out that nonmusicians would be able to detect numerosities as well, for instance in a simplified version of the paradigm. It remains to be tested further, whether or not the highly abstract sensory ability of detecting the number of tones in a series is (to some extent) available to all of us.

In improvised jazz music, rhythmic variation plays an important role. In *Study IV* it was shown that the larger a rhythmical violation the larger the evoked brain responses and that in jazz musicians with a high rhythmical aptitude these violations were processed predominantly in the language dominant left hemisphere. This effect correlates with the importance of rhythm variations in the communication between improvising jazz musicians and might imply language-like processing of rhythm in this expert group.

The studies mentioned above, investigated auditory processes by utilizing the MNN as an index of the ability of the auditory system to detect occasional changes in a repetitive/regular sound stream without the requirement of attention. Subject can ignore the stimuli and an MMN will nevertheless be elicited. However, when they subsequently attend to the sounds and try to detect the deviants they are usually able to do so (e.g., Tiitinen *et al.*, 1994). Based on these findings there has been a consensus in the literature that MMN elicitation correlates with perception. In *Study V* it was shown that this might not be a valid conclusion for all types of stimulation. In the case of simple standard-deviant contrasts, the outcome of the sensory change detection mechanism might enter awareness easily, but when there are more complex regularities involved then the outcome of this process might not, or only partially, reach awareness. This was indeed found for a stimulus sequence containing deviants violating an abstract regularity. Not all subjects were able to detect the deviants when attending to the stimuli even though an MMN was elicited in a prior recording under ignore conditions. This indicates that these subject had implicit knowledge. This finding is supported by a few other studies (Allen *et al.*, 2000; Tremblay *et al.*, 1998). Only after a short associative training session, about half of the subjects expressed explicit knowledge of the sound structure whereas the other

half expressed only intuitive knowledge. The latter group learned during the training to only partially access the knowledge contained in the auditory system. Explicit knowledge is thus not a necessity for deviant detection, subjects can detect deviants by knowing *that* but without knowing *in what way* something is deviating. That MMN elicitation can be based on implicit knowledge and that subjects can intuitively detect deviants, gives new evidence for the MMN generator mechanism having an automatic character. Even though the auditory change-detection function does not in all circumstances operate fully independent of attention (Woldorff *et al.*, 1991; Sussman *et al.*, 2003a) it can, in certain situations operate without its outcome reaching full conscious awareness in attending subjects.

## REFERENCES

- Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception*, **22**, 549–562.
- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear & Hearing*, **16**, 38–51.
- Alho, K., & Sinervo, N. (1997). Preattentive processing of complex sounds in the human brain. *Neuroscience Letters*, **233**, 33–36.
- Alho, K., Woods, D.L., & Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology*, **31**, 469–479.
- Alho, K., Woods, D.L., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography & Clinical Neurophysiology*, **82**, 356–368.
- Allen, J., Kraus, N., & Bradlow, A. (2000). Neural representation of consciously imperceptible speech sound differences. *Perception & Psychophysics*, **62**, 1383–1393.
- Altmann, G.T.M., Dienes, Z., & Goode, A. (1995). On the modality-independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 899–912.
- Amenedo, E., & Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *European Journal of Neuroscience*, **12**, 2570–2574.
- Bak, C.K., Lebech, J., & Saermark, K. (1995). Dependence of the auditory evoked magnetic field (100 msec signal) of the human brain on the intensity of the stimulus. *Electrophysiology & Clinical Neurophysiology*, **61**, 141–149.
- Biacabe, B., Chevallier, J.M., Avan, P., & Bonfils, P. (2001). Functional anatomy of auditory brainstem nuclei: application to the anatomical basis of brainstem auditory evoked potentials. *Auris Nasus Larynx*, **28**, 85–94.
- Brattico, E., Tervaniemi, M., & Picton, T.W. (2003). Effects of brief discrimination-training on the auditory N1 wave. *NeuroReport*, **14**, 2489–2492.
- Bregman, A.S. (1990). Auditory scene analysis. Cambridge, MA: MIT Press.
- Buchner, A., & Steffens, M.C. (2001). Simultaneous learning of different regularities in sequence learning tasks: limits and characteristics. *Psychological Research*, **65**, 71–80.
- Colombo, M., D'Amato, M.R., Rodman, H.R., & Gross, C.G. (1990). Auditory association cortex lesions impair auditory short-term memory in monkeys. *Science*, **247**, 336–338.
- Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory & Cognition*, **19**, 909–921.

