

FUNCTIONAL ORGANIZATION OF HUMAN AUDITORY CORTEX DURING ACTIVE AUDITORY TASKS

Suvi Häkkinen

Doctoral Programme in Psychology, Learning and Communication
Department of Psychology and Logopedics, Faculty of Medicine
University of Helsinki, Finland

ACADEMIC DISSERTATION

To be publicly discussed
by due permission of the Faculty of Medicine
at the University of Helsinki in Lecture Hall 3, Biomedicum 1,
Haartmaninkatu 8, Helsinki, on the 18th of June, at 12 PM

Supervisors Dr. Teemu Rinne
Turku Brain and Mind Center
Department of Clinical Medicine
University of Turku, Finland

Associate Professor G. Christopher Stecker
Department of Speech & Hearing Sciences
Vanderbilt University, TN, USA

Reviewers Dr. Pia Rämä
Laboratoire Psychologie de la Perception
Université Paris Descartes, France

Associate Professor Iiro P. Jääskeläinen
Department of Neuroscience and Biomedical Engineering
Aalto University, Finland

Opponent Professor Marc Schönwiesner
Department of Biology
University of Leipzig, Germany, and
Department of Psychology
University of Montreal, Canada

ISBN 978-951-51-4249-8 (pbk.)

ISBN 978-951-51-4250-4 (PDF)

Unigrafia

Helsinki 2018

CONTENTS

Abstract.....	5
Tiivistelmä	7
Acknowledgments.....	9
List of original publications.....	11
1 Introduction	13
1.1 Auditory cortex.....	14
1.2 The auditory “what”/”where” model	16
1.3 Attention-related modulations in human auditory cortex	17
1.4 Activation related to selective attention to pitch and location	20
1.5 Activation during discrimination and <i>n</i> -back memory tasks	21
2 Aims of the present thesis.....	25
3 Methods.....	27
3.1 Subjects.....	27
3.2 Stimuli.....	27
3.3 Tasks	28
3.4 Paradigms and procedures.....	30
3.5 Behavioral training and analysis of task performance	31
3.6 Study I. Processing of pitch and location in human auditory cortex during visual and auditory tasks	31
3.6.1 Methods.....	31
3.6.2 Results	33
3.7 Study II. Source analysis of event-related potentials during pitch discrimination and pitch memory tasks	35
3.7.1 Methods.....	35
3.7.2 Results	37

3.8	Study III. Intrinsic, stimulus-driven and task-dependent connectivity in human auditory cortex.	39
3.8.1	Methods.....	39
3.8.2	Results.....	41
4	Discussion	45
4.1	Imaging the modular organization of human AC	45
4.2	The effect of active listening on stimulus-level activation.....	47
4.3	Temporal dynamics of task-dependent activation	49
4.4	Hierarchical organization of human AC	50
4.5	Functional significance of task-dependent activation in STG and IPL.....	51
4.6	A tentative model of task-dependent activation in human auditory cortex	54
5	Conclusions	57
	References.....	58
	Original publications.....	76

ABSTRACT

Previous imaging studies have shown that activation in human auditory cortex (AC) is strongly modulated during active listening tasks. However, the prevalent models of AC mainly focus on the processing of stimulus-specific information and speech and do not predict such task-dependent modulation. In the present thesis, functional magnetic resonance imaging was used to measure regional activation in AC during discrimination and *n*-back memory tasks in order to investigate the relationship between stimulus-specific and task-dependent processing (**Study I**) and inter-regional connectivity during rest and active tasks (**Study III**). In addition, source analysis of scalp-recorded event-related potentials was carried out to study the temporal dynamics of task-dependent activation in AC (**Study II**).

In **Study I**, distinct stimulus-specific activation patterns to pitch-varying and location-varying sounds were similarly observed during visual (no directed auditory attention) and auditory tasks. This is consistent with the prevalent models which presume parallel and independent “what” (e.g. pitch) and “where” processing streams. As expected, discrimination and *n*-back memory tasks were associated with distinct task-dependent activation patterns. These activation patterns were independent of whether subjects performed pitch or location versions of these tasks. Thus, AC activation during discrimination and *n*-back memory tasks cannot be explained by enhanced stimulus-specific processing (of pitch and location). Consistently, **Study II** showed that the task-dependent effects in AC occur relatively late (200–700 ms from stimulus onset) compared to the latency of stimulus-specific pitch processing (0–200 ms). In **Study III**, the organization of human AC was investigated based on functional connectivity. Connectivity-based parcellation revealed a network structure that consisted of six modules in supratemporal plane, temporal lobe, and inferior parietal lobule in both hemispheres. Multivariate pattern analysis showed that connectivity within this network structure was significantly modulated during the presentation of sounds (visual task) and auditory task performance. Together the results

of this thesis show that (1) activation in human AC strongly depends on the requirements of the listening task and that task-dependent modulation is not due to enhanced stimulus-specific processing, (2) regions in inferior parietal lobule play an important role in the processing of both task-irrelevant and task-relevant auditory information in human AC, and (3) the activation patterns in human AC during the presentation of task-irrelevant and task-relevant sounds cannot be fully explained by a hierarchical model in which information is processed in two parallel processing streams.

TIIVISTELMÄ

Aiemmat kuvantamistutkimukset ovat osoittaneet, että aktiiviset kuuntelutehtävät vaikuttavat voimakkaasti ihmisen kuuloaivokuoren aktivaatioon. Kuuloaivokuoren toiminnalliset mallit kuitenkin keskittyvät äänten akustisten piirteiden ja puheen käsittelyyn, eivätkä ne siten ennusta tehtävisidonnaisia vaikutuksia. Tässä väitöskirjassa tutkittiin kuuloaivokuoren toimintaa äänten erottelu- ja *n*-back-muistitehtävien aikana toiminnallisella magneettikuvauksella ja herätevasterekisteröinnillä. Tutkimusten tavoitteena oli selvittää riippuvatko ärsyke- ja tehtävisidonnaiset aktivaatiot toisistaan (**Tutkimus I**) sekä tutkia kuuloaivokuoren eri alueiden välistä toiminnallista konnektiivisuutta lepo- ja tehtävätilanteiden aikana (**Tutkimus III**). Pään pinnalta mitattujen herätevasteiden lähdemallinnuksen avulla tutkittiin kuuloaivokuoren tehtävisidonnaisen aktivaation ajallista dynamiikkaa (**Tutkimus II**).

