CONTENTS

ABSTRACT ................................................................................................................................. 5
TIIVISTELMÄ .............................................................................................................................. 6
ACKNOWLEDGEMENTS ............................................................................................................. 7
LIST OF ORIGINAL PUBLICATIONS .......................................................................................... 9

1 INTRODUCTION ...................................................................................................................... 10
  1.1 Processing of sound identity and location in the brain ......................................................... 10
      1.1.1 The subcortical ascending auditory pathway ................................................................. 10
      1.1.2 The auditory cortex ....................................................................................................... 11
  1.2 Effects of auditory selective attention on auditory cortex activity ..................................... 14
  1.3 Näätänen’s attentional-trace theory ...................................................................................... 15
      1.3.1 The early PN and selection of attended sounds in the auditory cortex ......................... 15
      1.3.2 Attentional processing of pitch and location of sounds in the auditory cortex ............. 16
      1.3.3 The role of frontal and parietal cortical areas in auditory attention .............................. 17
      1.3.4 Audiovisual selective attention .................................................................................. 18

2 AIMS OF THE PRESENT THESIS .......................................................................................... 20

3 METHODS AND RESULTS ..................................................................................................... 21
  3.1 Overview of brain research methods used in Studies I-IV .................................................... 21
      3.1.1 fMRI ............................................................................................................................. 21
      3.1.2 EEG and MEG ............................................................................................................... 22
  3.2 Participants in Studies I-IV ................................................................................................. 23
  3.3 Stimuli and procedures in Studies I-IV ................................................................................. 24
  3.4 fMRI imaging and data analyses in Studies I, II and IV ....................................................... 27
  3.5 EEG and MEG recording, and data analyses in Study III .................................................... 27
  3.6 Study I. Modulation of auditory-cortex activation by sound presentation rate and attention ................................................................................................................................. 28
      3.6.1 Specific experimental setting and data analyses ........................................................... 28
      3.6.2 Results .......................................................................................................................... 29
  3.7 Study II. Selective attention to sound location or pitch studied with fMRI ........................ 31
      3.7.1 Specific experimental setting and data analyses .......................................................... 31
      3.7.2 Results .......................................................................................................................... 33
  3.8 Study III. Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields .................................................................................................. 36
      3.8.1 Specific experimental settings and data analyses ......................................................... 36
      3.8.2 Results .......................................................................................................................... 36
3.9 Study IV. Human brain activity associated with audiovisual perception and attention................................................................. 40
3.9.1 Specific experimental setting and data analyses .................. 40
3.9.2 Results.................................................................................. 42

4 DISCUSSION .......................................................... 45

4.1 Modulation of auditory cortex attention effects with increasing sound presentation rate (Study I)......................................................... 45
4.2 Attention-related processing of pitch and location of sounds in the brain (Studies II and III)................................................................. 47
4.3 Attention-related processing of audiovisual information in the brain (Study IV)................................................................. 49
4.4 Methodological considerations.................................................. 51
4.5 Conclusions........................................................................... 53

5 REFERENCES ........................................................................ 54

ORIGINAL PUBLICATIONS .................................................. 63
ABSTRACT

Selective attention refers to the process in which certain information is actively selected for conscious processing, while other information is ignored. The aim of the present studies was to investigate the human brain mechanisms of auditory and audiovisual selective attention with functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and magnetoencephalography (MEG). The main focus was on attention-related processing in the auditory cortex.

It was found that selective attention to sounds strongly enhances auditory cortex activity associated with processing the sounds. In addition, the amplitude of this attention-related modulation was shown to increase with the presentation rate of attended sounds. Attention to the pitch of sounds and to their location appeared to enhance activity in overlapping auditory-cortex regions. However, attention to location produced stronger activity than attention to pitch in the temporo-parietal junction and frontal cortical regions. In addition, a study on bimodal attentional selection found stronger audiovisual than auditory or visual attention-related modulations in the auditory cortex. These results were discussed in light of Näätänen’s attentional-trace theory and other research concerning the brain mechanisms of selective attention.
TIIVISTELMÄ

Valikoivalla tarkkaavaisuudella tarkoitetaan prosessia, jossa tietoiseen käsittelyyn valitaan aktiivisesti jotain tietoa ja muu tieto jätetään huomioimatta. Tämän väitöskirjatutkimuksen tavoite oli selvittää kuulotietoon kohdistuvan sekä kuulo- ja näkötietoa yhdistävän valikoivan tarkkaavaisuuden aivomekanismeja ihmisellä. Tutkimusmenetelminä käytettiin toiminnallista magneettikuvausta (fMRI), elektroenkefalografiaa (EEG) ja magnetoenkefalografiaa (MEG). Tutkimus keskittyi erityisesti tarkkaavaisuuden alaiseen tiedonkäsittelyyn kuuloaivokuorella.

ACKNOWLEDGEMENTS

This work was carried out in the Attention and Memory Networks (AMN) research group at the Department of Psychology, University of Helsinki. I wish to express my upmost gratitude to my supervisors Dr. Teemu Rinne and Professor Kimmo Alho for giving me the opportunity to work in their group, and for their supportive guidance throughout the work. Thank you for sharing your scientific knowledge with me, and for your generosity and patience. I am also grateful to the past and present student members of the AMN group for their contributions in scientific and other matters during the years. Mr. Juha Salmi, Mr. Ville Villberg, Ms. Anna Särkkä, Ms. Siiri Kirjavainen, Ms. Sonja Koistinen, Mr. Sebastian Cederström and Ms. Johanna Salonen, your presence was instrumental in creating a nice working atmosphere. A special thanks goes to Mr. Juha Salmi for his help in dealing with the challenges of the work.

The fMRI experiments were conducted at the Advanced Magnetic Imaging Centre (AMI Centre), Helsinki University of Technology, while the EEG data was measured at the Cognitive Brain Research Unit (CBRU), Department of Psychology, University of Helsinki, and the MEG data collected at the BioMag Laboratory, Hospital District of Helsinki and Uusimaa HUSLAB, Helsinki University Central Hospital. I would like to thank the Department of Psychology, Professor Riitta Hari, the director of AMI Centre, Professor Risto Näätänen and Professor Teija Kujala, the former and present directors of CBRU, and Dr. Jyrki Mäkelä, the director of BioMag, for the excellent research facilities. For financial support, I am indebted to the Academy of Finland, the University of Helsinki and the Finnish Graduate School of Psychology.

I give my warmest thanks to my co-authors Professor Kimmo Alho, Dr. Teemu Rinne, Mr. Juha Salmi, Dr. Johanna Pekkola, Professor Mikko Sams, Dr. Iiro Jääskeläinen, Ms. Anna Särkkä, Dr. Taina Autti and Dr. Oili Salonen for interesting and successful collaboration. I am grateful to Professor Synnöve Carlson for agreeing to be my opponent. I want to thank also the official reviewers of this thesis, Dr. Jyrki Ahveninen and Assistant Professor Christopher Stecker for constructive comments on this manuscript and Professor Christina Krause for agreeing to serve in the grading committee. An enormous thanks also goes to all the participants in the experiments who made this work possible.

I extend my thanks to colleagues at the Department of Psychology and to people elsewhere who have helped me on the way and made working more pleasant: Ms. Marja Junnonaho, Ms. Piiu Lehmus, Ms. Riitta Salminen, Ms. Erja Id, Mr. Kalevi Reinikainen, Dr. Petri Paavilainen, Dr. Markku Verkasalo, Mr. Pekka Lahti-Nuuttila, Mr. Jari Lipsanen, Mr. Pertti Keskivaara, Ms. Riikka Lovio, Dr. Minna Huotilainen, Mr. Lauri Parkkonen and many others. I especially thank Ms. Leena Wallendahr for lightening up the atmosphere at the work place many times. I am grateful to Mr. Miika Järvenpää, Mr. Markus Kalske, Mr.
Teemu Peltonen and Mr. Pasi Piiparinen for their invaluable technical help at the CBRU. Similarly, I thank Mr. Jussi Nurminen, Ms. Suvi Heikkilä and Dr. Anna Shestakova for their technical assistance at the BioMag laboratory, and Ms. Marita Kattelus and Mr. Antti Tarkiainen for their help at the AMI centre.

With overwhelming gratitude, I thank my mother Irja-Lea Degerman-Ubani, my brothers Dr. Martin Ubani and Anthony Degerman, my sister Alice Degerman and my grandparents Martta and Sten Degerman for their support and faith in me. Thank you for the encouraging examples of determination and strength. I am also grateful to my friends for taking my mind off work matters: Mr. Jukka Arppe, Mr. Väne Orava, “Suopajärven jengi”, the Pekuri family and others.

Finally, I express special gratitude to my fiancée Ms. Paula Pekuri, to whom this thesis is dedicated. Thank you for your continuous encouragement, understanding and faith in me even during the hardest phases of this work. Thank you also for the inspiration and motivation in real life.

Helsinki, October, 2008         Alexander Degerman
LIST OF ORIGINAL PUBLICATIONS


1 INTRODUCTION

Identifying relevant sounds in the environment and localizing them in space are two important functions of the auditory system. A sound is transmitted to the ear by pressure oscillations in the air at certain frequencies (i.e., a sound wave; Klinke, 1989). Therefore, the frequency content of a sound, such as the harmonic structure of a musical chord or the formant structure of a phoneme (Bendor and Wang, 2006), is essential for identifying it. Pitch is a perceptual attribute that for pure tones correlates with frequency (Bendor and Wang, 2006; Klinke, 1989). For complex harmonic sounds, pitch is computed in the auditory system using the available frequency information (Bendor and Wang, 2006; Moore, 2001; see also, Hall and Plack, 2007; Schönwiesner and Zatorre, 2008). The computation of sound location relies mainly on binaural cues, that is, differences in the timing and intensity of the sound waves arriving at the two ears (Cohen and Knudsen, 1999). For example, the sound waves caused by a word spoken from the left arrive later and are lower in intensity at the right ear than the left ear.

Auditory selective attention enables a rapid and precise selection of relevant sounds on the basis of their pitch or location (e.g., Cherry, 1953; Fritz et al., 2007; Näätänen, 1992). This active selective listening strongly modulates activity in the brain. The main focus of the present thesis was to examine the effects of voluntary auditory and audiovisual selective attention on auditory cortex activity in humans.

1.1 Processing of sound identity and location in the brain

Sound-identity cues (e.g., pitch) and sound-source location are processed in the auditory pathway and higher level temporal, parietal and frontal cortical regions. This section describes processing in the ascending auditory pathway that begins from the inner ear, projects across subcortical nuclei, and terminates in the auditory cortex.

1.1.1 The subcortical ascending auditory pathway

Sound waves entering the ear cause oscillation of the eardrum at their characteristic frequencies. The oscillation energy is transmitted by the middle-ear ossicles and oval window to the fluid in the cochlea of the inner ear. Vibration of the cochlear fluid sets the basilar membrane in motion stimulating hair cells that convert the mechanical sound signal into neural signals (e.g., Klinke, 1989). Regions of the cochlea systematically respond to sounds of a certain frequency. This provides a spatial representation of frequencies (i.e., tonotopical map) that is encoded in the auditory nerve fibers (Shamma, 2001). Frequency
information is also encoded temporally by neural firing that is phase-locked to the motion of the basilar membrane (Rose et al., 1967; Shamma, 2001). The tonotopical organization of frequency representations is preserved in the ascending auditory pathway up to the auditory cortex by neurons with frequency-specific responses (Klinke, 1989).

From the hair cells of the cochlea, the neural signals travel via the cochlear, superior olivary and lateral lemniscal nuclei, inferior colliculus and the medial geniculate body of the thalamus to the auditory cortex. The superior olivary complex in the midbrain is the first nucleus in the ascending auditory pathway at which inputs from the two ears may be compared. It contains neurons that respond selectively to sounds with a certain interaural time or intensity difference (Brand et al., 2002; Klinke, 1989; Yin and Chan, 1990). This suggests that computation of sound location begins already at an early subcortical level of information processing. Differential neural responses to binaural localization cues have also been found at the higher levels of the ascending auditory pathway, including the auditory cortex (Cohen and Knudsen, 1999; King et al., 2007; Kuwada et al., 2006; McAlpine et al., 2001; Stanford et al., 1992; Takahashi et al., 1984).