- Deliège, I. (1987). Grouping conditions in listening to music. An approach to Lerdahl & Jackendoff's grouping preference Rules. *Music Perception*, **4**, 325–360.
- Deutsch, D. (1999). Grouping mechanisms in music. In D. Deutsch (2nd Ed.), *The Psychology of Music*. New York: Academic Press.
- Dijksterhuis, A., Bos, M., Nordgren, L.F., & van Baaren, R.B. (2006) On making the right choice: the deliberation-without-attention effect, *Science*, **331**, 1005–1007.
- Doeller, C.F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., & Schroger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *NeuroImage*, **20**, 1270–1282.
- Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. *Annals of the New York Academy of Sciences*, **930**, 17–27.
- Drake, C., & Botte, M. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, **54**, 277–286.
- Erwin R, & Buchwald JS. (1986). Midlatency auditory evoked responses: differential effects of sleep in the human. *Electroencephalography & Clinical Neurophysiology*, **65**, 383–392.
- Escera, C., Alho, K., Schroger, E., & Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology & Neurootology*, **5**, 151–166.
- Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, **10**, 590–604.
- Fischer, C., Morlet, D., Bouchet, P., Luaute, J., Jourdan, C., & Salord, F. (1999). Mismatch negativity and late auditory evoked potentials in comatose patients. *Clinical Neurophysiology*, **110**, 1601–1610.
- Ford, J.M., & Hillyard, S.A. (1981). Event-related potentials (ERPs) to interruptions of a steady rhythm. *Psychophysiology*, **18**, 322–330.
- Fraisse, P. (1982). Rhythm and tempo. In D.Deutsch (Ed.), *The psychology of music* (pp. 194–180). London: Academic Press.
- Friedman, D., Cycowicz, Y.M., & Gaeta, H. (2001). The novelty P3a: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, **25**, 355–373.
- Friedman, D., & Simpson, G.V. (1994). ERP amplitude and scalp distribution to target and novel events: effects of temporal order in young, middle-aged and older adults. *Cognitive Brain Research*, **2**, 49–63.
- Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, **16**, 1010–1021.
- Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, **17**, 1578–1592.
- Giard, M.H., Lavikainen, J. Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience*, **7**, 133–143.

- Griffiths, T.D., Rees, G., Rees, A., Green, G.G., Witton, C., Rowe, D., Buchel, C., Turner, R., & Frackowiak, R.S. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nature Neuroscience*, **1**, 74–79.
- Grimm, S., Widmann, A. & Schröger, E. (2004). Differential processing of duration changes within short and long sounds in humans. *Neuroscience Letters*, **356**, 83–86.
- Hackett, T.A., Preuss, T.M., & Kaas, J.H. (2001). Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *Journal of Comparative Neurology*, **17**, 197–222.
- Halgren, E., Baudena, P., Clarke, J.M., Heit, G., Liegeois, C., Chauvel, P., & Musolino, A. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. I. Superior temporal plane and parietal lobe. *Electroencephalography & Clinical Neurophysiology*, **94**, 191–220.
- Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., & Lounasmaa, O.V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, **65**, 413–497.
- Hari, R., Hamalainen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R., & Sams, M. (1984). Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man. *Neuroscience Letters*, **50**, 127–32
- Herrmann, C.S., & Knight, R.T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience and Biobehavioral Reviews*, **25**, 465–76.
- Howard, J.H.J., & Ballas, J.A. (1980). Syntactic and semantic factors in the classification of nonspeech transient patterns. *Perception and Psychophysics*, **28**, 431–439.
- Imada, T., Hari, R., Loveless, N., McEvoy, L., & Sams, M. (1993). Determinants of the auditory mismatch response. *Electroencephalography Clinical Neurophysiology*, **87**, 144–153.
- Imberty, M. (2000). The question of innate competencies in musical communication. In N.L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music*. (pp. 449–462). Cambridge Massachusetts: The MIT Press.
- Jackendoff, F., & Lerdahl, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Jansma, J.M., Ramsey, N.F., Slagter, H.A., & Kahn, R.S. (2001). Functional anatomical correlates of controlled and automatic processing. *Journal of Cognitive Neuroscience*, **13**, 730–743.
- Jaramillo, M., Paavilainen, P., & Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters*, **25**, 101–104.
- Jerger, K., Biggins, C., & Fein, G. (1992). P50 suppression is not affected by attentional manipulations. *Biological Psychiatry*, **31**, 365–377.
- Jevons, W.S. (1871). The power of numerical discrimination. *Nature*, **3**, 281–282.
- Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, **23**, 367–383.
- Kandel, E.R., Schwartz, J.H., & Jessell, T.M. (2000). *Principles of neuroscience*. USA: McGraw-Hill.
- Koelsch, S., & Mulder, J. (2002). Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology*, **113**, 862–869.

- Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior attentive and pre-attentive auditory processing in musicians. *NeuroReport*, **10**, 1309–1313.
- Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Hartcourt, Brace and Company.
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity, *Psychophysiology*, **38**, 557–577.
- Korzyukov, O.A., Winkler, I., Gumenyuk, V.I., & Alho, K. (2003). Processing abstract auditory features in the human auditory cortex. *NeuroImage*, **20**, 2245–2258.
- Kropotov, J.D., Alho, K., Näätänen, R., Ponomarev, V.A., Kropotova, O.V., Anichkov, A.D., & Nechaev, V.B. (2000). Human auditory-cortex mechanisms of preattentive sound discrimination. *Neuroscience Letters*, **280**, 87–90.
- Kujala, A., Huotilainen, M., Uther, M., Shtyrov, Y., Monto, S., Ilmoniemi, R., & Näätänen, R. (2003). Plastic cortical changes induced by learning to communicate with non-speech sounds. *NeuroReport*, **14**, 1683–1687.
- Lieberman, M.D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin*, **126**, 109–137.
- Liegeois-Chauvel, C., de Graaf, J.B., Laguitton, V., & Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cerebral Cortex*, **9**, 484–496.
- Liegeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalography Clinical Neurophysiology*, **92**, 204–214.
- Locher, J.L. (1992). *M.C. Escher: Life and Work*. Portland, Or.: Book News, Inc.
- Loewy, D.H., Campbell, K.B., & Bastien, C. (1996). The mismatch negativity to frequency deviant stimuli during natural sleep. *Electroencephalography & Clinical Neurophysiology*, **98**, 493–501.
- Lütkenhöner, B., Seither-Preisler, A., & Seither, S. (2005). Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and nonmusicians. *NeuroImage*, **30**, 927–937
- Markand, O.N. (1994). Brainstem auditory evoked potentials. *Journal of Clinical Neurophysiology*, **11**, 319–342.
- Menning, H., Roberts, L.E., & Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *NeuroReport*, **11**, 817–822.
- Miller, E.K., Nieder, A., Freedman, D.J., & Wallis, J.D. (2003). Neural correlates of categories and concepts. *Current Opinion in Neurobiology*, **13**, 198–203.
- Mizuno, C., Schwartz, J.L., & Cazals, Y. (1994). Periodicity of long-term context can influence gap detection. *Hearing Research*, **78**, 41–48.
- Münste, T.F., Altenmüller, E., & Jancke, L. (2000) The musician's brain as a model of neuroplasticity. *Nature Review Neuroscience*, **3**, 473–478.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials. *Behavioral and Brain Sciences*, **13**, 201–288.

- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum Associated.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, **38**, 1–21.
- Näätänen, R., Gaillard, A.W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, **42**, 313–329.
- Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology*, **42**, 25–32.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, **385**, 432–434.
- Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1987). The mismatch negativity to intensity changes in an auditory stimulus sequence. *Electroencephalography & Clinical Neurophysiology: Supplement*, **40**, 125–131.
- Näätänen R., & Picton TW. (1986). N2 and automatic versus controlled processes. *Electroencephalography & Clinical Neurophysiology: Supplement*, **38**, 169–186.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, **24**, 375–425.
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., & Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *NeuroReport*, **4**, 503–506.
- Näätänen, R., Simpson, M., & Loveless, N.E. (1982). Stimulus deviance and evoked potentials. *Psychological Biology*, **14**, 53–89.
- Näätänen, R., Syssoeva, O., & Takegate, R. (2004). Automatic time perception in the human brain for intervals ranging from milliseconds to seconds. *Psychophysiology*, **41**, 660–663.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends in Neuroscience*, **24**, 283–288.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, **125**, 826–859.
- Nisbett, R.E., & DeCamp Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, **84**, 231–259.
- Novak, G.P., Ritter, W., Vaughan, H.G., Jr., & Wiznitzer, M.L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*, **75**, 255–275.
- Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, **20**, 26–36.
- Paavilainen, P., Jaramillo, M., Naatanen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, **23**, 179–182.



- Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology*, **38**, 359–365.
- Paavilainen, P., Tiitinen, H., Alho, K., & Näätänen, R. (1993). Mismatch negativity to slight pitch changes outside strong attentional focus. *Biological Psychology*, **37**, 23–41.
- Palmer, C., & Krumhansl, C.L. (1990). Mental representations of musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 728–741.
- Palmer, A.R., & Summerfield, A.Q. (2002). Microelectrode and neuroimaging studies of central auditory function. *British Medical Bulletin*, **63**, 95–105.
- Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography & Clinical Neurophysiology*, **94**, 26–40.
- Pantev, C., Elbert, T., Ross, B., Eulitz, C., & Terhardt, E. (1996). Binaural fusion and the representation of virtual pitch in the human auditory cortex. *Hearing Research*, **100**, 164–170.
- Pantev, C., & Lütkenhöner, B. (2000). Magnetoencephalographic studies of functional organization and plasticity of the human auditory cortex. *Journal of Clinical Neurophysiology*, **17**, 130–142.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, **392**, 811–814.
- Peretz, I. (1989). Clustering in music: an appraisal of task factors. *International Journal of Psychophysiology*, **24**, 157–178.
- Perrault, N., & Picton, T.W. (1984). Event-related potentials recorded from the scalp and nasopharynx. II. N2, P3 and slow wave. *Electroencephalography & Clinical Neurophysiology*, **59**, 261–278.
- Piazza, M., Mechelli, A., Butterworth, B., & Price, C.J. (2002). Are subitizing and counting implemented as separate or overlapping processes? *NeuroImage*, **15**, 435–446.
- Picton, T.W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: different water in the same river. *Audiology & NeuroOtology*, **5**, 111–139.
- Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R.J., & Näätänen, R. (2001). Memory traces for words as revealed by the mismatch negativity. *NeuroImage*, **14**, 607–616.
- Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, **77**, 317–327.
- Riedel, H., & Kollmeier, B. (2002). Auditory brain stem responses evoked by lateralized clicks: is lateralization extracted in the human brain stem? *Hearing Research*, **163**, 12–26.
- Rinne, T., Degerman, A., & Alho, K. (2005). Superior temporal and inferior frontal cortices are activated by infrequent sound duration decrements: an fMRI study. *NeuroImage*, **26**, 66–72.
- Ritter, W., Deacon, D., Gomes, H., Javitt, D.C., & Vaughan, H.D., Jr. (1995). The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review. *Ear & Hearing*, **16**, 52–67.