Tutkimuksessa I äänen korkeuden ja tulosuunnan vaihtelu aktivoivat erillisiä kuuloaivokuoren alueita sekä näkötehtävän (ei suunnattua kuulotarkkaavaisuutta) että kuuntelutehtävien aikana. Tämä tulos on yhtenevä vallitsevien kuuloaivokuoren mallien kanssa, joissa oletetaan, että äänen korkeus ja tulosuunta käsitellään rinnakkaisissa ja toisistaan riippumattomissa mitä- (esim. äänen korkeus) ja missä-järjestelmissä. Aktiivisten kuuntelutehtävien aikana kuuloaivokuoren aktivaatiojakauma riippui odotetusti siitä, tekivätkö koehenkilöt äänten erottelu vai *n*-back-muistitehtävää. Tehtävisidonnaiset aktivaatiojakaumat (erottelu- ja muistitehtävän erot) olivat kuitenkin hyvin samankaltaisia äänen korkeus- ja tulosuuntatehtävien aikana. Kuuloaivokuoren tehtävisidonnaisia aktivaatioita äänten erottelu- ja *n*-back-muistitehtävien aikana ei siten voida selittää ääni-informaation käsittelyyn liittyvien aktivaatioiden voimistumisella. Tätä johtopäätöstä tukevat myös **Tutkimuksen II** tulokset, joiden mukaan kuuloaivokuoren tehtävisidonnaiset aktivaatiot (n. 200–700 ms äänen alusta) havaitaan pääosin äänenkorkeustiedon käsittelyyn liittyvän aktivaation (0–200 ms) jälkeen. **Tutkimuksessa III**

selvitettiin kuuloaivokuoren toiminnallista organisaatiota ja sen eri aluiden muodostamia verkostoja konnektiivisuusanalyysien avulla. Näissä analyyseissä havaittiin modulaarinen rakenne, jossa kuuloaivokuori ja sen lähialueiden muodostama verkosto jakaantuu kuuteen osaan (moduuliin). Toiminnallisen konnektiivisuuden muutoksia eri koetilanteissa tarkasteltiin monimuuttujakuvioanalyysillä. Tulokset osoittivat, että konnektiivisuus kuuloaivokuoren ja sen lähialueiden muodostamassa verkostossa muuttui merkittävästi verrattuna lepo-tilanteeseen, kun koehenkilöille esitettiin ääniä (näkötehtävän aikana) ja kun he tekivät kuuntelutehtäviä.

Väitöskirjan tutkimusten tulokset osoittavat, että (1) ihmisen kuuloaivokuoren aktivaatio riippuu voimakkaasti kuuntelutehtävän vaatimuksista, (2) päälaenlohkon alaosat osallistuvat merkittäväällä tavalla kuuloaivokuoren toimintaan ääni-informaation käsittelyn aikana riippumatta siitä, ovatko äänet olennaisia vai epäolennaisia kulloisenkin tehtävän kannalta, ja (3) ihmisen kuuloaivokuoren aktivaatiota ei voida täysin selittää hierarkisella mallilla, jossa ääni-informaatiota käsitellään kahdella rinnakkaisella tiedonkäsittelyradalla.

ACKNOWLEDGMENTS

This work was carried out in the @brain research group, Department of Psychology and Logopedics (former Institute of Behavioral Sciences), University of Helsinki. The fMRI measurements were conducted at the Advanced Magnetic Imaging Centre (AMI), Aalto University School of Science, and EEG data was collected at the Cognitive Brain Research Unit (CBRU), University of Helsinki. The thesis was funded by the Academy of Finland, the Doctoral Programme in Psychology, Learning and Communication, and the Instrumentarium Science Foundation.

First and foremost I want to thank my supervisor Dr. Teemu Rinne for the opportunity to work in his project, encouragement throughout the thesis work, and hard work in keeping my terminology straight. I may now understand attention. I give my warmest thanks also to my second supervisor Associate Professor Chris Stecker for his insightful comments over the years.

I thank the official reviewers Dr. Pia Rämä and Associate Professor Iiro Jääskeläinen for valuable comments on this thesis, and Professor Marc Schönwiesner for kindly agreeing to serve as my opponent. I also want to thank Professor Kimmo Alho for collaboration on the first study and for agreeing to be my custos.

I highly appreciate the excellent research facilities at both AMI and CBRU. Marita Kattelus was great help in getting started with MRI measurements and research assistants Noora Ovaska and Leena Tirkkonen in subject recruitment and training. I also wish to thank the subjects one more time for taking all the tasks including the direction 3-back memory task in stride.

I was fortunate to share the office with Drs. Emma Salo, Kirsi Harinen and Annika Svedholm-Häkkinen and Mr. Patrik Wikman. Thank you for the discussions on all things important, the absurd amount of faith you had in me, and the candy. Sincere thanks also to other dear friends inside and outside the academic bubble for support and encouragement over the years.

I extend my thanks to everyone at the FinnBrain project, University of Turku, for the friendliness and understanding during the last few busy months.

Finally, an enormous thank you to my parents Eila and Aulis Talja, for the love and support, and to my sister Niina Keränen, for letting me defend first. My deepest gratitude goes to my husband Teemu Häkkinen and our son Mikael for the patience, joy, and curiosity in life.

Helsinki, May 2018

Suvi Häkkinen

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

- Study I** Häkkinen (née Talja), S., Ovaska, N., and Rinne, T. (2015) Processing of pitch and location in human auditory cortex during visual and auditory tasks. *Frontiers in Psychology*, 6.
- Study II** Talja, S., Alho, K., and Rinne, T. (2015) Source analysis of event-related potentials during pitch discrimination and pitch memory tasks. *Brain Topography*, 28, 445–458.
- Study III** Häkkinen, S., and Rinne, T. (2018) Intrinsic, stimulus-driven and task dependent connectivity in human auditory cortex. *Brain Structure & Function*, in press.