1.1.2 The auditory cortex

The auditory cortex in humans is located on the superior temporal cortex (Fig. 1a). It participates in processing sound-identity and location cues, although much of this information is processed already subcortically (e.g., Bendor and Wang, 2006; King et al., 2007). It has been proposed that neural activity in the auditory cortex does not merely reflect computation of physical sound features, but may also reflect higher-order functions (Irvine, 2007; Weinberger, 2004), such as integrative processing of auditory objects (Nelken, 2004). Moreover, the auditory cortex participates in multisensory processing (Winer and Lee, 2007), and appears to be activated by visual speech perception even in the absence of auditory sensory input (Calvert et al., 1997; Pekkola et al., 2005; Pekkola et al., 2006). Based on animal data (de la Mothe et al., 2006; Hackett et al., 1999; Nakamoto et al., 2008; Romanski et al., 1999; Suga and Ma, 2003; Winer and Lee, 2007), the auditory cortex influences both subcortical and higher-level cortical processing of auditory information through ascending and descending connections with, for instance, the thalamus and frontal and parietal cortical regions.

Anatomical and neurophysiological studies suggest that the primate auditory cortex is organized into primary and secondary regions including various subregions (Fig. 1; Brugge et al., 2008; Fullerton and Pandya, 2007; Hackett et al., 2001). Similarly, an organization of the auditory cortex into primary and secondary regions has been proposed in other mammals, such as in the cat (e.g., see, Fig. 1 Malhotra et al., 2008). Yet, the number, borders or functional properties of auditory cortex regions are not well known (Brugge et al., 2008; Fullerton and Pandya, 2007; Petkov et al., 2006).
Fig. 1. (a) A lateral surface of the human brain (left) and an axial view of the human brain (right; Collins et al., 1994; Evans et al., 1993). The position of the axial slice is indicated by the white line on the lateral brain image. An approximation of the auditory cortex in the superior temporal cortex is given in shaded red and blue. The blue color covers Heschl’s gyrus, which is the approximate landmark of the primary auditory cortex (Hackett et al., 2001; Morosan et al., 2001). The red color depicts the secondary auditory cortices presumably surrounding the primary auditory cortex. A = anterior, P/L = posterior/left. (b) A schematic illustration of the auditory “what” (green) and “where” (red) streams in primates (Rauschecker and Tian, 2000). The primary auditory cortex (core): A1 = auditory area 1, R = rostral area; the secondary auditory cortex (belt): AL = anterolateral area, ML = middle lateral area, CL = caudolateral area, CM = caudomedial area; MGd and MGv = dorsal and ventral regions of the medial geniculate nucleus of the thalamus; PB = parabelt cortex; PP = posterior parietal cortex, PFC = prefrontal cortex, T2/T3 = anterior pole of the temporal lobe. Illustration (b) reprinted from: Rauschecker, J. P. & Tian, B. (2000) Mechanisms and streams for processing of “what” and “where” in auditory cortex. Proceedings of the National Academy of Sciences U. S. A., 97, 11800–11806. Copyright (2000) National Academy of Sciences, U.S.A.
Consistent with animal data (Kosaki et al., 1997; Petkov et al., 2006; Rauschecker et al., 1995), some human functional magnetic resonance imaging (fMRI) studies (Formisano et al., 2003; Petkov et al., 2004; Talavage et al., 2004; Wessinger et al., 1997; Yang et al., 2000) have observed signs of tonotopical organization in subregions of the human auditory cortex. These tonotopic frequency representations are subject to learning-related plasticity, as suggested by fMRI and positron emission tomography (PET) results (Morris et al., 1998; Ohl and Scheich, 2005; Thiel et al., 2002). Correspondingly, animal studies (Dahmen and King, 2007; Fritz et al., 2007; Irvine, 2007; Polley et al., 2006; Recanzone et al., 1993; Rutkowski and Weinberger, 2005; Weinberger, 1995; Weinberger, 2004) have demonstrated that feature-specific responses in the auditory cortex may be modulated according to behavioral needs and salient properties of task-related stimuli. For example, Polley et al. (2006) trained rats to attend independently to either certain frequency cues or intensity cues, while presented with an identical set of auditory stimuli. The authors observed an expanded representation in the rat auditory cortex for the trained feature range (i.e., certain frequency or intensity) but no apparent change in the representation of the irrelevant feature. The degree of plastic changes in the relevant feature representations was correlated with the degree of perceptual learning in the tasks. In addition, animal studies suggest that neurons in the secondary auditory cortices respond to more complex acoustic stimulation than those in the primary auditory cortex (Petkov et al., 2006; Rauschecker et al., 1995; Tian and Rauschecker, 2004). Correspondingly, fMRI results in humans (Hall et al., 2002; Wessinger et al., 2001) have shown that complex sounds activate more widespread auditory cortex regions than simple tones.

Based on electrophysiological recordings and anatomical tract-tracing in non-human primates, it has been proposed that subregions of the auditory cortex are functionally specialized for processing sound identity and location (Kaas and Hackett, 1999; Rauschecker and Tian, 2000; Recanzone, 2000; Romanski et al., 1999; Tian and Rauschecker, 2004; see also, Lomber and Malhotra, 2008). Proposedly, neurons in the anterior auditory cortex respond primarily to identity-related information as part of a “what” stream that projects to the ventral prefrontal cortex, whereas, neurons in the posterior auditory cortex are more sensitive to sound location, and are part of a “where” stream with projections to the parietal cortex and dorsal prefrontal cortex (Fig. 1b). In line with this auditory “what” and “where” model, some human lesion studies (Clarke and Thiran, 2004) and fMRI and PET studies (Arnott et al., 2004; Rämä et al., 2004) have shown different involvement of frontal, temporal and parietal cortical regions in processing sound identity and location. However, the question of a “what”−“where” dichotomy in the human auditory cortex has remained unresolved (Cohen and Wessinger, 1999; Deouell et al., 2007; see also Section 1.3.2), although there is evidence that inferior parietal areas
Some studies in cats indicate that neurons in both anterior and posterior regions of the auditory cortex code sound location (Malhotra et al., 2008; Middlebrooks, 2002; Stecker and Middlebrooks, 2003), although there may be enhanced spatial sensitivity in the posterior regions (Stecker et al., 2005). Quite recently, Lomber and Malhotra (2008) found compelling evidence of a double-dissociation of “what” and “where” processing in the auditory cortex of the cat. The authors trained cats to discriminate between different temporal patterns of auditory stimuli or to localize the spatial position of auditory stimuli. During testing, anterior or posterior regions of the cats’ auditory cortex were deactivated by reversible cooling. It was shown that deactivation of anterior auditory field produces deficits in auditory pattern discrimination but not in sound localization, while deactivation of posterior field results in deficits in sound localization but not in pattern discrimination. This suggested that processing in the anterior auditory cortex is necessary for accurate auditory identification, while processing in the posterior auditory cortex is necessary for accurate localization of sounds.

1.2 Effects of auditory selective attention on auditory cortex activity

Perception of sounds in our environment is not merely passive interpretation of information received by the ears. Auditory perception is influenced by goal-directed behavior, such as learning and active listening (Moore et al., 2007; Palmer et al., 2007). The process in which certain information is actively selected for conscious processing, while other information is ignored, is called selective attention. Auditory selective attention (i.e., active selective listening to sounds) strongly modulates brain activity depending on the behavioral task (Fritz et al., 2007; Palmer et al., 2007). As such, auditory selective attention appears to focus neural processing on the most relevant sensory input, facilitating goal-directed behaviour.

Auditory selective attention produces enhanced activity in the auditory cortex already within 100 ms from sound onset, as indicated by source analyses of event-related brain potentials (ERPs) and magnetic fields (ERFs) recorded with electroencephalography (EEG) and magnetoencephalography (MEG), respectively (Arthur et al., 1991; Giard et al., 1988; Hari et al., 1989; Rif et al., 1991; Woldorff et al., 1993). The magnitude of the ERP attention effects in the auditory cortex depends on stimulation rate (Alho et al., 1990; Neelon et al., 2006; Näätänen, 1990; Teder et al., 1993). fMRI and PET studies have also found prominent enhancements of auditory cortex activity during auditory selective
attention (Alho et al., 1999; Grady et al., 1997; Petkov et al., 2004; Rinne et al., 2007; Woodruff et al., 1996; Zatorre et al., 1999). Based on the results of these studies, it appears that auditory selective attention activates both the primary and secondary auditory cortices. Furthermore, in general, auditory-cortex attention effects increase with task difficulty (Alho et al., 1992; Jäncke et al., 1999; O’Leary et al., 1997).

### 1.3 Näätänen’s attentional-trace theory

In ERP studies, the effects of auditory selective attention on brain activity are observed as a negative difference (Nd; Hansen and Hillyard, 1980; Näätänen et al., 2002) between ERPs to attended sounds and those to unattended sounds when measured from fronto-central scalp sites. The Nd usually has two peaks, the first one (“early Nd”) at about 100–200 ms and the second peak (“late Nd”) after 300 ms from sound onset (Alho et al., 1994; Hansen and Hillyard, 1980; Michie et al., 1993; Näätänen et al., 2002; Salmi et al., 2007a). According to Näätänen’s (1982, 1990, 1992) attentional-trace theory, the auditory Nd results from a separate attention-related response called the processing negativity (PN) consisting of a sensory-specific early component and a frontal late component.¹

#### 1.3.1 The early PN and selection of attended sounds in the auditory cortex

Näätänen (1982) proposed that the early PN reflects a matching process in which sensory input is compared with an “attentional trace”. The auditory attentional trace was described as a pattern of facilitated neurons that resides in the auditory cortex, and represents the physical stimulus feature(s) common to the attended sounds. Presumably, the auditory early PN is produced by an increase of neural responses in the auditory cortex lasting as long as the sensory input matches with the facilitated part of this brain region. The auditory cortex origin of the auditory early PN is supported by source analyses of ERPs and ERFs indicating that the main sources of the auditory early Nd and its magnetic counterpart (early Ndm) are located in the supratemporal plane (Arthur et al., 1991; Giard et al., 1988; Hari et al., 1989).

¹ Some researchers have suggested that the auditory early Nd at the latency of the N1 response (peak around 100 ms from sound onset in ERPs) or its ERF counterpart, the N1m, is simply caused by a larger N1/N1m to the attended sounds than to the unattended sounds (Hillyard et al., 1973; Rif et al., 1991; Woldorff and Hillyard, 1991; Woldorff et al., 1993; for contradictory evidence, see Näätänen, 1992; for discussion, see Näätänen et al., 2002). The auditory N1 is elicited by a rapid change in stimulus energy (e.g., sound onset; Näätänen, 1992), and its main sources are located in the auditory cortex (Näätänen et al., 2002).
According to Näätänen (1982, 1990), the attentional trace is formed and maintained with active selective rehearsal of the attended stimulus features, and supported by sensory reinforcement provided by each occurrence of the attended stimuli. The more frequently the attended stimuli are presented, the better the attentional trace is presumably maintained, and the more prominent is the resulting PN. Consistent with this proposal, some ERP studies (Alho et al., 1990; Neelon et al., 2006; Näätänen, 1990) have shown that the amplitude of attention effects at the early PN latency increases with presentation rate of attended sounds. In contrast to Näätänen’s proposal, however, others (Teder et al., 1993) have found an opposite rate-dependency of auditory attention effects at the early PN latency.

1.3.2 Attentional processing of pitch and location of sounds in the auditory cortex

ERP studies investigating the brain mechanisms of selective attention have often required participants to focus on sounds with a designated pitch, location, or both while ignoring other stimuli intermixed with the to-be-attended sounds (e.g., Alho et al., 1994; Hansen and Hillyard, 1980; Hillyard et al., 1973; Näätänen et al., 1978). Results of these ERP studies indicate that selectively attending to either sound feature produces prominent attention effects in the auditory cortex.