- Rosburg, T., Trautner, P., Dietl, T., Korzyukov, O.A., Boutros, N.N., Schaller, C., Elger, C.E., & Kurthen M. (2005). Subdural recordings of the mismatch negativity (MMN) in patients with focal epilepsy. *Brain*, **128**, 819–828.
- Rossing, T.D. (1990). *The science of sound*. USA: Addison-Wesley.
- Rupp, A., Uppenkamp, S., Gutschalk, A., Beucker, R., Patterson, R.D., Dau, T., & Scherg, M. (2002). The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans (1). *Hearing Research*, **174**, 19–31.
- Rüsseler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Münte, T.F. (2001). Event-related brain potentials to sound omissions differ in musicians and nonmusicians. *Neuroscience Letters*, **308**, 33–36.
- Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, M., & Näätänen, R. (1992). Representations of abstract attributes of auditory stimuli in the human brain. *NeuroReport*, **3**, 1149–1151.
- Sallinen, M., Kaartinen, J., & Lyytinen, H. (1994). Is the appearance of mismatch negativity during stage 2 sleep related to the elicitation of K-complex? *Electroencephalography & Clinical Neurophysiology*, **91**, 140–148.
- Sams, M., Alho, K., & Näätänen, R. (1983). Sequential effects on the ERP in discriminating. *Biological Psychology*, **17**, 41–58.
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography & Clinical Neurophysiology*, **62**, 437–448.
- Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, **267**, 699–701.
- Schmithorst, V.J., & Holland, S.K. (2003). The effect of musical training on music processing: a functional magnetic resonance imaging study in humans. *Neuroscience Letters*, **348**, 65–68.
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H.J., Dosch, H.G., Bleeck, S., Stippich, C., & Rupp, A. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience*, **8**, 1241–7.
- Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, **8**, 527–539.
- Shahin, A., Bosnyak, D.J., Trainor, L.J., & Roberts, L.E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience*, **23**, 5545–5552.
- Shaw N.A. (1995). The temporal relationship between the brainstem and primary cortical auditory evoked potentials. *Progress in Neurobiology*, **47**, 95–103.
- Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and general theory. *Psychological Review*, **84**, 127–190.
- Shtyrov Y, & Pulvermüller F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *NeuroReport*, **13**, 521–525.
- Sinkkonen, J., Kaski, S., Huottilainen, M., Ilmoniemi, R.J., Näätänen, R., & Kaila, K. (1996). Optimal resource allocation for novelty detection in a human auditory memory. *NeuroReport*, **7**, 2479–2482.

- Squires, K.C., Wickens, C., Squires, N.K., & Donchin E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, **193**, 1142–1146.
- Stapells, D.R., & Oates, P. (1997). Estimation of the pure-tone audiogram by the auditory brainstem response: a review. *Audiology & Neurootology*, **2**, 257–280.
- Sussman, E., Ritter, W., & Vaughan, H.G., Jr. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research*, **789**, 130–138.
- Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., & Näätänen, R. (2002). Top-down effects can modify the initially stimulus-driven auditory organization. *Cognitive Brain Research*, **13**, 393–405.
- Sussman, E., Winkler, I., & Schröger, E. (2003b). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin & Review*, **10**, 630–637.
- Sussman, E., Winkler, I., & Wang, W. (2003a). MMN and attention: competition for deviance detection. *Psychophysiology*, **40**, 430–435.
- Sutton, S., Braren, M., Zubin, J., & John, E.R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, **150**, 1187–1188.
- Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letters*, **266**, 109–112.
- Tervaniemi, M. & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Research Reviews*, **43**, 231–246.
- Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. (1997). The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters*, **226**, 1–4.
- Tervaniemi, M., Rytkönen, M., Schröger E., Ilmoniemi RJ, & Näätänen R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning & Memory*, **8**, 295–300.
- Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, **372**, 90–92.
- Trehub, S. (2000). Human processing predispositions and musical universals. In N.L. Wallin, B. Merker, & S. Brown (Eds.). *The origins of music*. (pp. 427–448). Cambridge Massachusetts: The MIT Press.
- Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *NeuroReport*, **9**, 3557–3560.
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear & Hearing*, **22**, 79–90.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, **101**, 80–102.
- Van Noorden, L.P.A.S. (1975). *Temporal coherence in the perception of tone sequences*. Doctoral dissertation, Eindhoven University of Technology.