1 INTRODUCTION

Human auditory cortex (AC) in the superior temporal lobe receives its acoustic input from the subcortical ascending pathway via the inferior colliculus (IC) and the medial geniculate body (MGB) of the thalamus. AC is often seen as a low-level sound analyzer, while more cognitive functions occur in cortical areas outside AC. However, this view is not fully consistent with literature. For example, most of the input to primary AC (area A1) originates from other cortical sources and not from the ascending pathway (Budinger and Scheich, 2009). In addition to the ascending pathway, there is a massive descending pathway from AC to MGB and IC (Winer, 2005), which allows AC to modulate subcortical auditory processing (Bajo et al., 2010; Rinne et al., 2008; Slee and David, 2015; Suga and Ma, 2003; Winer and Lee, 2007). Moreover, human and animal studies have demonstrated that operations in AC are strongly modulated by active listening tasks (Fritz et al., 2003; Mesgarani and Chang, 2012; Petkov et al., 2004; Rinne et al., 2009). Taken together, these anatomical and functional findings suggest that AC does not only process acoustical information from the ascending pathway but is also involved in higher-order cognitive processes and the control of auditory subcortical processing. In fact, it has been argued that task-specificity is a key feature of auditory representations in AC (Scheich et al., 2007) and that, rather than being a low-level sound analyzer, the overall function of AC is to solve higher auditory problems and challenges that face an organism (Weinberger, 2011). The idea that responses in AC are not a fixed function of the acoustical stimulus properties and that operations in AC should be investigated in a cognitive context is becoming increasingly accepted in auditory neuroscience (David, 2018; Higgins et al., 2017; Irvine, 2017; Walker and King, 2011). Yet, current models of AC operations focus on feature-specific analysis of auditory information. The present thesis used fMRI activation and connectivity analysis as well as source-analysis of scalp-recorded EEG to examine the functional dynamics and organization of human AC during active listening tasks.

1.1 AUDITORY CORTEX

Based on cytoarchitecture and anatomical connections in nonhuman primates, it is hypothesized that AC consists of anatomically and functionally distinct core, belt and parabelt regions that can be further divided into several areas (Hackett et al., 2001; Hackett et al., 1998; Kaas and Hackett, 2000). The core is strongly connected with the surrounding belt and parabelt regions in a hierarchical manner (core -> belt -> parabelt; **Figure 1 a**). Consistent with a hierarchical organization, studies in primates have shown that core neurons also show shorter response latencies, narrower tuning and greater temporal precision than lateral belt neurons (Kosaki et al., 1997; Lakatos et al., 2005; Rauschecker et al., 1995; Recanzone et al., 2000). Belt and parabelt areas are connected with other parts of the cortex, and these connections are organized into at least two parallel pathways (**b**; Recanzone and Cohen, 2010; Romanski et al., 1999; see also Section 1.2). Besides the feedforward hierarchy, there are feedback connections, for instance, from parabelt to core and from other cortical regions to parabelt and belt (Hackett, 2015; Hackett et al., 2014).

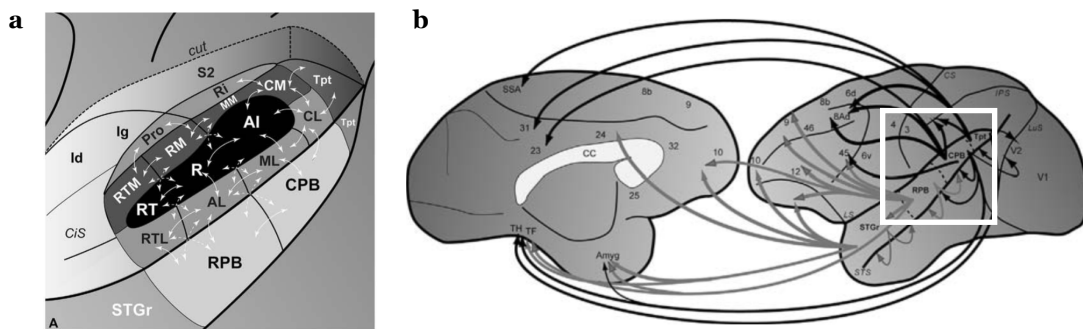


Figure 1 Topographical organization of anatomical connections in monkey superior temporal cortex. **(a)** Connections within AC are organized hierarchically between core (black; areas AI, R, RT), belt (grey; CL, ML, AL, RTL, RTM, RM, MM, CM) and parabelt (light grey; CPB, RPB) regions so that adjacent areas are reciprocally connected. **(b)** Connections from AC to frontal and parietal cortex. Connections are shown on medial (left) and lateral (right) surfaces of the left hemisphere. The connections are organized into at least two parallel pathways and are mostly reciprocal. The white rectangle in **(b)** indicates the location of the cortical region in **(a)**. Adapted from Hackett (2015). Reproduced with permission from Elsevier.

Post-mortem studies have revealed a similar core–belt–parabelt organization also in human AC (Fullerton and Pandya, 2007; Rivier and Clarke, 1997; Sweet et al., 2005). Human primary auditory cortex (A1 area of the core) is located in the vicinity of the medial Heschl’s gyrus on the superior temporal plane. However, based on macroanatomical markers, the localization of A1 is inaccurate. A1 can be noninvasively defined by mapping myelin content and tonotopy using magnetic resonance imaging (MRI; Da Costa et al., 2011; Rademacher et al., 2001). MRI of myelin content suggests that A1 (high myelin content) is located in medial HG (Dick et al., 2012; Glasser and Van Essen, 2011; Wallace et al., 2016). Consistently, also tonotopic mapping studies suggest that core areas A1 and R are located near medial HG although their exact location and tonotopic patterns in the surrounding areas are still debated (Baumann et al., 2013; Moerel et al., 2014; Saenz and Langers, 2014). Currently, there are no generally accepted non-invasive approaches to define the regions in human AC outside the core.

The similarities and differences between human and primate AC are not well understood. This complicates the use of the current primate models in research on the neural basis of human auditory cortical functions. Previous studies have identified key differences between human and monkey AC in terms of morphology (e.g. non-human primates lack a proper Heschl's gyrus; Baumann et al., 2013), cytoarchitecture (Fullerton and Pandya, 2007; Hackett, 2015; Rivier and Clarke, 1997), and connectivity (Frey et al., 2008; Neubert et al., 2014; Ruschel et al., 2014). For example, in humans, but not in monkeys, regions in inferior parietal lobule (IPL) are strongly connected with AC regions, which suggests that IPL plays an important role in the operations of human AC. Most studies on monkeys have focused on stimulus-driven effects during passive conditions, partially because of the difficulties in training nonhuman primates on auditory tasks (Rinne et al., 2017). By contrast, human studies have shown that operations in human AC are strongly modulated during active listening tasks (Mesgarani and Chang, 2012; Petkov et al., 2004; Rinne et al., 2009). These modulations cannot be predicted based on the current models of AC based on studies in nonhuman primates. Taken together, it is clear that more systematic work is needed to

better understand the functional organization of human AC and how it is affected by listening tasks.