In his review, Näätänen (1990) hypothesized that because an attentional trace “involves an area in the auditory cortex specific to the feature represented by this trace”, “its location may differ as a function of whether stimuli are selected, for example, on the basis of pitch or spatial position. In addition, if an attentional trace is two-dimensional (i.e., if the relevant stimuli are defined by two features) it is, presumably, distributed to two loci in the auditory cortex” (Näätänen, 1990, p. 223–224). As suggested by the auditory “what” and “where” model (e.g., Rauschecker and Tian, 2000), attention to the pitch of sounds and attention to their location should activate mainly the anterior and posterior auditory cortex, respectively.

Consistent with the auditory “what” and “where” model, previous ERP studies using two-dimensional attention tasks (Woods and Alain, 1993; Woods et al., 1994; Woods and Alain, 2001) have found a more anterior scalp distribution for the pitch-related early Nd than the location-related early Nd, indicating different generators for these two attention effects in the auditory cortex. Similarly, ERF studies (Ahveninen et al., 2006; Anourova et al., 2001) and some fMRI (Alain et al., 2001; Barrett and Hall, 2006; Maeder et al., 2001; Obleser et al., 2006; Warren and Griffiths, 2003) and PET (Zatorre et al., 2004) studies have shown that attention-related processing of sound identity activates
especially the antero-lateral auditory cortex, and processing of sound location especially the postero-medial auditory cortex. However, such “what” and “where” segregation in the auditory cortex is challenged by results in other studies.

For example, Alho et al. (1994), using one-dimensional (i.e., pitch or location) attention tasks, found no scalp distribution differences between the pitch-related and location-related early Nds in ERPs. Correspondingly, several fMRI and PET studies (Alain et al., 2005; Arnott et al., 2004; Obleser et al., 2007; Zatorre et al., 1999; Zatorre et al., 2002; see also Barrett and Hall, 2006) and human lesion data (Zatorre and Penhune, 2001) indicate that anterior and posterior auditory cortex areas play a role in attention-related processing of both sound identity and location. Moreover, human behavioral experiments (Mondor et al., 1998) have demonstrated that target detection based on the pitch of sounds is affected by modulation of task-irrelevant sound location, and vice versa. Based on these behavioral results and the PET results of similar pitch-related and location-related attention effects in different cortical regions (Zatorre et al., 1999), it was suggested that attention cannot be directed independently to pitch or location of sounds (Mondor et al., 1998; Zatorre et al., 1999). Instead, the attentional processing of these two auditory features may be integrated, and facilitate neural responses in similar cortical areas (Zatorre et al., 1999).

1.3.3 The role of frontal and parietal cortical areas in auditory attention

Näätänen (1982, 1990) proposed that the late Nd between ERPs to attended and unattended sounds is caused by a late PN component possibly generated in the frontal cortex. The late PN was hypothesized (Näätänen, 1982) to reflect further processing or selective rehearsal of the attended stimuli. The importance of the frontal cortex in generation of the late PN is supported by results of ERP scalp distribution analysis (Giard et al., 1988) showing that the auditory late Nd has its negative maximum at frontal sites. In addition, ERP studies have demonstrated an attenuated auditory late Nd in patients with frontal cortex lesions compared to the late Nd in healthy participants (Knight et al., 1981; Näätänen et al., 2002). These ERP results coincide with those of fMRI and PET studies in healthy participants (Alho et al., 1999; Salmi et al., 2007b; Tzourio et al., 1997; Wu et al., 2007; Zatorre et al., 1999) indicating that auditory attention modulates activity in several frontal cortex regions, including the superior, middle and inferior frontal gyri.

The parietal cortex is also involved in auditory selective attention, as suggested by ERP results of an attenuated late Nd in patients with lesions in the temporo-parietal cortex (Woods et al., 1993). In addition, patient studies have shown that unilateral lesions in the inferior parietal cortex as well as in areas of the frontal cortex may be associated
with neglect characterized by deficits in directing spatial attention to the contralesional hemifield and maintaining attention there (Heilman and Valenstein, 1972; Mesulam, 1999). Correspondingly, previous fMRI and PET studies in healthy participants have found that parietal regions, such as the inferior and superior parietal lobules and precuneus, are involved in shifting and maintenance of attention (Salmi et al., 2007b; Shomstein and Yantis, 2004; Wu et al., 2007; Zatorre et al., 1999). Thus, it appears that areas in both the frontal and parietal cortices may participate in auditory attention-controlled processing (Alho et al., 1999; Driver and Frackowiak, 2001; Näätänen, 1992; Näätänen, 1990; Näätänen et al., 2002; Wu et al., 2007).

1.3.4 Audiovisual selective attention

Attention may be important for successful integration of stimulus features into unified percepts of objects. This is suggested by human behavioral data showing, for instance, that distraction and high attentional demands reduce accuracy of audiovisual integration in speech perception (Alsius et al., 2005; Tiippana et al., 2004; see also, Treisman and Gelade, 1980). Correspondingly, other behavioral data indicate that attention to sounds in a certain location facilitates processing of visual stimuli presented in the same location, and vice versa (Driver and Spence 1998; McDonald, Teder-Sälejärvi, Hillyard 2000). These behavioral results suggest cross-modal links in attentional processing. However, human brain mechanisms of bimodal (e.g., auditory and visual) selective attention are not yet well understood.

Näätänen’s (1982, 1990, 1992) attentional-trace theory focuses on selection of stimuli within a single sensory modality (e.g., audition or vision), that is, unimodal selective attention. Although the theory suggests the possibility that stimuli across sensory modalities may be matched against a unimodal attentional trace, it leaves open the question of how the attentional trace mechanism operates during audiovisual attention requiring integration of information in two sensory modalities. Based on a possible analogue in unimodal processing, it might be that audiovisual selective attention engages a two-dimensional attentional trace with its two loci in the auditory and visual cortices for processing the respective sensory-specific information. This is supported by ERP and fMRI studies (Busse et al., 2005; Driver and Spence, 1998; Eimer and Schröger, 1998; Hillyard et al., 1984; Molholm et al., 2007; Teder-Sälejärvi et al., 1999) indicating that spatial and temporal congruence of auditory and visual stimuli during attention to stimuli in one of the sensory modalities can automatically lead to audiovisual attention, and facilitation of neural responses in both auditory and visual cortices.
Audiovisual selective attention may also utilize multisensory cortical representations. This is suggested by ERP studies demonstrating that attention-related integration of auditory and visual features produces specific activity in the so-called sensory-specific auditory and visual cortices, and in multimodal frontal, temporal and parietal regions (Fort et al., 2002b; Giard and Peronnet, 1999; Molholm et al., 2006; Talsma et al., 2007). Similarly, fMRI and PET studies using various levels of control for participants’ attention have found activity associated with audiovisual integration in brain regions such as the auditory and visual cortices, the superior and middle frontal gyri and the superior and inferior parietal lobules (Calvert et al., 1999; Calvert et al., 2000; Calvert et al., 2001; Lehmann et al., 2006; Saito et al., 2005; Sekiyama et al., 2003; Wright et al., 2003). The involvement of auditory, visual, frontal and parietal cortical areas in integration of audiovisual features is supported by animal data showing that these areas contain multimodal neurons and are interconnected (Fuster et al., 2000; Ghazanfar et al., 2005; Mazzoni et al., 1996; Meredith, 2004; Mountcastle, 1978; Schroeder and Foxe, 2005; Vaadia et al., 1986; Winer and Lee, 2007).
2 AIMS OF THE PRESENT THESIS

The aim of the present thesis was to study the effects of selective auditory attention (Studies I–IV) and audiovisual attention (Study IV) on brain activity in humans. fMRI (Studies I, II and IV), ERP (Study III: Exp 1) and ERF (Study III: Exp 2) methods were used (see, Section 3.1). The specific aims of Studies I–IV are described below.

Study I aimed to determine with fMRI how auditory cortex activity is modulated by attention when sound presentation rate is systematically varied. Based on previous ERP results described in Section 1.3.1, and on fMRI results demonstrating rate-dependency of auditory cortex activity (e.g., Binder et al., 1994; Harms and Melcher, 2002), it was hypothesized that both attention and increasing stimulation rate enhance activity in the auditory cortex. In addition, it was examined whether the effects of attention and stimulation rate interact.

The aim of Studies II and III was to determine whether selective attention to pitch or location of sounds enhances activity in different regions of auditory cortex. Previous brain research studies using various experimental designs have found conflicting results on this matter (Section 1.3.2). Therefore, in order to increase the ability to detect possible differences between the pitch-related and location-related attention effects in the auditory cortex, fMRI, ERP and ERF data were collected using a similar experimental design.

Study IV aimed at using fMRI to determine which brain areas are involved in audiovisual selective attention requiring integration of information from two sensory modalities. Based on previous brain research results (Section 1.3.4), it was hypothesized that audiovisual attention activates the auditory and visual cortices and multimodal regions involved in integrating audiovisual information.
3 METHODS AND RESULTS

3.1 Overview of brain research methods used in Studies I–IV

3.1.1 fMRI

fMRI measures non-invasively changes in blood flow and blood oxygenation (i.e., hemodynamic changes) associated with neural activity (Logothetis, 2007). In blood oxygenation level dependent (BOLD) fMRI, the main signal arises from behavior of hydrogen nuclei of the brain under a strong magnetic field (e.g., 3 Tesla; Heeger and Ress, 2002). After perturbation by radio-frequency (RF) pulses, these nuclei emit the absorbed RF energy until returning to their equilibrium state (relaxation; Matthews, 2001). The emitted RF energy is detected by the RF coil of the MRI scanner. This allows for construction of anatomical brain images based on distinctive density and relaxation times of hydrogen nuclei in different tissues (Jezzard and Clare, 2001). The BOLD signal reflects local variations in inhomogeneity of the magnetic field, which are caused by changes in the concentration of deoxygenated hemoglobin in the blood during oxygen metabolism. These inhomogeneities affect the relaxation times of nearby hydrogen nuclei, and the emitted RF energy detected by the MRI scanner (Matthews, 2001). Peak BOLD signal changes in the auditory cortex induced by acoustic stimuli are typically around few percent above the baseline (Hall et al., 2000). In general, BOLD fMRI provides high spatial resolution (1–10 mm) in studying neural activity in the brain (Matthews, 2001). The exact way in which neural activity is associated with changes in local hemodynamics, however, is not known (Heeger and Ress, 2002; Logothetis, 2007; Ugurbil et al., 2003).

The temporal resolution of BOLD fMRI is generally low as compared to the time course of neural activations. This is because the BOLD response takes several seconds to evolve after stimulus onset (Menon and Goodyear, 2001; Miezin et al., 2000). In addition, commonly a blocked fMRI design is used where experimental stimuli are presented in task blocks lasting from tens of seconds to a minute, and mean activity across an entire block is compared with that of another block (Donaldson and Buckner, 2001). Such a design can provide robust attention-related activity. An event-related fMRI design makes it possible to examine BOLD responses to events separated by a few seconds (Donaldson and Buckner, 2001). However, because the evolution of the BOLD response takes significantly longer (seconds) than the evolution of neural activity itself (milliseconds), a more accurate detection of the temporal dynamics of brain activity requires the use of electrophysiological measures.
3.1.2 EEG and MEG

EEG and MEG provide high temporal resolution for measuring non-invasively electrical and electromagnetic neural activity in the human brain. The electroencephalogram is recorded with electrodes attached to the scalp, and depicts the electric potential difference between two electrodes as a function of time (Luck, 2005). The amplitude range of this potential fluctuation is about 1 mV. MEG measures the small magnetic fields produced by electrical currents within the brain (amplitude range 1 pT; Hämäläinen et al., 1993). Both, EEG and MEG signals are mainly generated by synchronous postsynaptic potentials in large groups of pyramidal cells (Hämäläinen et al., 1993; Luck, 2005; Okada, 2003; Picton et al., 1995). For studying brain function, the EEG or MEG signals are commonly averaged across several presentations of experimental stimuli in order to reveal the ERPs and ERFs time-locked to processing the stimuli (Hämäläinen et al., 1993; Luck, 2005; Picton et al., 1995). This allows for studying information processing in the brain at a millisecond time scale.