- Winkler, I., Korzyukov, O., Gumenyuk, V., Cowan, N., Linkenkaer-Hansen, K., Ilmoniemi, R.J., Alho, K., & Näätänen, R. (2002). Temporary and longer term retention of acoustic information. *Psychophysiology*, **39**, 530–534.
- Winkler, I., Teder-Sälejärvi, W.A., Horváth, J., Näätänen, R., & Sussman, E. (2002). Human auditory cortex track task-irrelevant sound sources. *NeuroReport*, **14**, 2053–2056.
- Woldorff, M.G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, **30**, 98–119.
- Woldorff, M.G., Gallen, C.C., Hampson, S.A., Hillyard, S.A., Pantev, C., Sobel, D., & Bloom, F.E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 8722–8726.
- Woldorff, M.G., Hackley, S.A., & Hillyard, S.A. (1991). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, **28**, 30–42.
- Woldorff, M.G., Hillyard, S.A., Gallen, C.C., Hampson, S.R., & Bloom, F.E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology*, **35**, 283–292.
- Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport*, **8**, 1971–1974.
- Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., & Pantev, C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cerebral Cortex*, **11**, 411–423.
- Zatorre, R.J., Belin, P., & Penhune, V.B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Science*, **6**, 37–46.

## APPENDIX: METHODS

### EEG/MEG SETUP

Study I: EEG was recorded with Ag/AgCl electrodes placed at three midline positions Fz, Cz, and Pz, at the left and right mastoids (Lm and Rm), and at sites along the coronal chain at one thirds (L1 and R1) and two thirds (L2 and R2) between Fz and the mastoids, left and right. The reference electrode was placed on the nose. The horizontal electro-oculogram (EOG) was recorded from electrodes placed lateral to the outer canthi of both eyes and the vertical EOG from electrodes placed above and below the left eye. The reference electrode was placed on the nose. Study II: EEG was recorded with Ag/AgCl electrodes from 8 scalp locations (F3, F4, C3, C4, P3, P4, Lm, and Rm). EOG electrodes and reference were placed as in study I. Study III: See study I. Study IV: MEG was recorded with a 306-channel whole-head device (Elektra Neuromag) in a magnetically shielded room. The EOG was recorded like in study I. Study V: EEG was recorded with Ag/AgCl electrodes at 32 recording sites, placed according to the 10–20 system by using an electrode cap. In addition, electrodes were placed at the left (Lm) and right (Rm) mastoids. EOG and reference electrodes were placed as in Study I.

### DATA ACQUISITION AND FILTERING:

Study I: The signal was sampled at 250 Hz and band-pass filtered (1.5–35 Hz). Study II: The signal was sampled at 250 Hz and band-pass filtered (2.5–16.0 Hz). Study III: See study I. Study IV: The signal was sampled at 300 Hz and band-pass filtered (2–30 Hz). Study V: The signal was sampled at 500 Hz and down-sampled offline to 250 Hz (with an anti-aliasing function) and band-pass filtered (2–35 Hz).

### EPOCHS AND ARTIFACT REJECTION:

Study I: The EEG signal was divided in epochs starting 100 ms before and ending 375 ms after the onset of a tone. Baseline correction was applied on single epochs after which epochs containing signal exceeding  $\pm 75 \mu\text{V}$  at any electrode were rejected from further analysis. Study II: Epochs started 200 ms pre-stimulus and ended 400 ms post-stimulus. Epochs with voltage difference between temporally adjacent sampling points exceeding  $8 \mu\text{V}$  on any channel were rejected from further analysis. Due to the fast semi-random-SOA presentation, long-latency ERP components elicited by previous stimuli were expected to overlap the ERP components of interest. To reduce this effect, ADJAR level 1 procedure was applied (Woldorff, 1993). Study III: The same as in study I but with a 50 ms baseline. Study IV: Epochs started 50 ms before the deviant stimulus and ended 300 ms post-stimulus. Online EOG artifact-rejection criteria were set to  $\pm 100 \mu\text{V}$ . Study V: The EEG signal was divided into epochs from –100 to 600 ms to the onset of the second tone of the stimulus pairs. Single epochs were baseline corrected after which all epochs containing signal exceeding  $\pm 60 \mu\text{V}$  were removed.