1.2 THE AUDITORY “WHAT”/“WHERE” MODEL

In an influential study, Tian et al. (2001) investigated single-neuron responses in lateral belt areas of monkey AC to species-specific communication calls occurring at different spatial locations. They found that neurons in anterolateral belt (AL in **Figure 1**) were selective to the type of monkey call irrespective of the spatial location of the stimulus, whereas neurons in the caudolateral belt (CL) were selective to location independent of the vocalization type. Based on these results, the authors suggested that auditory processing is organized in specialized “what” (identity of the source) and “where” (location of the source) streams analogous to the visual system (Goodale and Milner, 1992; Mishkin and Ungerleider, 1982).

Numerous human imaging studies have tested the predictions of this model (for recent reviews see Ahveninen et al., 2014; Alho et al., 2014). Ahveninen et al. (2006) compared sensitivity in AC to phonetic content and location of sounds using magnetoencephalography (MEG). They used the so-called stimulus adaptation paradigm, which is based on the finding that the neural response to a stimulus is reduced when it is preceded by a physically similar stimulus. The authors hypothesized that if phonetic content and location are processed in separate regions of anterior and posterior AC, respectively, then these regions would show different adaptation effects to phonetic content and location. Their subjects were presented with pairs of vowels from two locations during passive listening or during selective attention to phonetic content or location. As hypothesized, the source estimates of the N1m response (100 ms from stimulus onset) showed adaptation to phonetic content in anterolateral HG, anterior STG and posterior planum polare, whereas adaptation to location was observed in planum temporale (PT) and posterior superior temporal gyrus (STG). The results of this study, and those of several other studies using the adaptation paradigm (e.g. Altmann et al., 2007; Barrett and Hall, 2006; De Santis et al.,

2007; van der Zwaag et al., 2011; Warren and Griffiths, 2003), are consistent with the view that that stimulus-specific processing of nonspatial (“what”) and spatial (“where”) features involves at least partially distinct regions in AC.

However, not all results of previous human imaging studies appear compatible with the “what”/“where” model. For example, a number of previous studies suggest that PT, within the putative “where” pathway, is involved in the analysis of pitch sequences (Barker et al., 2013; Barrett and Hall, 2006; Griffiths et al., 1998; Overath et al., 2007; Patterson et al., 2002; Puschmann et al., 2010; Stewart et al., 2008; Warren and Griffiths, 2003). It has been suggested that PT may contain functional subdivisions so that sensitivity to location is observed in posterior PT and sensitivity to pitch in anterior PT (Ahveninen et al., 2014; Hickok and Saberi, 2012). That is, according to this idea anterior and posterior part of PT would belong to the “what” and “where” stream, respectively. Other studies have reported that PT activation does not increase with the amount of spatial variation (Smith et al., 2010; Zatorre et al., 2002). These and several other studies have suggested that PT activation is related to a more general auditory function such as source segregation, which often is based on spatial cues, rather than processing location as such. Griffiths and Warren (2002), for example, suggested that PT is a processing hub that performs source segregation and analyzes the nature of the signal based on previously learned patterns. In addition, the STG areas associated with the “what” and “where” pathways are strongly modulated by attention and active listening tasks (Grady et al., 1997; Hall et al., 2000; Petkov et al., 2004; Rinne et al., 2009; Woods et al., 2009). However, previous studies have not systematically investigated the relationship between the stimulus-specific processing of “what” and “where” information and effects associated with active listening tasks.

1.3 ATTENTION-RELATED MODULATIONS IN HUMAN AUDITORY CORTEX

In studies using noninvasive brain imaging, the effect of attention on auditory processing is typically measured as the difference between

responses to the same sounds presented during attended and ignored conditions. In a pioneering event-related potential (ERP) study by Hillyard et al. (1973), subjects were presented with asynchronous streams of short tones to both ears in which 10% of the sounds had a slightly higher pitch. Their subjects were instructed to attend to the sounds in one ear, to ignore similar sounds in the other ear, and to report the higher-pitch targets among the attended stream. ERPs were stronger to the attended than to ignored tones at 80–110 ms from sound onset. According to the authors, this was because the N1 wave of ERP was enhanced during attended conditions. It is generally accepted that the N1 wave, peaking circa 100 ms from sound onset, reflects onset of sound energy and the stimulus-driven processing of the physical aspects of stimuli and that its main sources are in AC (Arthur et al., 1991; Godey et al., 2001; Liégeois-Chauvel et al., 1994; Näätänen and Winkler, 1999). The difference between the attended and ignored ERPs was later termed as “the negative difference” (Nd; Hansen and Hillyard, 1980).

Since this initial study, it has been debated to what extent the stimulus-specific (N1) and attention-dependent (Nd) activity are related. Similar to N1, the main sources of Nd are estimated to be in bilateral superior temporal cortices (Arthur et al., 1991; Degerman et al., 2008; Giard et al., 1988; Hari et al., 1989; Jemel et al., 2003; Kauramäki et al., 2012; Rif et al., 1991; Ross et al., 2010; Tiitinen et al., 2006; Woldorff et al., 1993). However, some studies have reported that N1 and Nd do not always overlap in time and show different scalp topographies (reviewed in Näätänen, 1990). Such results suggest that N1 and Nd are generated by distinct brain processes and that attention effects cannot be explained by enhanced processing of stimulus-level information. However, the results of other studies suggest that attentional modulation could be due to enhanced stimulus-specific processing. For example, monaural stimulation elicits stronger responses in the contralateral than ipsilateral hemisphere (Loveless et al., 1994; Suzuki et al., 2002). Thus, if attentional modulation were due to enhanced stimulus-specific processing, then attention effects should also be stronger in the contralateral hemisphere. Consistently, contralateral attention effects have been reported in several previous studies (Neelon et al., 2011; Woldorff and

Hillyard, 1991; Woods et al., 1992). Further, the aforementioned study by Ahveninen et al. (2006) showed that selective attention enhances adaptation effects to phonetic content and location. Together these results suggest that auditory attention related effects in ERPs are due to both modulations in stimulus-level processing and effects that are more directly associated with attention or the attention-engaging task.

In addition to Nd, previous ERP studies have reported various effects at later latencies (300–900 ms from sound onset) associated with active listening task, task-relevance, response selection, and memory (Donchin and Coles, 1988; Herrmann and Knight, 2001; Picton, 1992; Polich, 2007). However, it is not known whether sources in AC contribute to these effects (Linden, 2005; Nieuwenhuis et al., 2005).

It has also been long debated whether attention effects arise due to “early” or “late” selection (Broadbent, 1958; Deutsch and Deutsch, 1963; Murphy et al., 2017). According to so-called early selection theories, attention enhances sensory analysis of task-relevant features (e.g. pitch or location of sounds that vary in both features) and dampens the processing of task-irrelevant features. Late selection theories, in contrast, suggest that acoustical analysis takes place irrespective of attention and that selection is based on higher-level “objects” attributed with meaning (e.g. a bird singing, a car passing, a particular voice; Bregman, 1990; Shinn-Cunningham, 2008). Thus, late selection theories typically predict that attentional and task processing are distinct from acoustical processing.