EEG and MEG are rather limited in spatial resolution. Under optimal conditions, the spatial resolution of these measures is on the order of several millimeters (Matthews, 2001). In general, the EEG signal is distorted by different conductivities of the anatomical structures (e.g., brain tissue, scull and scalp) it passes through, which makes it difficult to segregate simultaneous activity of different sources (Luck, 2005; Picton et al., 1995). MEG is usually more sensitive in detecting differences between sources than EEG, because the MEG signal is not affected by conductivities of the head structures (Hämäläinen et al., 1993; Okada, 2003). However, MEG as well as EEG source localization is limited by the signal-to-noise ratio of the recordings and by the location of the recording sites relative to the sources (Kaufman and Lu, 2003). To obtain higher spatial resolution, EEG studies may utilize dense electrode arrays of over 100 electrodes (Gevins et al., 1995; Gevins et al., 1999), and MEG instruments may have hundreds of sensors covering the whole scalp area (Lounasmaa and Hari, 2003). In addition, source estimation can be performed using realistically shaped head models constructed on the basis of individual MR images (Darvas et al., 2004; Luck, 2005). Moreover, combined EEG and MEG source modeling can be used, allowing for more accurate detection of, for instance, deeper sources in the brain (Molins et al., 2008).

The EEG and MEG source localizations rely on extra-cranially recorded signals in identifying sources within the brain. This is problematic because theoretically a signal with a particular extracranial distribution may be generated by an infinite number of different source configurations (Luck, 2005; Picton et al., 1995). Often, this so-called inverse problem is solved with source models constrained based on prior anatomical or neurophysiological knowledge (Luck, 2005). Sources may be modeled as equivalent
current dipoles if it is assumed that the activity is generated in a few restricted brain regions (Cuffin, 1998; Darvas et al., 2004; Kucukaltun-Yildirim et al., 2006). Distributed source models, such as the minimum norm estimate (MNE) of source currents (Hämäläinen and Ilmoniemi, 1994; Komssi et al., 2004), may be used to represent the pattern of neural activity with a large number of dipoles covering the whole brain (Darvas et al., 2004; Kucukaltun-Yildirim et al., 2006; Luck, 2005). Such distributed models have the benefit of not imposing strict constraints on the number of sources in the model. They can also be constructed by constraining the source currents to the cortex (Lin et al., 2006), where the main sources of the EEG and MEG signals are located. Furthermore, MEG source current estimates may be weighted toward hemodynamically activated areas measured with fMRI (Dale et al., 2000).

### 3.2 Participants in Studies I–IV

In Studies I–IV, participants were healthy right-handed adults with normal hearing and normal or corrected-to-normal vision. Details of the participants in each study are given in Table 1. Seven of the participants in the fMRI experiment of Study II also took part in the ERP and ERF experiments of Study III, but ERP data of one of these participants were excluded from the analyses because of extensive artifacts of extracerebral origin in her electroencephalogram. All participants gave written informed consent prior to testing in accordance with the experimental protocol approved by the Ethical Committee of the Hospital District of Helsinki and Uusimaa.

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Males</th>
<th>Age (mean)</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>12</td>
<td>7</td>
<td>18–45 (27)</td>
<td>fMRI</td>
</tr>
<tr>
<td>II</td>
<td>10</td>
<td>5</td>
<td>19–46 (25)</td>
<td>fMRI</td>
</tr>
<tr>
<td>III: Exp 1</td>
<td>16</td>
<td>8</td>
<td>19–47 (26)</td>
<td>ERPs</td>
</tr>
<tr>
<td>III: Exp 2</td>
<td>11</td>
<td>7</td>
<td>20–48 (28)</td>
<td>ERFs</td>
</tr>
<tr>
<td>IV</td>
<td>12</td>
<td>4</td>
<td>18–31 (26)</td>
<td>fMRI</td>
</tr>
</tbody>
</table>

Exp = experiment, N = number of participants, Age is in years, Method is the brain research measure used in the study.
3.3 Stimuli and procedures in Studies I–IV

**Auditory stimuli.** The auditory stimuli in Studies I–IV were harmonic sounds. These sounds were chosen in order to produce prominent activity in the auditory cortex (Hall et al., 2002). Study I used binaurally presented sounds that had a fundamental frequency (F0) of 186 Hz with five harmonics (372, 558, 744, 930 and 1116 Hz) of equal intensity. Studies II–IV used sounds of two pitches (high and low) presented monaurally (Studies II and III) or binaurally (Study IV). The high sounds had an F0 of 1500 Hz with four harmonics (3000, 4500, 6000 and 7500 Hz), and the low sounds had an F0 of 150 Hz with four harmonics (300, 450, 600 and 750 Hz) of equal intensity. In all studies, there were frequent and infrequent sounds presented in a random order. In Study I, the duration of sounds was 200 ms, but there was a 3% frequency glide starting at 150 ms from sound onset. The frequent sounds had an upward frequency glide and the infrequent sounds had a downward frequency glide. In Studies II–IV, the high and low frequent sounds had a duration of 150 ms, while the duration of the infrequent high and low sounds was 50 ms. In Studies I, II and IV, the sounds were delivered through earplugs via headphones, and in Study III, via headphones (Exp 1) or plastic tubes and earpieces (Exp 2). The estimated effective level of the sounds at the eardrum was 75 dB SPL (Study I), 60 dB SPL (Studies II and III) or 70 dB SPL (Study IV).

**Visual stimuli.** In all studies, the visual stimuli were randomized colored circles with a diameter of approximately 3.5°. The filled circles were presented at the center of the visual field on a gray background. In Study I, the circles were each presented for 100 ms and contained a color change 50 ms from their onset: frequent circles changed from yellow to orange and infrequent circles from red to orange. In Studies II–IV, there were blue and red circles. The frequent circles of each color had a duration of 150 ms and the infrequent circles of each color a duration of 50 ms. In all studies, a fixation mark was presented embedded in the circles or alone. The visual stimuli were projected onto a mirror fixed to the head coil of the fMRI scanner (Studies I, II and IV), or viewed from a computer screen (Study III: Exp 1) or projector screen (Study III: Exp 2). The presentation of the circles and sounds was asynchronous in all other studies, except Study IV. The stimulation parameters in different Studies are summarized in Table 2.

**Procedure.** All studies consisted of attention conditions (see, Table 3) during which the participants were instructed to focus on the fixation mark, attend to designated stimuli, and respond to infrequent target stimuli \( P = 0.05 \) appearing among the attended stimuli. The to-be-attended stimuli were indicated by a simultaneous auditory and visual instruction presented whenever the attention task changed (e.g., in Study II: “right-ear sounds”, “high sounds”, “blue circles” etc., in Finnish). In unimodal auditory-attention and visual-attention conditions of Studies I–IV the target was a designated auditory or
visual infrequent stimulus, respectively. In the audiovisual-attention conditions of Study IV, the target was a designated combination of auditory and visual infrequent stimuli. In Studies I, II and IV and in Experiment 1 of Study III, the participants responded to targets by pressing a response button with their left index finger. In Experiment 2 of Study III, the participants’ responses were detected with thumb-movement electromyograms recorded from their left hand. A response was defined as a hit if it occurred 200–1100 ms from target onset, otherwise it was classified as a false alarm. Hit rates were calculated as the number of hits divided by the number of targets (Studies I–IV) while, in Studies II–IV, false-alarm rates were calculated by dividing the number of false-alarms by the number of all responses.

<table>
<thead>
<tr>
<th>Table 2. Stimulation parameters in Studies I–IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>II–IV</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

F0 = fundamental frequency
Overall, to ensure that changes in neural activity were associated with changes in selective attention, the following steps were taken. First, the behavioral responses were collected in all conditions of Studies I–IV to ensure that the participants indeed performed the designated attention task. Second, the fMRI experiments of Studies I, II and IV used active baseline conditions to avoid effects of uncontrolled activity in a passive baseline condition, such as a silent rest period (see, Alho et al., 2006), on fMRI results. Third, the attention conditions in different Studies (I–IV) were designed to produce an approximately equal number of manual responses so that brain activity associated with them would be subtracted in between-condition comparisons. Fourth, the experimental stimulation was similar in different conditions of Studies I-IV (except for stimulation rate differences between conditions of Study I, the visual-attention condition with no sounds in Study II and Study III: Exp 1, and the mental counting condition with no audiovisual stimuli in Study IV) so that between-conditions differences would reflect task-related rather than stimulus-dependent effects.

<table>
<thead>
<tr>
<th>Table 3. Attention conditions in Studies I–IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
</tr>
<tr>
<td>Conditions</td>
</tr>
<tr>
<td>Blocks in a Condition</td>
</tr>
<tr>
<td>Duration of Blocks</td>
</tr>
<tr>
<td>Exp = experiment</td>
</tr>
</tbody>
</table>

Overall, to ensure that changes in neural activity were associated with changes in selective attention, the following steps were taken. First, the behavioral responses were collected in all conditions of Studies I–IV to ensure that the participants indeed performed the designated attention task. Second, the fMRI experiments of Studies I, II and IV used active baseline conditions to avoid effects of uncontrolled activity in a passive baseline condition, such as a silent rest period (see, Alho et al., 2006), on fMRI results. Third, the attention conditions in different Studies (I–IV) were designed to produce an approximately equal number of manual responses so that brain activity associated with them would be subtracted in between-condition comparisons. Fourth, the experimental stimulation was similar in different conditions of Studies I-IV (except for stimulation rate differences between conditions of Study I, the visual-attention condition with no sounds in Study II and Study III: Exp 1, and the mental counting condition with no audiovisual stimuli in Study IV) so that between-conditions differences would reflect task-related rather than stimulus-dependent effects.
3.4 fMRI imaging and data analyses in Studies I, II and IV

The fMRI scanning in Studies I, II and IV was performed using a 3.0-T GE Signa scanner and a quadrature head coil. Functional gradient-echo planar (EPI) MR images (TR, 2800 ms; TE, 32 ms; flip angle, 90°; voxel matrix, 64 x 64; in-plane resolution 3.4 mm x 3.4 mm) were acquired with an imaging area consisting of 28 contiguous 3.4-mm thick axial oblique slices (Studies I and IV) or 28 4.0-mm thick axial oblique slices with a 1-mm inter-slice gap (Study II). In Studies I and IV, the lowest slice was positioned approximately 2 cm caudal to the AC–PC line. In Study II, the imaging area covered the whole brain. A blocked fMRI design was used in all Studies. The conditions included several blocks (Table 3) presented in semi-randomized order. For each participant, 60 functional volumes in each condition were acquired. In addition, a T1-weighted inversion recovery spin-echo volume was acquired for anatomical alignment. The T1 image acquisition used a denser in-plane resolution (matrix 256 x 256), but otherwise the same slice prescription as the functional image acquisition.

The fMRI data were analyzed using fMRI Expert Analysis Tool software (FEAT; www.fmrib.ox.ac.uk/fsl; Smith et al., 2004). To allow for initial stabilization of the fMRI signal, the first five volumes were excluded from data analyses. The data were motion-corrected, spatially smoothed with a Gaussian kernel (5 mm; FWHM), and high pass filtered (cutoff 676 s, 80 s and 132 s in Studies I, II and IV, respectively). Statistical analyses were performed using the FMRIB Improved Linear Model (FILM). The hemodynamic response was modeled using a gamma-function (mean lag 6 s, SD 3 s) and its temporal derivative. The model was high-pass filtered the same way as the data. Several contrasts were specified to create individual Z-statistic images. For group analyses, the individual Z-statistic images were transformed into a standard brain (MNI152; Montreal Neurological Institute). The Z-statistic images for the attention-related modulations were thresholded with \( Z > 2.3 \) (Study I), \( Z > 3.1 \) (Study II), or \( Z > 3.5 \) (Study IV) with a corrected cluster significance threshold of \( P < 0.05 \) (Studies I, II and IV). In addition, region-of-interest (ROI) analyses were conducted to determine mean percent signal changes in several cortical regions during different conditions.