### COMPONENT CHARACTERIZATION AND STATISTICAL ANALYSIS:

Study I: The MMN was measured by taking the average in a 40 ms window around the peak of the grand-average for both the standard and the deviant. MMN elicitation was statistically tested with a two-way ANOVA (Electrode [Fz, L1, R1, Lm, Rm]  $\times$  Stimulus Type [standard, deviant]) separately for each

condition. To test differences between subject group and condition, the difference (standard minus deviant) in a 40 ms window around the peak on Fz (re-referenced with the average of the mastoids) was taken and an additional ANOVA was performed (Expertise [musicians, nonmusicians] × Condition [pitch-similarity, good-continuation-of-pitch]). Study II: The mean amplitudes in the MMN latency range of 212–236 ms was calculated and the deviant-minus-standard differences tested against zero using a *t*-test (pooled over the two groups of subjects and over F3 and F4) for each of the three conditions. In addition, an ANOVA (Group [high, low] × Condition [Perceived-Pattern-Deviant, Alternative-Pattern-Deviant] × Stimulus Type [standard, deviant] × Electrode [F3, F4]) was performed.

Study III: MMN elicitation was determined by taking the mean amplitude in a 32 ms window centered on the peak. For each subject group and condition, one-sided *t*-tests (deviant < standard) were performed on Fz and the average of L1 and R1. In addition, an ANOVA was performed on the MMN amplitude (Expertise [musicians, nonmusicians] × Condition [time, number] × Electrode [Fz, L1, R1]). Study IV: From the ERF responses to the standard, syncopation and incongruent beat, the maximum mean-gradient amplitude (MGA) in the interval 100–170 ms was taken. Equivalent current dipoles (ECD) were estimated at the latency of the maximal MGA for each hemisphere with a spherical head model. The amplitudes (*A*) of the dipoles were used to calculate an asymmetry index  $(A_{\text{right}} - A_{\text{left}}) / (A_{\text{right}} + A_{\text{left}})$  and a *t*-test was performed across subject groups. Study V: The MMN amplitudes were taken from Fz in a 64 ms window centered on the peak amplitude of the grand-average deviant-minus-standard wave. P3 amplitudes were determined in the same way from Pz. Two-tailed *t*-tests were used to determine the presence of MMN and P3, to test whether the MMN amplitudes were different between subject groups, and to test for differences in behavioral measures.

#### SUBJECTS:

Study I: Eleven musicians and 12 nonmusicians participated in the experiment. Musicians had reached, as a minimum, the level of acceptance into a music academy (Sibelius Academy, Helsinki). Nonmusicians had never studied any form of music at a formal/professional level. Study II: 21 subjects participated in the experiment. The data of one subject were rejected due to artifacts. Study III: 13 musicians and 15 nonmusicians participated in the experiment (selection criteria the same as in study I). Study IV: Subject groups were selected with a rhythm aptitude test that is used as part of entry examinations to music conservatories in Denmark. Eight nonmusicians with a low score and nine jazz musicians with a high score (educated at the Sibelius Academy of Music, Helsinki, Finland) participated in the study. Study V: Twenty-four subjects participated in the experiment. The data of one subject were discarded because of a lack of motivation to participate.

#### SUBJECT INSTRUCTIONS AND TASKS:

Study I: Subjects were instructed to watch a self-selected subtitled movie without the soundtrack. Study II: 10 subjects were instructed to group the intermediate tone together with the high tones and maintain this perception throughout the stimulus blocks (high stream group). The other 10 subjects were instructed to group the intermediate tone with the low tones (low stream group). All subjects were instructed to press a

button on a response pad with three buttons to indicate whether a target tone appeared on the first, second or third position in the pattern that they were instructed to hear. Subjects were trained in both the pattern perception and the task before starting the EEG recordings. Study III: See study I. Study IV: Subjects were instructed to press a button to the down- or up-tuned beat in the fourth measure of the stimuli. Study V: In the ignore conditions subjects were instructed to watch a self-selected subtitled movie without hearing the soundtrack. In the attend condition the subjects were instructed to detect the ‘sounds that sometimes were different’ and that in the associative training session had coincided with a flash on a screen.