The so-called reverse hierarchy theory (Hochstein and Ahissar, 2002; Nelken and Ahissar, 2006) combines aspects of early and late selection theories. The theory posits that immediate auditory analysis is crude, categorical (e.g. identify a spoken word and its speaker), and rapid (within 100 ms). More detailed analysis occurs later if it is required by the current task. For example, categorical same–different tasks do not require detailed acoustical analysis, whereas detection of small pitch changes does. Unlike the early and late selection theories, the reverse-hierarchy theory does not assume auditory processing to be fully hierarchical. Instead, it posits that

attention-dependent modulation of processing in AC may occur early or late depending on the specific requirements of the listening task.

Only a few previous EEG/MEG studies have used source analysis methods to investigate attention-related modulations in AC. However, several fMRI studies have shown that activation, particularly in posterolateral STG, is strongly modulated during attention-engaging auditory tasks (Alho et al., 2014; Hall et al., 2000; Petkov et al., 2004; Woods et al., 2009; see also Section 1.5). As the temporal resolution of fMRI is relatively poor, the time course of these modulations is not well understood.

1.4 ACTIVATION RELATED TO SELECTIVE ATTENTION TO PITCH AND LOCATION

Within the auditory “what”/“where” model, it is often assumed that selective attention to “what” and “where” features enhances activation in anterior and posterior AC, respectively. As pitch is an important auditory identity cue, several studies have compared selective attention to pitch and location. Degerman et al. (2008) used EEG and MEG to compare Nd/Ndm associated with selective attention to pitch or location. During attention to pitch, high or low sounds were randomly delivered to constant location (left or right ear) and subjects were required to selectively attend either the high or low sounds. During attention to location, sounds had a fixed pitch, their location varied randomly (left or right ear), and subjects selectively attended sounds in one ear. During all auditory attention conditions, subjects detected occasional shorter-duration targets among the attended sounds. The EEG results showed that scalp distribution of Nd (vs. visual task) was more anterior during attention to pitch than during attention to location. However, in MEG, the authors found no differences between Ndm responses and source estimation did not find significant differences in the superior temporal cortex. They concluded that, as MEG should be more sensitive than EEG to sources in anterior and posterior STG, selective attention to pitch and location is not associated with an activation difference in superior temporal cortex. However, several other EEG studies have reported differences in AC activity during selective attention to pitch or location at several latencies

(ca. 100–600 ms; Alain et al., 2001; 2009; Anourova et al., 2001; Bidet-Caulet et al., 2007; Woods and Alain, 2001; Woods et al., 1994).

A number of previous fMRI studies have shown attention to pitch is associated with more anterior activation than attention to location within superior temporal cortex (reviewed in Alho et al., 2014). However, some studies also report enhanced activation in anterior STG during location tasks and enhanced activation in posterior STG/IPL during pitch tasks (Griffiths and Warren, 2002; Hill and Miller, 2010; Rinne et al., 2009; 2012; Schadwinkel and Gutschalk, 2010). These conflicting results could reflect the fact that a variety of different tasks have been used in these studies to probe selective attention effects. Thus, the results could have been affected by activation due to the specific requirements of the listening task.

Studies in animals and humans using high-spatial resolution methods have shown that auditory representations are modulated during active listening. For example, selective attention enhances the representation of the attended frequency in the tonotopic AC (Da Costa et al., 2013; Fritz et al., 2010; Riecke et al., 2017) and spatial tasks sharpen the spatial tuning of A1 neurons (Lee and Middlebrooks, 2010). Thus, activation modulations during pitch and location tasks may be due to both local (at the level of single or small cortical area) and global (larger areas) effects.

1.5 ACTIVATION DURING DISCRIMINATION AND N-BACK MEMORY TASKS

Only a few previous human studies have investigated whether the specific requirements of attention-engaging auditory tasks modulate activation in AC. Rinne et al. (2009) compared activation in AC during discrimination and *n*-back memory tasks with each other and with activation during a visual task. During all tasks, similar pitch-varying sounds consisting of two 100-ms parts were presented. In the discrimination task, subjects were required to press a button when the parts of a pair were identical in pitch. The task required detailed acoustical analysis and comparison of the sound pair parts. In pitch *n*-back tasks, subjects indicated when the sound pair belonged to the same pitch category (low, medium, high) as the one presented one, two, or

three trials before. In contrast to the pitch discrimination task, this task was performed based on pitch categories and the slight within-pair pitch differences were not relevant. During the visual task, subjects were to ignore the sounds and to detect changes in visual stimuli. The visual task was used to measure stimulus-dependent activation to pitch-varying sounds in the absence of directed auditory attention. Results showed that both discrimination and *n*-back memory tasks were associated with enhanced activations (vs. visual task) in posterior STG and anterior insula (**Figure 2 a**). Regions in anterior to mid-STG showed stronger activation during the pitch discrimination than *n*-back task, whereas the pitch *n*-back task showed enhanced activation in posterior STG and IPL (**b**). Further, they found that activation in IPL increased (not shown) and activation in wide areas of anterior STG and insula decreased with increasing *n*-back difficulty (**c**).

Based on these results, the authors (Rinne et al., 2009) argued that the enhanced activation in anterior STG during pitch discrimination was due to detailed pitch processing and that the enhanced IPL activation during pitch *n*-back memory tasks was related to working memory and categorical processing. Further, they suggested that the activation decrease in STG during *n*-back memory tasks was due to halting of pitch analysis in order to save resources and time for the actual memory task. That is, the authors suggested that operations in STG and IPL are dynamically connected.