3.5 EEG and MEG recording, and data analyses in Study III

In Experiment 1 of Study III, the electroencephalogram (bandwidth 0–134 Hz, sampling rate 512 Hz) was recorded with 130 scalp-attached electrodes. An electrode placed at the nose served as a common reference (calculated offline). Eye movements and blinks were monitored by recording electro-oculogram (EOG) with electrodes attached to the outer canthi of the eyes and above and below the right eye. The electroencephalogram was
digitally filtered (passband 0.1–20 Hz) and epoched starting 100 ms before and ending 800 ms after each stimulus onset. In each block, the epochs for the first two auditory stimuli and epochs with EEG or EOG amplitude exceeding ±150 µV at any electrode were excluded from further analyses, because so large changes probably result from extracerebral artefacts such as eye movements, blinks, or muscle activity. ERPs were averaged separately for attended and unattended frequent sounds in different auditory attention conditions. For comparing scalp distributions of attention effects, mean Nd amplitudes were measured at different latencies from attended – unattended ERP difference waves at 25 electrode sites (Fig. 6a, right).

In Experiment 2 of Study III, magnetoencephalogram (MEG) was recorded with a 306-channel whole head magnetometer (passband 0.01–200 Hz, sampling rate 601 Hz). The MEG device contains 204 planar gradiometers and 102 magnetometers, but only the former were used in the analyses. MEG data with deflections exceeding ±200 µV at EOG channels, 10000 fT/cm at magnetometers, or 5000 fT/cm at gradiometers were rejected. ERFs were analyzed offline in a similar fashion as the ERPs (epoch: -100–800 ms; baseline: -100–0 ms; bandpass filter: 0.1–20 Hz; Nd: attended sound ERF – unattended sound ERF). ERF minimum-norm estimation (MNE) was conducted to estimate Nd source-current distributions in the auditory cortex. The MNE solution was calculated over a 30-ms time window centered separately for each participant at the mean global field power peak found at the latencies of 150–250 ms and 400–500 ms for the early Nd and late Nd, respectively. The source analyses were performed using realistically-shaped head models based on individual MR images, and the MNEs constrained to the reconstructed cortical surface. MNE amplitudes of the Nd responses were measured in 9 lateral and 9 medial ROIs set in the auditory cortex of each hemisphere (Fig 8, bottom right).

Statistical analyses. In all Studies (I–IV), the between-condition differences in performance and brain responses measured within ROIs (Studies I–IV) or from the scalp (Study III: Exp 1) were tested using repeated-measures analyses of variance (ANOVAs), t-tests or Newman-Keuls tests.

3.6 Study I. Modulation of auditory-cortex activation by sound presentation rate and attention

3.6.1 Specific experimental setting and data analyses

The experiment consisted of 6 auditory-attention and 6 visual-attention conditions. Frequent and infrequent sounds and circles (see, Table 2) were presented in independent streams during all other conditions, except for one auditory-attention and one visual-
attention condition in which the auditory stimulation consisted of only the infrequent sounds. The presentation rate of the binaural sounds in different conditions was 0.1, 0.5, 1, 1.5, 2.5, or 4 Hz, and the average presentation rate of the circles in all conditions 1 Hz.

In the fMRI data analyses, the effects of attention and stimulation rate on auditory cortex activity were studied by contrasting each of the 6 auditory-attention conditions and five of the visual-attention conditions with a baseline condition (visual-attention condition with the infrequent-sound presentation), or with each other. In addition, auditory cortex activity was studied within ROIs defined based on activity clusters obtained in the comparison of all conditions vs. the baseline.

3.6.2 Results

Performance. On average 76% (SEM ± 5%) of the targets in different conditions were correctly detected. Performance in the visual task was more accurate but slower than in the auditory task (hit rate: visual task 81 ± 5%, auditory task 71 ± 6%, ANOVA: main effect of TASK: $F(1,10) = 7.84, P < 0.05$; reaction time: mean difference 62 ± 14 ms, main effect of TASK: $F(1,10) = 46.57, P < 0.001$). The rate of sound presentation had no systematic effect on the accuracy or speed of target detection.

fMRI results. As expected, both, attention to sounds ($F(1,11) = 21.9, P < 0.01$) and increasing sound presentation rate ($F(4,44) = 52.2, P < 0.001$) enhanced activity bilaterally in the auditory cortex (Fig. 2a). In addition, there was a significant interaction ($F(4,44) = 2.8, P < 0.05$) between attention and presentation rate, that is, the attention effects were larger at higher stimulation rates (Fig. 2b).
Fig. 2. Results of study I. (a) Auditory-cortex areas showing significant $(N = 12; \text{threshold: } Z > 2.3, \text{corrected cluster threshold } P < 0.05)$ activity associated with auditory attention (left) and increased sound presentation rate (right). The activity is projected onto a standard brain (MNI152; Montreal Neurological Institute). Data from the 4-Hz auditory stimulation condition is shown. L/A = left/anterior; R = right. (b) Mean percent signal changes ($\pm$SEM) in the auditory cortex of each hemisphere. Both, auditory attention and increasing sound presentation rate enhanced auditory cortex activity. In addition, the attention effects were larger at higher stimulation rates. STC: superior temporal cortex, attAud = attend auditory, attVis = attend visual. (b) is from: Rinne, T., Pekkola, J., Degerman, A., Autti, T., Jääskeläinen, I.P., Sams, M. & Alho, K. (2005). Modulation of auditory cortex activation by sound presentation rate and attention. Human Brain Mapping, 26, 94–99, reprinted with permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.
3.7 Study II. Selective attention to sound location or pitch studied with fMRI

3.7.1 Specific experimental setting and data analyses

The experiment included 13 attention conditions (Table 3; Fig. 3). In four different pitch-attention conditions, high and low sounds were delivered randomly to one ear (left or right), and the participants attended to the high or low sounds. In four different location-attention conditions, sounds of a constant pitch (high or low) were randomly presented to the left and right ears, and the participants attended to the left-ear or to the right-ear sounds. Blue and red circles were presented during all auditory-attention conditions. In five visual-attention conditions, the participants attended to circles with a designated color. The sounds in these conditions were either like in the pitch-attention or location-attention conditions, or there was no auditory stimulation. The sounds and circles were presented in independent streams, the offset-to-onset interval randomly varying between 300 and 600 ms, in 50-ms steps.

In the fMRI analyses, the auditory attention-related modulations were obtained by contrasting the pitch-attention and location-attention conditions, respectively, with the visual-attention conditions (with the same auditory and visual stimulation). In addition, differences between attention to pitch and attention to location were determined with direct comparisons between the conditions. In all contrasts, the visual-attention condition without auditory stimulation served as the baseline. ROI analyses were conducted using spherical ROIs with a diameter of 8 mm (Fig. 5, bottom). Four ROIs covered left superior temporal gyrus (i.e., auditory cortex; ROI 3), right inferior parietal lobule (ROI 6), and bilateral middle frontal gyrus (premotor/supplementary motor; ROIs 2 and 5) activation maxima obtained in the comparison of location-attention conditions vs. pitch-attention conditions (see, Fig. 4c). Further two ROIs were set at right middle frontal (ROI 4) and left inferior frontal gyrus (prefrontal; ROI 1) attention-related activation maxima produced by attention to location (Fig. 4b).
Fig. 3. Experimental design in Study II. (Top) The experiment consisted of 13 different 33.6-s task blocks presented 5 times in a semi-randomized order. The participants’ task was to attend to a designated sound location or pitch, or to circles of a certain color, and to press a response button for the shorter-duration targets appearing among the attended stimuli. Before each block, a 2-s auditory and visual instruction indicated which task to perform. Each block was followed by a 4-s period during which only the fixation cross was shown. (Bottom) Independent streams of sounds and pictures were presented (offset-to-onset interval 300–600 ms) during every task block, except for the Pics-only block during which only pictures were shown. In the location tasks, the pitch of sounds was constant, and vice versa. Low-left: Attend to Low Sounds (high and low sounds presented to the left ear); Pics-right: Attend to Pictures (high and low sounds presented to the right ear); Left-high: Attend to Left Ear (high sounds presented to opposite ears) etc. From: Degerman, A., Rinne, T., Salmi, J., Salonen, O. & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. Brain Research, 1077, 123–134, with permission of Elsevier, Inc.
3.7.2 Results

Performance. There were no significant differences in the accuracy of target detection between the pitch-attention (hit rate: 71 ± 3%; false-alarm rate: 17 ± 2%) and location-attention conditions (hit rate, 71 ± 4%; false-alarm rate, 16 ± 2%). Nor were there significant differences in the speed of target detection between the pitch-attention (reaction time: 765 ± 21 ms) and location-attention conditions (750 ± 16 ms).

fMRI results. Both selective attention to the pitch of sounds and attention to their location enhanced activity in widespread areas of the auditory cortex in both hemispheres, the left prefrontal cortex and the temporo-parietal cortex bilaterally in comparison with visual-attention conditions with the same sounds (Figs. 4 and 5). Activity in the left posterior auditory cortex and the right inferior parietal cortex was stronger during attention to location than during attention to pitch. In addition, attention to location but not attention to pitch significantly enhanced activity in the premotor/supplementary motor cortices of both hemispheres and the right prefrontal cortex. No brain areas showed stronger activity during attention to pitch than during attention to location.
Fig. 4. Results of Study II. (a–c) Areas of significant ($N = 10$; threshold: $Z > 3.1$, corrected cluster threshold $P < 0.05$) activity projected (depth 0–1.5 cm) onto a standard brain (Collins et al., 1994; Evans et al., 1993). (a) Attention-related modulations during attention to the pitch of sounds. (b) Attention-related modulations during attention to the location of sounds. (c) Brain areas showing stronger attention-related activity during attention to location than attention to pitch. L = left, R = right. Modified from: Degerman, A., Rinne, T., Salmi, J., Salonen, O. & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. Brain Research, 1077, 123–134, with permission of Elsevier, Inc.
Fig. 5. Results of Study II. Mean percent signal changes (SEM; \(N = 10\)) during attention to pitch and attention to location of sounds in the inferior (ROI 1) and middle frontal cortices (ROIs 2, 4 and 5) and in the temporo-parietal cortices (ROIs 3 and 6). Pitch = attention-related modulations during attention to the pitch of sounds, Loc = attention-related modulations during attention to the location of sounds. The asterisks depict p-values of one-tailed t-tests (mean percent signal changes compared with zero percent, \(*P < 0.05\), \(**P < 0.01\), \(***P < 0.001\)). Attention to location produced significantly \((P \leq 0.05)\) stronger activity than attention to pitch in all studied areas, except the left prefrontal cortex. The approximate position of spherical regions of interest (ROIs 1–6) used for analyzing signal changes in these areas are shown at the bottom of the figure projected onto the brain surface. Modified from: Degerman, A., Rinne, T., Salmi, J., Salonen, O. & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. *Brain Research*, 1077, 123–134, with permission of Elsevier, Inc.
3.8 Study III. Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields

3.8.1 Specific experimental settings and data analyses

The ERP and ERF experiments were conducted in different sessions, and used a similar experimental design as the fMRI experiment in Study II (see above, Section 3.7.1). Differing from the fMRI experiment, however, the auditory stimulus offset-to-onset interval in the ERP and ERF experiments randomly varied between 300 and 600 ms, and the visual stimulus offset-to-onset interval between 307 and 607 ms, in 10-ms steps. In addition, the ERF experiment consisted of only the eight auditory-attention conditions (with the to-be-ignored visual stimuli) used in the fMRI experiment.