#### STIMULI:

Study I: ‘pitch-similarity condition’: 100 ms sine-wave tones (50 dB above hearing threshold; 10 ms rise and 10 ms fall times) were presented with a constant inter-tone interval of 87.5 ms. Stimulus sequences consisted of identical-pitch four-tone segments varying on five frequency levels, ranging from 311.1 Hz to 392 Hz in semitone steps. Ten percent of these four-tone groups were prolonged by a fifth tone of the same pitch violating the standard group-length. ‘Good-Continuation-of-Pitch Condition’: The same five tones with the same inter-tone-interval were used as the ‘pitch-similarity’ condition. Standard groups of four-tones were ascending in pitch. A four-tone segment could start one of the two lowest pitch levels. Ten percent of the tone segments contained an additional ascending step. This was always a continuation of the tone group starting at the lowest pitch level in order to avoid the introduction of a new (and therefore deviating) frequency. Study II: Three tones were presented (low pitch: 548 Hz, 50dB above individual hearing threshold (AHT), intermediate pitch: 740 Hz, 48 dB AHT, and high pitch 1155, 45 dB AHT) in alternating order. These stimuli induced the streaming effect. The tone duration was 30 ms and the tone-onset-to-onset time (SOA) contained variation with medians between 60 and 280 ms. A cycle of five tones (high, low, high, low, intermediate or low, high, low, high, intermediate) had an average duration of 732 ms. The timing of the tones encouraged the perception of repeating triplets. In 8% of the cycles the interval preceding the intermediate tone was shortened from having a median of 320 ms to a median of 210 ms. The tone sequence also contained occasional (5%) intensity deviants (+12 dB) in the perceived pattern. These deviations could occur in any of the three tones of the perceived pattern and served as targets in the task. Study III: Each condition consisted of a certain number of tones of the same pitch (a segment) followed by a number of tones of another pitch. The segments were varying on five pitch levels, ranging from 311.1 Hz to 392 Hz in semitone steps on the musical scale. In the ‘number’ condition there were always four tones of the same pitch followed by four tones of another pitch. The duration of the segments varied between 610 and 890 ms. The duration of the individual tones of a tone segment varied between 65 and 100 ms and the inter-tone-intervals (ITI) between 87.5 and 122.5 ms. Within a tone segment the individual tone duration and ITI were always the same. The deviant was an additional tone violating the number of tones in the standard segments but never extended the duration of the segments beyond 890 ms, to avoid a temporal violation. 10% of the segments contained a deviant. In the ‘time’ condition the segment onset-to-onset was constant at 750 ms but the number of tones within that time varied between two and six. The tones used to construct this sequence varied in duration between 60 and 200 ms and the ITI between 58.3 and 195 ms in

combinations that resulted in segment durations of 750 ms. Within each segment the individual tone duration and ITI were always the same. 10% of the segments contained a deviating tone (starting at 750 ms) that extended the segments duration. Only the standard segments with 2 to 5 tones could have deviants added so that they did not form a number-of-tone violation. Study IV: Rhythmic sequences were made of realistic broadband drum sounds. The stimulus was a simple four-beat rock rhythm with 5% of the last snare drum beats in the fourth measure tuned up or down (to serve as targets) and with 33% of the third measures containing a syncopation and with 33% of the third measures containing a beat incongruent with the meter.

Study V: Subject were presented with randomly roving ascending tone pairs as standards, of which 10% was replaced by roving descending tone pairs as deviants. The individual tone duration was 75 ms, the within-pair interval 20 ms and the inter-pair interval 300 ms. Tones ranged on the musical scale from C4 to C5 (261.6 Hz to 523.3 Hz) in semitone steps. The frequency step within the standard pairs was 5 semitones ascending and the frequency change within the deviant pairs was 5 semitones descending.