In a subsequent study, Rinne et al. (2012) reported quite similar activation patterns during spatial discrimination and spatial *n*-back tasks (**Figure 2 d–f**). Thus, it is clear that the activation patterns observed during discrimination and *n*-back memory tasks cannot be due to the processing of pitch or spatial features of the sounds but are due to specific requirements of the discrimination and *n*-back memory tasks. This conclusion was further supported by the results of Harinen and Rinne (2013), who reported similar task-dependent activation patterns during discrimination and *n*-back tasks performed on vowels. Harinen and Rinne (2014) also investigated whether the activation patterns during *n*-back memory tasks are due to general requirements for working memory or categorical processing. To this end,

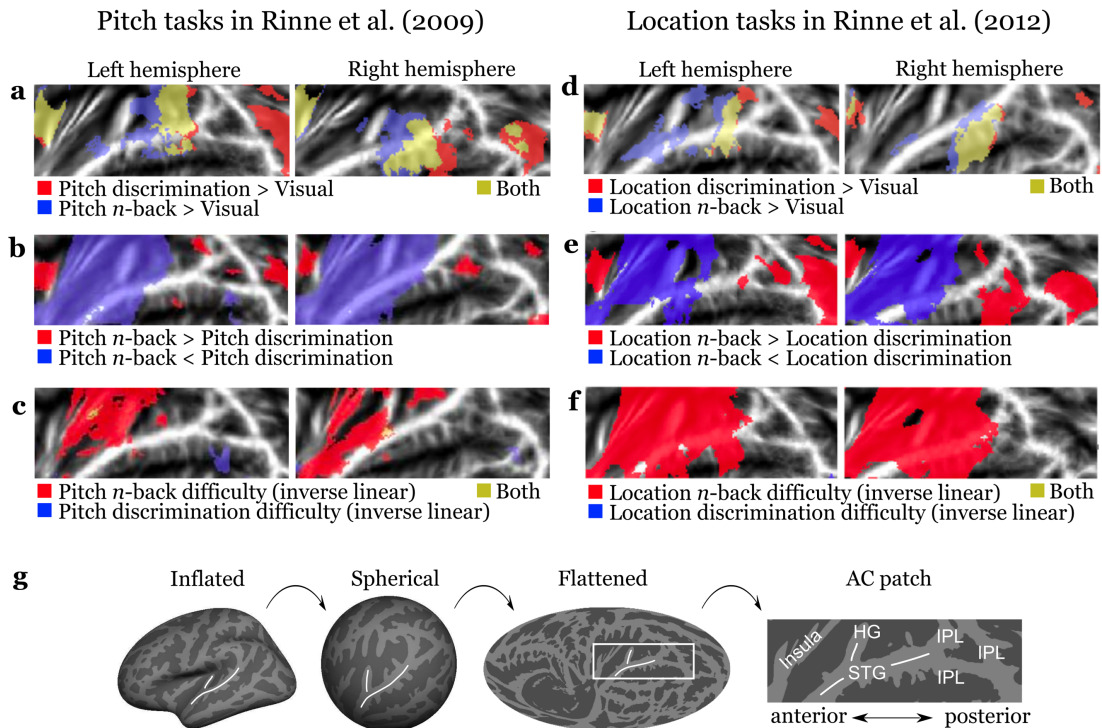


Figure 2 Task-dependent activation during discrimination and *n*-back memory tasks in Rinne et al. (2009, 2012) studies ($N = 17$, threshold $Z > 2.3$, cluster-corrected $p < 0.05$ in both studies). **(a)** Areas showing enhanced activation during pitch discrimination or pitch *n*-back memory task as compared with visual task with the same sounds. **(b)** Comparison of discrimination and *n*-back memory task activation. **(c)** Results of a linear inverse contrast revealing areas where activations decreased with increasing task difficulty. **(d–f)** The corresponding comparisons for location discrimination and location *n*-back memory tasks. **(g)** Data were projected onto cortical surfaces, aligned in spherical space, and flattened to 2D. Data analysis focused on a cortical patch including AC and adjacent areas (white rectangle). STG superior temporal gyrus, HG Heschl's gyrus, IPL inferior parietal lobule. **(a–f)** Adapted from Rinne et al. (2009, 2012). Reproduced with permission from Society for Neuroscience and Elsevier.

their subjects also performed a category discrimination task, in which the discrimination task was performed based on categorical rather than acoustical information (i.e. in a target vowel pair both vowels belonged to the same vowel category). They found that unlike the vowel (acoustical) discrimination task, the category discrimination task and 2-back memory tasks were both associated with strong IPL activation. Based on these results,

the authors suggested that IPL activation is associated particularly with operations on categorical representations.

The prevalent theoretical models are not able to predict the task-dependent modulation observed in AC during discrimination and *n*-back memory tasks. Thus, to better understand the functional significance of these effects, the dynamics of AC activation during active listening tasks have to be investigated in more detail.

2 AIMS OF THE PRESENT THESIS

The present thesis used fMRI and EEG in order to better characterize the spatial, temporal, and network dynamics of activation in human AC during active listening tasks.

Study I investigated the relationship between stimulus-dependent and task-dependent activation in AC during pitch and location processing. Activation in AC to sounds varying in pitch, location or both was measured using fMRI during discrimination, *n*-back memory and visual tasks (similar tasks as in Rinne et al., 2009, 2012). It was hypothesized that (1) stimulus-dependent processing of pitch and location is associated with distinct activation in anterior and posterior STG, respectively, and that these activation patterns are observed independent of the task, (2) task-dependent activation patterns during discrimination and *n*-back memory tasks are similarly observed irrespective of whether these tasks are performed based on sound pitch or location, and (3) pitch and location tasks enhance activation especially in those areas that also show stimulus-dependent sensitivity to these dimensions.

Study II addressed the question of whether the task-dependent activation patterns observed in fMRI by Rinne et al. (2009) can also be detected and investigated using source analysis of the scalp-recorded EEG. This would make it possible to investigate the temporal dynamics of task-dependent activation in AC. It was hypothesized that EEG source analysis would show (1) enhanced AC activation during auditory tasks, (2) distinct activation patterns during pitch discrimination and pitch *n*-back memory tasks, and (3) a systematic modulation of activation in AC as a function of task difficulty in the pitch *n*-back memory task.

Study III investigated whether fMRI connectivity at rest and during active tasks is informative of the functional organization of human AC. It was

hypothesized that (1) an analysis of functional connectivity would reveal a modular and hierarchical organization in AC, and that (2) functional connectivity patterns in AC differ across pitch discrimination and pitch n -back memory tasks. In addition, (3) this study tested the hypothesis (Rinne et al., 2009) that STG and IPL regions are dynamically connected during active listening.

3 METHODS

3.1 SUBJECTS

In all studies of the present thesis, subjects were healthy right-handed adults with normal hearing and normal or corrected-to-normal vision (**Table 1**). The subjects gave written informed consent prior to testing in accordance with the experimental protocol approved by the appropriate ethical review board at the University of Helsinki.