In data analyses, ERPs and ERFs were extracted from the EEG and MEG signals, respectively. Nds in the pitch-attention and location-attention conditions were obtained by subtracting the ERPs to unattended sounds from the ERPs to attended sounds, and Ndm responses similarly by comparing ERFs to attended and unattended sounds. For studying differences between the pitch-related and location-related attention effects, scalp distribution analyses in ERPs and MNE source analyses in ERFs were conducted. The ERP scalp distribution analyses used a 25-electrode array (Fig. 6a, right), while the ERF MNE analyses used 9 lateral and 9 medial ROIs in the auditory cortex of each hemisphere (Fig. 8, bottom right).

3.8.2 Results

Performance. In the ERP experiment, there were no significant differences in the accuracy of target detection based on false-alarm rates (pitch conditions, 7 ± 1%; location conditions, 6 ± 1%). However, hit rates were significantly lower during attention to pitch (87 ± 2%) than during attention to location (91 ± 2%; \( F(2,28) = 20.77, P < 0.001 \) for differences between the pitch-attention condition, location-attention condition and visual-attention condition with sounds, 79 ± 3%; Newman-Keuls tests: \( P < 0.05 \) for differences between conditions). There were no significant differences in the speed of target detection between the pitch-attention (reaction time: 562 ± 16 ms) and location-attention conditions (559 ± 16 ms).

In the ERF experiment, there were no significant differences in hit rates or reaction times between the pitch-attention (hit rate: 78 ± 4%, reaction time: 449 ± 19 ms) and location-attention (hit rate: 80 ± 4%, reaction time: 458 ± 17 ms) conditions. However, mean false-alarm rates were significantly higher in the pitch-attention conditions (14 ± 2%) than location-attention conditions (9 ± 2%; two-tailed \( t(10) = 3.11, P < 0.05 \)).
ERP scalp distribution analyses. Both, attention to pitch and attention to location produced prominent Nds (Fig. 6a). The early Nds and late Nds peaked at 200–300 ms and after 400 ms from sound onset, respectively. The scalp distribution analyses showed that, at 150–200 ms and 200–250 ms, the pitch-related early Nd had a more anterior scalp distribution than the location-related early Nd (Fig. 6b; 150–200 ms and 200–250 ms: Condition * Frontality interaction: $F(4,56) = 4.46$ and 2.97, respectively, $P < 0.05$, for both time windows). The same scalp distribution difference was observed between the pitch-related and location-related late Nds (Fig. 6b; 400–500 ms, 500–600 ms and 600–700 ms: $F(4,56) = 4.48$, 4.06 and 4.59, respectively, $P < 0.01$, for all time windows).

Furthermore, in the pitch-attention condition, the late Nd at 400–600 ms had a more anterior scalp distribution than the early Nd at 150–250 ms (Response * Frontality * Laterality interaction: $F(16,224) = 2.63$, $P < 0.01$).

ERF source analyses. At the early-Nd and late-Nd latencies (150–250 ms and 400–500 ms, respectively), the Ndm in the pitch-attention and location-attention conditions was associated with activity especially in the auditory cortex (Fig. 7). The lateral auditory cortex showed the most prominent attention effects that differed clearly from those in medial sites (Fig. 8, left). Therefore, only activity in the lateral auditory-cortex regions were further analyzed. The Ndm activity in the auditory cortex had a centro-posterior maximum (early Ndm: main effect of ROI: $F(8,80) = 7.68$, $P < 0.001$; late Ndm: $F(8,80) = 7.33$, $P < 0.001$). There were no significant differences in distribution of the Ndm activity in the auditory cortex between the pitch-attention and location-attention conditions (Fig. 8).

Nor were there significant differences between the early Ndm and late Ndm distributions in the auditory cortex.

ERP source analyses. The lack of significant differences between the pitch-related and location-related attention effects in the auditory cortex prompted for investigating possible differences between the pitch-related and location-related electrical Nds in the temporo-parietal cortex. In the fMRI experiment of Study II, the temporo-parietal cortex showed stronger location-related than pitch-related attention effects (Fig. 4c and Fig. 5). Therefore, in Study III, additional ERP ROI analyses were conducted using an MNE source model based on a standard brain (MNI152; Montreal Neurological Institute), and spherical ROIs with a diameter of 30 mm set in the temporo-parietal cortex in each hemisphere (Fig. 6c, right). These ROI analyses revealed that, at 150–250 ms, the location-related early Nd was significantly stronger than the pitch-related early Nd in the left temporo-parietal cortex (one-tailed $t(14) = 1.84$, $P < 0.05$), and a similar non-significant tendency was observed also in the right hemisphere (Fig. 6c).
Fig. 6. Results of Study III. (a) Grand-average \( (N = 15) \) ERPs at a selected electrode (Cz of the 10/20 system) with an attention-related negative difference (Nd) response elicited during attention to pitch and attention to location of sounds. A schematic illustration of the electrode layout (130 electrodes) is shown on the right. The white and gray circles depict the electrode matrix used for statistical analysis of the Nd scalp distributions. The Cz electrode is represented by the gray colored circle. A = anterior, P = posterior, L = left, R = right. (b) Scalp distributions of the grand-average early Nds shown for mean amplitudes over the time windows of 150–200 ms and 200–250 ms, and the late Nd distributions for the mean amplitude over 400–500 ms. Amplitude values are scaled (to 0–1 range) so that maximum negativity is represented with darkest gray. The black dots represent the electrode locations. The pitch-related Nd had a more anterior scalp distribution than the location-related Nd over the latencies shown. (c) Mean (±SEM; \( N = 15 \)) early Nd activity in the left and right temporo-parietal cortex during attention to pitch and attention to location of sounds. Squared sum of minimum-norm estimate amplitudes within spherical regions of interest were calculated for the early Nds over the time window of 150–250 ms. The right side of the figure shows the approximate position of the regions of interest projected onto a standard brain (MNI152; Montreal Neurological Institute). The location-related early Nd was significantly (*) stronger than the pitch-related early Nd in the left temporo-parietal cortex (cf. Fig. 4c and Fig. 5). Pitch = pitch-related early Nd, Loc = location-related early Nd. Modified from: Degerman, A., Rinne, T., Särkkä, A-K., Salmi, J. & Alho, K. (2008). Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields. European Journal of Neuroscience, 27, 3329–3341, Federation of European Neuroscience Societies, Blackwell Publishing.
Fig. 7. Results of Study III. (a) Two participants’ (P1 and P2) ERFs at selected channels with the magnetic negative difference (Ndm; cf. Fig. 6a) response during attention to pitch and attention to location of sounds. (b) Minimum-norm estimates (MNEs) of the early Ndm and late Ndm responses calculated over the time window 205–235 ms and 430–460 ms for P1, and 205–235 ms and 465–495 ms for P2. Ndm activity was observed especially in the auditory cortex. Modified from: Degerman, A., Rinne, T., Särkkä, A-K., Salmi, J. & Alho, K. (2008). Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields. European Journal of Neuroscience, 27, 3329–3341, Federation of European Neuroscience Societies, Blackwell Publishing.
3.9 Study IV. Human brain activity associated with audiovisual perception and attention

3.9.1 Specific experimental setting and data analyses

The experiment consisted of 9 different conditions during which audiovisual stimuli were presented, and one condition with no experimental stimulation. The audiovisual stimuli were combinations of the same frequent and infrequent sounds (presented binaurally) and colored circles used in Studies II and III. All combinations of synchronous sounds and circles (blue-high, blue-low, red-high and red-low) were presented with an offset-to-onset interval randomly varying from 300 to 600 ms, in 50-ms steps (Fig. 9).
In two auditory-attention conditions, the participants attended selectively to high or low sounds (cf. pitch conditions of Studies II and III), and in two visual-attention conditions, they attended to the blue or red circles. In four audiovisual-attention conditions, the participants attended to designated audiovisual feature combinations. In addition, there were two conditions in which participants counted mentally backwards from 100 and pressed a response button whenever they reached 90, 80, 70, etc.

In the fMRI analyses, attention-related modulations were revealed by comparing the auditory-, visual- or audiovisual-attention conditions with the baseline (mental counting condition with the unattended audiovisual stimuli). Differences between conditions were studied with direct comparisons. In addition, 8-mm spherical ROIs were used to further investigate attention-related activity in the auditory and visual cortices, and in the frontal cortex during different attention conditions. The auditory-cortex ROIs covered auditory attention-related activity maxima found in the superior temporal gyrus, and the visual-cortex ROIs, visual attention-related activity maxima found in the left middle occipital gyrus and the right middle occipito-temporal cortex (Fig. 10a, bottom). The frontal-cortex ROIs were set to attention-related activity maxima in the left precentral gyrus and right middle frontal gyrus revealed by the comparison of audiovisual-attention conditions vs. auditory-attention conditions (Fig. 11a).

**Fig. 9.** Schematic illustration of the audiovisual stimuli in Study IV. Participants were presented with audiovisual stimuli at a fast rate during all but one task. During the auditory, visual and audiovisual tasks, the participants focused their attention on a designated auditory or visual feature, or audiovisual feature combination in order to detect infrequent shorter-duration targets among the attended stimuli. During two additional tasks, the participants performed a mental backward counting task during which the audiovisual stimuli were ignored, if presented (see text for more details). From: Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I.P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *NeuroImage*, 34, 1683–1691, with permission of Elsevier, Inc.
3.9.2 Results

Performance. There were no significant differences in the accuracy of target detection between the auditory-, visual- and audiovisual-attention conditions as measured with false-alarm rates (overall 26 ± 2%). However, hit rates were lower in the audiovisual-attention conditions (63 ± 6%) than in the auditory-attention (82 ± 3 %) and visual-attention conditions (75 ± 6%; \(F(2,22) = 7.53, P < 0.01\); Newman-Keuls tests: \(P < 0.05\) in both cases). The speed of target detection was similar in the auditory-attention (reaction time: 764 ± 10 ms), visual-attention (761 ± 17 ms) and audiovisual-attention (763 ± 16 ms) conditions.

fMRI results. Attention to auditory, visual and audiovisual stimuli all produced widespread activity in largely overlapping areas, including the so called sensory-specific auditory and visual cortices (Fig 10a). There were no significant differences in auditory cortex activity between the auditory-attention and visual-attention conditions (Fig. 10b). However, activity in the auditory cortex during audiovisual attention exceeded that during either one of the unimodal attention conditions (left and right hemisphere: \(F(3,33) = 20.20\) and 24.93, respectively, \(P < 0.001\) in both hemispheres for differences between the auditory-attention, visual-attention and audiovisual-attention conditions and the mental counting condition with audiovisual stimulation; Newman-Keuls tests: \(P < 0.05\) for comparisons of audiovisual-attention conditions vs. the auditory- or visual-attention conditions). In addition, activity in the right visual cortex and right frontal cortex (Fig. 11) was significantly stronger during both audiovisual and visual attention than during auditory attention (\(F(3,33) = 17.28–30.08, P < 0.01\) for the right visual cortex ROI and the two right frontal cortex ROIs; Newman-Keuls tests: \(P < 0.05\) for the ROIs in the comparisons of audiovisual- or visual-attention conditions vs. the auditory-attention conditions). These brain regions showed no significant differences in activity between the audiovisual-attention and visual-attention conditions.
Fig. 10. Results of Study IV. (a) Areas of significant ($N = 12$; threshold: $Z > 3.5$, corrected cluster threshold $P < 0.05$) activation projected (depth 0–2.0 cm) for illustrative purposes onto an average brain of ten of the participants. Approximate location of regions of interest (ROIs 1–4) used for analyzing mean percent signal changes in (b) are shown (bottom) projected to the brain surface. L = left, R = right. (b) Mean percent signal changes ($\pm$SEM; $N = 12$) in the superior temporal and middle occipito-temporal cortices of each hemisphere. The solid lines depict signal changes during the counting condition with audiovisual stimuli (C), auditory-attention conditions (A), visual-attention conditions (V), and audiovisual-attention conditions (AV). Activity during C was weaker than during A, V and AV in all areas. In the auditory cortex (ROIs 1 and 2), the magnitude of the attention effects in A and V did not differ. However, the magnitude of the auditory cortex attention effect during AV was larger than that during A or V. Both, AV and V produced stronger attention-related modulations than A in the right visual cortex (ROI 4). The dashed lines represent mean percent signal changes ($N = 10$) in the auditory- and visual-cortex ROIs of this Study (IV) applied on data of Study II that had similar A (pitch conditions) and V conditions, but monaural auditory stimuli presented asynchronously with visual stimuli. In Study II, signal changes differed between A and V in all other areas, except the right visual cortex. Modified from: Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I.P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *NeuroImage*, 34, 1683–1691, with permission of Elsevier, Inc.
Fig. 11. Results of Study IV. (a) Significant ($N = 12$; threshold: $Z > 3.5$, corrected cluster threshold $P < 0.05$) attention-related modulations revealed by the comparison of audiovisual-attention conditions with auditory-attention conditions. The comparison of audiovisual attention vs. visual attention showed no significant activity. Approximate locations of regions of interest (ROIs 1-3) used in the analyses described in (b) are illustrated projected to the brain surface. L = left, R = right. (b) Mean percent signal changes ($\pm$SEM) in the frontal-cortex ROIs. The solid lines represent signal changes in ROIs in this study (IV). Activity during audiovisual-attention conditions (AV) and visual-attention conditions (V) was stronger than during the auditory-attention conditions (A) or the counting condition with audiovisual stimuli (C) in all areas except the left frontal cortex, which showed no significant activity differences between AV and C, V and C, or V and A. In addition, activity during A was stronger than during C in the right frontal cortex (ROI 3). The dashed lines represent signal changes in the frontal ROIs applied to data in Study II ($N = 10$), which had similar A (pitch conditions) and V conditions as Study IV (for other details, see Fig. 10b). The signal changes in the frontal areas during the pitch-attention and visual-attention conditions of Study II did not markedly differ from each other. Modified from: Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I.P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *NeuroImage*, 34, 1683–1691, with permission of Elsevier, Inc.
4 DISCUSSION

In the present studies, fMRI, ERP and ERF measures were used to investigate effects of selective auditory attention and audiovisual attention on human brain activity. The main focus was on attention effects in the auditory cortex. Study I showed that the amplitude of attention-related modulations in the auditory cortex increases with the presentation rate of attended sounds as measured with fMRI. In addition, Studies II and III using fMRI, ERP and ERF measures suggested that attention to pitch and attention to location activate overlapping regions in the auditory cortex. Furthermore, the fMRI results of Study IV showed stronger audiovisual than auditory or visual attention-related modulations in the auditory cortex. Below, these findings are discussed in light of other research concerning the brain mechanisms of selective attention and Näätänen’s (1982, 1990, 1992) attentional-trace theory.

4.1 Modulation of auditory cortex attention effects with increasing sound presentation rate (Study I)

Study I showed that activity in the auditory cortex increased with sound presentation rate from 0.5 to 4 Hz, in line with results of other fMRI studies (Binder et al., 1994; Giraud et al., 2000; Harms and Melcher, 2002; Harms et al., 2005; Tanaka et al., 2000). In addition, auditory cortex activity was strongly modulated by attention to the sounds. The auditory attention-related enhancements were of the same magnitude as the activity observed for the experimental auditory stimulation as such. Importantly, there were also larger attention effects in the auditory cortex with higher stimulation rates. This result together with those of previous ERP studies (Alho et al., 1990; Neelon et al., 2006) give some support for Näätänen’s (1990) proposal that the amplitude of attention effects in the auditory cortex increases with presentation rate of the attended sounds. However, in Study I, the auditory attention-related activity in the auditory cortex appeared to reach a plateau at the highest presentation rate (4 Hz; Fig. 2b). Such non-monotonic rate-dependency of attention effects in the auditory cortex would seem to be at odds with Näätänen’s (1982, 1990, 1992) attentional-trace theory.

According to Näätänen’s theory, the highest presentation rate of the attended sounds in Study I should have produced the strongest attention-related activity in the auditory cortex, as it provided the most frequent sensory reinforcement to the attentional trace. The plateau effect observed at the highest stimulation rate in Study I, therefore, suggests that Näätänen’s proposal concerning the rate-dependency of the auditory cortex attention effects holds only for lower (< 4 Hz) presentation rates. However, the plateau effect might
be explained by refractoriness of the neurons producing the attention effect at the highest stimulation rate (Teder et al., 1993). It might also be that activity averaged over 28-s blocks provides only a partial picture of auditory-cortex rate-dependency. The relatively small number of blocks in Study I for each auditory stimulation rate did not allow reliable examination of BOLD responses at different phases of the blocks (i.e., onset, steady-state response, and offset). Yet, for example, Harms et al. (2002, 2005) demonstrated that these different components of the fMRI signal measured from the auditory cortex during a stimulation block are modulated differently by sound presentation rate. In addition, the characteristically sustained fMRI signal becomes increasingly phasic with higher presentation rates (Harms and Melcher, 2002; Harms et al., 2005). This may correspond to a perceptual change where discrete sounds start to be perceived as a single continuous event (Harms and Melcher, 2002; Harms et al., 2005; Weiss et al., 2008), possibly leading to continuous attentional selection, rather than selection of the individual stimuli (see, Teder et al., 1993).

It should be noted that it is not clear to which extent auditory-cortex attention effects actually reflect amplification of stimulus-dependent activations or engagement of other processes (Petkov et al., 2004). The interaction between attention and presentation rate in Study I might be explained in terms of attention-related enhancements of stimulus-dependent activations in the auditory cortex. In this case, the detection of stronger fMRI signals for higher rates could be associated with accumulation of larger attention-related activity to individual sounds (Hillyard et al., 1973; Näätänen et al., 1978) during the higher stimulation rates than during lower rates. However, it is also possible that the attention effects were caused by activation of additional processes required by the task, such as recognition and memory (Hillyard et al., 1973; Näätänen, 1982; Petkov et al., 2004). Thus, the interaction between attention and presentation rate could also be explained by differences in the dynamics of attention-related and stimulus-related processes in the auditory cortex. At least a partial segregation of attention-related modulations and stimulus-dependent activations has been suggested by other fMRI studies (Petkov et al., 2004; Rinne et al., 2007) showing that auditory-cortex attention effects are distributed more laterally than stimulus-dependent activations. Study I did not reveal differences in the distribution of the attention-related modulations and stimulus-dependent activation, possibly because of limitations in the fMRI design (e.g., voxel size) in revealing activation differences. Yet, Studies II and IV (see, Original publications: II and IV) suggested that auditory-cortex attention effects were more widespread than stimulus-dependent activations. These results support the suggestion of at least a partial segregation of attention-related and stimulus-dependent processing in the auditory cortex.
4.2 Attention-related processing of pitch and location of sounds in the brain (Studies II and III)

Consistent with results of some previous ERP (Woods and Alain, 1993; Woods et al., 1994; Woods and Alain, 2001) and ERF (Ahveninen et al., 2006) studies, the ERP experiment of Study III found a more anterior scalp distribution for the pitch-related early Nd than the location-related early Nd. This ERP result suggested different generators for the two attention effects, in line with Näätänen’s (1990) hypothesis that attentional selection of pitch or location of sounds may involve attentional traces with different loci in the auditory cortex (see also Woods et al., 1994). In addition, Study II using fMRI found stronger activity in the left posterior auditory cortex during attention to location than attention to pitch. Correspondingly, previous fMRI (Alain et al., 2001; Barrett and Hall, 2006; Maeder et al., 2001; Warren and Griffiths, 2003) studies have shown that the attentive processing of sound location produces specific activity in more posterior auditory-cortex regions than attentive processing of other auditory features, in accordance with the auditory “what” and “where” model (e.g., Rauschecker and Tian, 2000).

However, in contrast to Näätänen’s (1990) hypothesis, minimum-norm estimation in the ERF experiment of Study III did not reveal differences in the loci of the pitch-related and location-related attention effects in the auditory cortex. Furthermore, no auditory cortex areas showed stronger pitch-related than location-related attention effects in Studies II and III. Moreover, although fMRI (Study II) revealed stronger location-related than pitch-related attention effects in the left posterior auditory cortex, the location-related effects were not restricted to the posterior auditory-cortex regions and the pitch-related effects were not restricted to the anterior regions. This is consistent with studies suggesting that neurons distributed throughout the auditory cortex process sound location (Malhotra et al., 2008; Middlebrooks, 2002; Stecker and Middlebrooks, 2003; Stecker et al., 2005; Zatorre et al., 1999; Zatorre et al., 2002), and that also the posterior auditory cortex participates in processing sound identity cues (Arnott et al., 2004; Barrett and Hall, 2006; Obleser et al., 2007; Tian and Rauschecker, 2004; Zatorre et al., 1999; Zatorre et al., 2002).

The fact that the present Studies II and III could not reliably separate the pitch-related and location-related attention effects in the auditory cortex could be due to limitations in the experimental methodology (discussed in Section 4.4). However, it could also be that attention to pitch and attention to location modulated activity in the same or overlapping neural populations in the auditory cortex. This raises the possibility that attention-related feature processing is not as regionally segregated in the auditory cortex as it appears to be in the visual cortex: For example, Corbetta et al. (1990, 1991) have shown that attention to color, shape or motion of visual stimuli activates different regions of the visual cortex. Yet, based on results in Studies II and III, the posterior auditory cortex
may be activated more by attention to location than attention to pitch. Such increased activity in the auditory cortex during attention to location could reflect incorporation of spatially sensitive neurons in the attentional trace, or enhanced spatial sensitivity in the posterior auditory cortex, in general (see, Stecker et al., 2005; Tian and Rauschecker, 2004; Zatorre et al., 2002).

In Study II, both attention to pitch and attention to location also activated areas outside the auditory cortex. In line with previous fMRI (Salmi et al., 2007b; Wu et al., 2007) and PET (Alho et al., 1999; Tzourio et al., 1997; Zatorre et al., 1999) studies, auditory attention-related activity was observed, for instance, in the prefrontal and inferior parietal cortices. These prefrontal and inferior parietal activations were possibly associated with the control of the attention-related processing in the auditory cortex, and maintenance of attention on the designated sounds when other stimuli were to be ignored (Alho et al., 1999; Driver and Frackowiak, 2001; Näätänen, 1990; Wu et al., 2007). However, the inferior parietal attention effects in Study II might also be associated with task-related processing of the attended stimulus features, rather than attention-related control functions as such. In effect, it has been suggested that the whole concept of the auditory cortex as a restricted area in the superior temporal cortex subserving higher cortical areas should be revised to include substantial temporal, as well as, parietal and frontal cortical regions (Moore et al., 2007; Poremba and Mishkin, 2007). In accordance with previous proposals (e.g., Fritz et al., 2007), the present results suggest that the extent to which these auditory areas are activated depends at least to some extent on the specific task (e.g., attention to pitch or attention to location) performed.

Stronger activity in the premotor/supplementary motor (Study II) and temporo-parietal cortices (Studies II and III) during attention to location than during attention to pitch is consistent with previous fMRI, PET and EEG results (Alain et al., 2001; Alain et al., 2008; Arnott et al., 2004; Bushara et al., 1999; Clarke and Thiran, 2004; De Santis et al., 2007; Maeder et al., 2001; Weeks et al., 1999) suggesting that these brain regions belong to the auditory “where” stream. It could be argued, however, that the premotor/supplementary motor activity in Study II was not associated with spatial processing, but rather with motor responses. This is unlikely, since approximately equal numbers of manual responses were given in different conditions of Study II and therefore there should be no prominent differences in movement-related brain activity between the location and pitch conditions. The results of differences in location-related and pitch-related attention effects in the frontal (Study II) and temporo-parietal (Studies II and III) cortices suggest that there are some functional differences in the attentional selection of pitch and location. These results further suggest that auditory attention operates, at least to some extent, on individual features, instead of operating at a level where the individual features have
already been integrated (for contradicting behavioral and PET results, see Mondor et al., 1998; Zatorre et al., 1999).