Table 1. Subjects in **Studies I–III**

	<i>N</i>	Males/Females	Age (mean)
Study I	22	14/8	18–36 (25.5)
Study II	17	9/8	20–41 (25.1)
Study III	19	4/15	20–53 (28.2)

3.2 STIMULI

Auditory stimuli. Auditory stimuli were sound pairs that consisted of either iterated rippled noise (IRN) bursts (**Studies I and III**) or pure tones (**Study II; Table 2**). IRN bursts were generated by iteratively delaying and adding noise (Yost, 1996). This results in broadband sounds that are associated with a distinct pitch. In **Study I**, virtual spatial locations were generated using head-related transfer functions (Wightman et al., 1987). To produce diffuse pitch, the IRN bursts were demodulated (Barker et al., 2012) and mixed with white noise (ratio 3:7). To create sounds with a diffuse location, the left and right channels were decorrelated (Culling et al., 2001). In **Study II**, stimuli were identical to those used in Rinne et al. (2009).

Visual stimuli. In **Studies I and III**, the visual stimuli were Gabor gratings and in **Study II** gray rectangles (**Table 2**).

Table 2. Stimulation parameters in **Studies I–III**

	Study I	Study II	Study III
Auditory stimulation			
Stimuli	IRN bursts	Tones	IRN bursts
Duration	90 + 90 ms	100 + 100 ms	90 + 90 ms
Onset-to-onset rate	800–1000 ms	800–1000 ms	800–1000 ms
Pitch ¹	Corresponding to 200–1400 Hz (200 levels) or diffuse	18 levels (200–2517 Hz)	Corresponding to 100–1400 Hz (60 levels)
Location	-120–120° (25 levels) or diffuse	Fixed (middle of the head)	Fixed (middle of the head)
Within-pair difference ²	Pitch 18–88 Hz/ Location 20–60°	Pitch 20–472 Hz	Pitch 10–20 Hz
Target ratio ³	0.13–0.2	0.11–0.22	0.14–0.18
Visual stimulation			
Stimuli	Gabor gratings	Rectangles	Gabor gratings
Duration	100 ms	200 ms	100 ms
Color	Fixed	(186, 186, 186) or (194, 194, 194)	Fixed
Orientation	18 levels ($\pm 180^\circ$)	-	18 levels ($\pm 180^\circ$)
Frequency	6 levels (0.4–1 cpd)	-	6 levels (0.4–1 cpd)
Difference	Orientation 20°/ Frequency 0.2 cpd	Color	Orientation 20°
Onset-to-onset rate	240–320 ms	800–1000 ms	240–320 ms
Target ratio	0.04–0.07	0.11–0.22	0.04–0.05
Experiment			
Blocks per condition	10	16	10
Task block duration	12.7 s	15 s	20 s
Rest block duration	7 s	8 s	12 s
Duration	62 min	40 min	37 min
¹ IRN pitch equals the reciprocal of used delay.			
² The difference was chosen individually during training in Studies I and II .			
³ There were no targets in discrimination PoLo tasks in Study I .			

3.3 TASKS

In all studies, the stimuli were presented during auditory discrimination, auditory *n*-back memory, and a visual target detection tasks (**Table 3**).

Table 3 Experimental conditions in **Studies I–III**

Discrimination tasks ¹	N-back memory tasks	Visual control task
Study I²		
Pitch discrimination P1L1	Pitch category 1-back P1L1	Target detection (P1L1)
Pitch discrimination P1LO	Pitch category 1-back P1LO	Target detection (P1LO)
Pitch discrimination POLO	Pitch category 2-back P1L1	Target detection (POL1)
Location discrimination P1L1	Pitch category 2-back P1LO	Target detection (POL1)
Location discrimination POL1	Location category 1-back P1L1	
Location discrimination POLO	Location category 1-back POL1	
	Location category 2-back P1L1	
	Location category 2-back POL1	
Study II		
Pitch discrimination (easy, medium, hard)	Pitch category 1-back Pitch category 2-back Pitch category 3-back	Target detection
Study III		
Pitch discrimination	Pitch category 1-back Pitch category 2-back Pitch category 3-back Pitch direction 1-back Pitch direction 2-back Pitch direction 3-back	Target detection
¹ In Study I discrimination targets were sound pairs with a pitch or location change (in different blocks). In Study II , targets were sounds without a pitch change. In Study III , all pitch change types (rising, falling, no change) were targeted in different blocks. ² In Study I , sounds had a salient or diffuse pitch (P1/o) and location (L1/o).		

Discrimination task. In discrimination tasks, subjects had to compare the two parts of the sound pair and press a button when the parts were identical or different, depending on the specific task instructions (**Figure 3 a**).

N-back memory task. In *n*-back memory tasks, subjects were required to indicate when a sound pair belonged to the same category as the one presented one, two or three trials before, respectively (**Figure 3 b**). In the pitch category *n*-back tasks, there were three categories defined by pitch level range (low, middle, high). In the location category *n*-back tasks, there were three categories defined by location range (left, middle, right). In the pitch

direction n -back tasks (**Study III**), subjects first had to discriminate the pitch of sounds in a pair (rising pitch, falling pitch, no pitch change) and then to perform the n -back task using the discrimination result.

Visual task. Visual tasks were simple but demanding target detection tasks. These tasks were used to measure activation to sounds in the absence of directed auditory attention. In **Studies I** and **III** subjects reported changes in Gabor orientation or spatial frequency (in separate blocks) and in **Study II** luminance changes of a flickering gray rectangle.

In **Study I**, the discrimination and n -back memory tasks were performed on sounds in which either both pitch and location were salient (P1L1), one feature was salient and the other diffuse (P1LO or PoL1), or both features were diffuse (PoLo). In **Studies II** and **III**, pitch was salient and location fixed to the middle of the head.

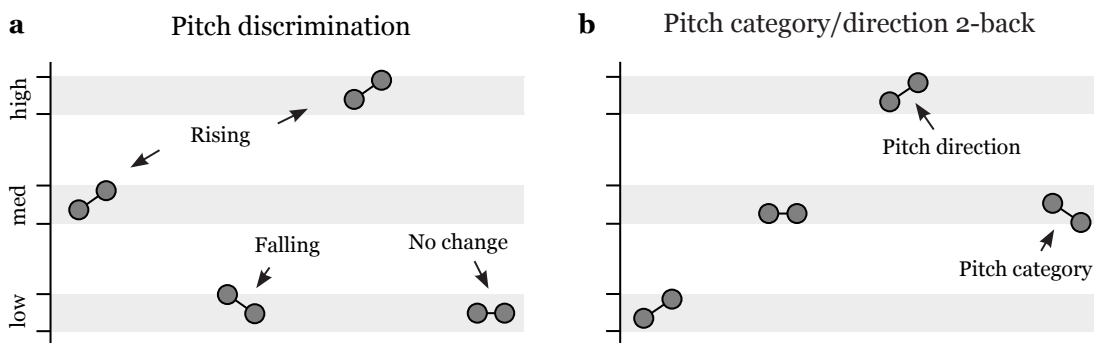


Figure 3 Auditory tasks in **Studies I–III** (different target types indicated by arrows). In all studies, discrimination (a) and n -back memory (b) tasks were performed on pairs of sounds in which the sounds were either identical or had a slight difference in pitch and/or location.