Näätänen’s (1982, 1990, 1992) attentional-trace theory postulates that the auditory early Nd is caused by an early component of the PN response generated in the sensory-specific auditory cortex, while the late Nd is caused by a later frontal PN component. The ERF results of Study II support the proposal that the early Nd is generated in the auditory cortex. However, the ERP source analysis of Study III applying the fMRI results of Study II suggested that stronger activity during attention to location than attention to pitch in the inferior parietal cortex contributed to the scalp distribution differences observed between the pitch-related early Nd and location-related early Nd in ERPs (Study III). Moreover, the ERF source analyses in Study III showed that auditory-cortex activity contributes to the Ndm at both the early Nd and late Nd latencies (cf. Arthur et al., 1991; Hari et al., 1989), although these results do not rule out possible contributions of a frontal source to the electrical late Nd suggested by Giard et al. (1988).

4.3 Attention-related processing of audiovisual information in the brain (Study IV)

As expected, Study IV showed that audiovisual attention modulates activity both in the auditory and visual cortices. In terms of Näätänen’s (1990) theory, this result suggests that audiovisual selective attention engages a two-dimensional attentional trace with its two loci in the auditory and visual cortices for processing the respective sensory-specific information. However, in the postero-lateral auditory cortex, the audiovisual attention effects exceeded those during unimodal auditory or visual attention. This raises the possibility that audiovisual selective attention utilizes a neural representation in the auditory cortex that possibly integrates information from both sensory modalities (Beauchamp et al., 2004a; Beauchamp et al., 2004b). A multisensory representation of auditory and visual stimulus features in the auditory cortex could be accomplished via multimodal neurons, as suggested by animal data (Ghazanfar et al., 2005; Meredith, 2004; Schroeder and Foxe, 2005) and by human ERP (Fort et al., 2002a; Fort et al., 2002b; Giard and Peronnet, 1999; Klucharev et al., 2003), ERF (Möttönen et al., 2004) and fMRI results (Calvert et al., 2000). It should be noted, however, that stronger audiovisual than unimodal attention effects found in the postero-lateral auditory cortex may be partly associated with increased task demands during the audiovisual-attention conditions in relation to the unimodal-attention conditions, as suggested by lower hit rates in the audiovisual-attention conditions.

The results showing similar attention effects in the auditory and visual cortices during the auditory-attention and visual-attention conditions in Study IV are in contrast
to the results obtained in Study II showing stronger attention-related modulations in the auditory cortex during auditory attention than during visual attention, and stronger attention-related activity in the visual cortex during visual-attention than auditory-attention conditions (see also, Alho et al., 1999; Kawashima et al., 1999; Petkov et al., 2004; Woodruff et al., 1996). In Study II, the auditory-attention and visual-attention conditions and stimuli were otherwise similar to those in Study IV, except that Study II used monoaural sounds and asynchronous auditory and visual stimulation, while Study IV used binaural sounds and synchronous auditory and visual stimulation. Therefore, it seems possible that the synchronous presentation of auditory and visual stimuli during unimodal auditory or visual attention in Study IV produced unintentional audiovisual integration and thus attentional processing also in the irrelevant sensory modality. This is consistent with previous ERP and fMRI studies (Busse et al., 2005; Driver and Spence, 1998; Eimer and Schröger, 1998; Hillyard et al., 1984; Molholm et al., 2007; Teder-Sälejärvi et al., 1999) indicating that spatial or temporal congruence of sounds and visual objects during attention to one of the sensory modalities may lead to cross-modal attentional spread producing attention-related activity in both auditory and visual cortices (see also, Martuzzi et al., 2007).

In Study IV, frontal areas showed stronger activity during audiovisual attention than during auditory attention. According to animal data (Fuster et al., 2000) and previous fMRI results (Calvert et al., 2000; Calvert et al., 2001) these frontal areas may participate in audiovisual integration. However, as the frontal cortices showed no differences in activity between the audiovisual and visual conditions, stronger audiovisual than auditory activity in these areas appears to be caused by enhanced attentional processing of the visual part of the bimodal stimuli. This is in contrast to the result of Study II showing no marked differences between visual and auditory attention-related activity in the same frontal areas. A possible reason for stronger frontal activity during visual attention than during auditory attention in Study IV is that the visual-attention condition required more effort. This explanation is not supported by the behavioral results of Study IV, as there were no differences in response accuracy or speed between these unimodal attention conditions. The enhanced activity in the frontal cortices during visual-attention conditions in Study IV might also reflect audiovisual integration caused by the synchronous presentation of the auditory and visual stimuli. However, this would mean that integration-related activity in the frontal cortices would be elicited only when at least the visual part of the audiovisual stimulus is attended. Moreover, if attention to the visual portion of the audiovisual stimulus was enough to produce audiovisual integration, then one would expect this to also produce activations of same magnitude during audiovisual-attention and visual-attention conditions in other areas involved in audiovisual integration, that is, also in the postero-lateral auditory cortex.
where audiovisual attention was associated with higher activity than visual (or auditory) attention to bimodal stimuli.

It might be argued that, in Study IV, stronger audiovisual than auditory attention-related activity in the frontal cortices was caused by attentional shifting during the audiovisual-attention conditions (Shomstein and Yantis, 2004), that is, the participants would have continuously shifted their attention between the two modalities rather than attended to them at the same time. This is unlikely, however, as the audiovisual-attention conditions did not show longer reaction times than the unimodal-attention conditions. Slower performance in the audiovisual-attention conditions would have been expected if auditory and visual information were processed serially rather than in parallel in these conditions. Moreover, shifting-related activity should have been revealed by comparisons of audiovisual-attention conditions vs. auditory- or visual-attention conditions, but only the comparison of audiovisual-attention conditions vs. auditory-attention conditions revealed significant activity in the frontal cortices.

4.4 Methodological considerations

Näätänen’s (1982, 1990, 1992) attentional trace theory is largely based on results obtained with ERPs and ERFs. Both ERPs and ERFs allow for highly accurate detection of the temporal dynamics of activity associated with attended stimuli. A large part of this thesis (Studies I, II and IV), however, focused on brain networks of selective attention as revealed by fMRI. An inherent limitation in the blocked fMRI designs used in Studies I, II and IV was the lack of temporal resolution. The fMRI experiments studied activity averaged across blocks lasting tens of seconds, which is in sharp contrast to the millisecond accuracy provided by ERPs and ERFs in Study III. Therefore it is not clear to what extent the attention-related activations measured with fMRI (Studies I, II and IV) actually reflect attentional selection suggested to underlie the early PN (Näätänen, 1990) that peaks within 200 ms from stimulus onset in ERPs. Nevertheless, Studies II and III demonstrated at least some correspondence between the hemodynamic and electrophysiological responses. For instance, both fMRI (Study II) and ERFs (Study III) showed prominent pitch-related and location-related attention effects with overlapping loci in the auditory cortex. In addition, both fMRI and ERPs (Study III) found stronger activity for attention to location than for attention to pitch in the temporo-parietal cortex. Although the precise way in which neural activity is associated with hemodynamic responses is unknown (e.g., Logothetis, 2007), the present results suggest correlation between electrophysiological and hemodynamic measures.
Another consideration in the fMRI design is that the background scanner noise produced by the continuous imaging sequence in Studies I, II and IV also activates the auditory cortex (Shah et al., 1999). After attenuation by acoustically shielding materials wrapped around the head coil, earphones and earplugs, the effective loudness of the scanner noise at the eardrum was still approximately 70 dB SPL. As the scanner noise may induce a change in auditory cortex activity of approximately a few percent relative to the baseline (Hall et al., 2000), this may have affected detectability of the experimental stimulus-related activity. It should be noted, however, that the auditory stimuli in all present fMRI studies were clearly distinguishable from the scanner noise that remained stable across each experiment. In addition, ERP and fMRI results suggest that although the scanner noise affects auditory stimulus-dependent activations, it does not have a significant effect on activations related to voluntary or involuntary attention (Novitski et al., 2001; Woods et al., 2008). Moreover, the fMRI results of Studies I, II and IV dealt with auditory-cortex activity from which the effects of the scanner noise had been subtracted. Yet, in Study II, the scanner noise may have added multidimensionality to the auditory conditions, as participants had to attend to sounds with a certain pitch and location while ignoring the scanner noise with a different pitch and location. This might have caused some insensitivity in Study II in detecting differences in brain activity between attention to pitch and attention to location.

There might also be some limitations in the ERF measures of Study III worth considering. First, although the result of stronger ERF attention effects in the lateral auditory cortex than in the medial auditory cortex is consistent with results of previous fMRI studies (Petkov et al., 2004; Rinne et al., 2007), this does not necessarily mean that the lateral areas were activated more by auditory selective attention than the medial areas. This is because the lack of strong ERF attention effects in the medial auditory cortex could also be explained by limitations of MEG in detecting deeper sources in the brain (Hämäläinen et al., 1993). Second, it might be argued that the ERF source analyses were insensitive to the differences between the pitch-related and location-related attention effects in the auditory cortex, because the differences observed in the ERP scalp distribution analyses were actually generated in the lateral surface of the auditory cortex by source components oriented radially to the skull. Activity of radial source components could be detected with EEG, but would be difficult to detect with MEG, which is sensitive to activity of tangential components of sources (Hämäläinen et al., 1993). However, radial source components, for instance, in the auditory cortex on the lateral aspect of the superior temporal gyrus could hardly explain ERP scalp-distribution differences along the anterior-posterior dimension. Third, it is possible that the spatial resolution of ERF source analysis with minimum-norm estimates was not high enough for revealing the possibly subtle differences between the pitch-related and location-related attention effects in the auditory
cortex. In future studies, the possibility for detecting differences between pitch-related and location-related attention effects could be increased by using combined EEG/MEG source estimation methods (Molins et al., 2008) and, for instance, by weighting MEG source estimates toward activated areas revealed with fMRI (Dale et al., 2000).

Although ERPs in Study III showed scalp distribution differences between the pitch-related and location-related attention effects, the ERPs provide limited information about the possible generators of these differences in the auditory cortex. Initially, in the data analyses of Study III, minimum-norm estimates were used to model the auditory cortex sources of the early electrical Nd. This, however, did not yield reliable source models for auditory-cortex attention effects. In addition, the Nd sources were studied with equivalent current dipoles. In these models, an early-Nd dipole was fixed in the auditory cortex of each hemisphere to a locus where the N1 response was localized. The dipole models allowed for estimating differences in the auditory cortex sources of the pitch-related and location-related attention effects by looking at the dipole orientations. These analyses revealed no significant differences in the orientation of the dipoles associated with attention to pitch and attention to location. Therefore, as MEG is more sensitive to source-level differences than EEG, the auditory cortex source analyses in Study III focused on the ERF data. Note that for ERFs, dipole modeling similar to that used in the initial ERP analyses showed no differences between the Ndm for attention to pitch and the Ndm for attention to location.

4.5 Conclusions

The main focus of the present studies was on attention-related processing in the auditory cortex. As expected, it was found that activity in the auditory cortex is strongly modulated by selective attention. In addition, the present results suggest that the amplitude of attention-related activity in the auditory cortex increases with presentation rate of attended sounds. These results emphasize the importance of controlling for both attention and stimulation rate in studies of auditory processing. Importantly, the present studies could not reliably separate pitch-related and location-related attention effects in the auditory cortex. This suggests that attention-related processing of sound identity and location may not be segregated in the auditory cortex. However, there appears to be enhanced attentional processing of sound location in the temporo-parietal junction and areas of the frontal cortices. In addition, a study on bimodal attention found that activity in the auditory cortex was more prominent during audiovisual attention than during auditory or visual attention. This result suggests that audiovisual attention utilizes a multisensory representation in the auditory cortex. Taken together, results of the present studies contribute to the knowledge about rate-dependent, feature-specific and integrative attention-related functions in the auditory cortex.
5 REFERENCES