3.4 PARADIGMS AND PROCEDURES

In all studies, the stimuli were presented in task blocks alternating with rest blocks. A task instruction symbol appeared a few seconds prior to a task block and was presented until the end of the block. During rest blocks, subjects were instructed to focus on a fixation cross presented in the middle of the screen. In all task blocks, the auditory and visual stimuli were presented in concurrent but asynchronous streams.

During fMRI (**Studies I and III**), auditory stimuli were delivered via plastic tubes through a porous EAR-tip (ER3, Etymotic research, Elk Grove Village, IL, USA) at a comfortable listening level. Scanner noise (ca. 97 dB LAeq) was attenuated by earphones, circumaural ear protectors and viscoelastic mattresses inside and around the head coil. The visual stimuli and task instruction symbols were presented in the middle of a screen viewed through a mirror attached to the head coil. During EEG (**Study II**), auditory stimuli were presented via circumaural headphones (MDR-7509HD, Sony Corporation, Tokyo, Japan) at 60 dB SPL above hearing threshold. Visual instructions and stimuli were presented on a computer screen placed circa 1.5 m in front of the subject.

3.5 BEHAVIORAL TRAINING AND ANALYSIS OF TASK PERFORMANCE

Subjects were carefully trained in the tasks before measurement. During the behavioral training, subjects were informed that the tasks were intentionally very demanding and required maximum effort especially during the most difficult levels. In addition, behavioral data was acquired during fMRI and EEG measurements to make sure that the subjects successfully performed the intended tasks during the measurement. In the analysis of task performance, responses occurring between 200 and 1300 ms from target onset were accepted as hits. Mean hit rates, false alarm rates, reaction times, and performance (d') were calculated separately for each task. Subjects successfully performed the demanding tasks in all studies.

3.6 STUDY I. PROCESSING OF PITCH AND LOCATION IN HUMAN AUDITORY CORTEX DURING VISUAL AND AUDITORY TASKS

3.6.1 METHODS

Study I investigated the relationship between stimulus-dependent and task-dependent activation in AC during pitch and location processing. Subjects

were presented with sound pairs in which both pitch and location varied (P1L1), only one feature varied (e.g. pitch P1Lo) and the other feature was diffuse, or both features were diffuse (PoLo). In different blocks, subjects performed pitch/location discrimination, pitch/location category 1–2-back memory, and visual target detection tasks (**Table 3**).

MRI imaging was carried out at 3.0 T (Siemens MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) using a standard 20-channel head coil. First, a high-resolution anatomical image (sagittal slices, slice thickness 1.0 mm, in-plane resolution 1.0 mm × 1.0 mm) was acquired. The functional images (GE-EPI; TR 2070 ms, TE 30 ms, FOV 18.9 cm × 18.9 cm², slice thickness 2.0 mm with no gap, interleaved acquisition, 27 slices) were acquired in two runs with a short break in between. Based on the anatomical image, the EPI slices were tilted to be in line with the Sylvian fissures (i.e. the slices were nearly axial). The imaged area covered the superior temporal lobe, insula, and most of the inferior parietal lobe in both hemispheres.

The preprocessing and statistical analysis of fMRI data was performed using the FSL software (www.fmrib.ox.ac.uk/fsl). FreeSurfer (www.freesurfer.net) was used for reconstruction of cortical surfaces and coregistration. Functional data were motion-corrected, resampled to the standard cortical surface, surface-smoothed (5 mm FWHM), and high-pass filtered (cutoff 100 s). Global vertex-wise analysis was performed in surface space. The hemodynamic response to task blocks was modeled using a gamma function (mean lag 6 s, SD 3 s) and its temporal derivative. A second level analysis using fixed effects combined the data from the two runs. Third level group analysis compared parameter estimates (from second level analysis) using permutation inference (Winkler et al., 2014). Correction for multiple comparisons was performed using cluster-mass correction (cluster-forming threshold $Z > 2.3$).

Region-of-interest (ROI) analysis was conducted to study mean percent signal changes in inferior parietal lobule.

3.6.2 RESULTS

The general patterns of fMRI activation to task-irrelevant pitch and location were very similar during all visual and auditory tasks. Stimulus-specific activation patterns to task-irrelevant pitch and location during the visual task were analyzed using a repeated-measures ANOVA with factors pitch (salient, diffuse) and location (salient, diffuse). This ANOVA showed a main effect of pitch in anterior–middle STG and lateral HG and a main effect of location in middle–posterior STG and PT (**Figure 4 b**). The interaction between pitch and location was not significant. Direct contrasts showed that the activation patterns were similar irrespective of whether one or both features varied (e.g. similar pitch activation in P1L1 vs. POL1 and P1LO vs. POLo comparisons). Comparisons of location tasks with or without task-irrelevant pitch (P1L1 vs. POL1) showed activation differences in anterior–middle STG, as during the visual task, but also decreased activation in IPL (**c–e**; decreased activation in IPL was observed during discrimination task at $p < 0.05$ uncorrected). Consistently, task-irrelevant location (P1L1 vs. P1LO) was associated with activation in posterior STG and PT (**h–j**). In sum, the observed activation pattern to pitch and location is in line with the auditory “what”/“where” model (Rauschecker and Tian, 2000) in which pitch and location are processed in separate regions in anterior and posterior AC, respectively.

Although the stimulus-specific activation patterns during visual task were consistent with the idea of independent parallel pathways for pitch and location, task-irrelevant pitch modulated activation in PT and IPL during location tasks. Comparison of location POL1 and pitch P1LO tasks revealed significant activation enhancements for location POL1 tasks (i.e. the total effect of an active location task) mainly in IPL (**k**). However, when activation during location P1L1 tasks (i.e. location task with task-irrelevant pitch variation) and pitch P1LO tasks was compared, activation enhancement associated with location P1L1 tasks was observed also in PT (**l**). Further, during location tasks IPL activation was lower when the sounds contained task-irrelevant pitch (**e**) and during pitch discrimination tasks when pitch

was salient (i.e. pitch discrimination P1Lo vs. PoLo). These results suggest that pitch and location processing are not fully independent during active listening tasks.

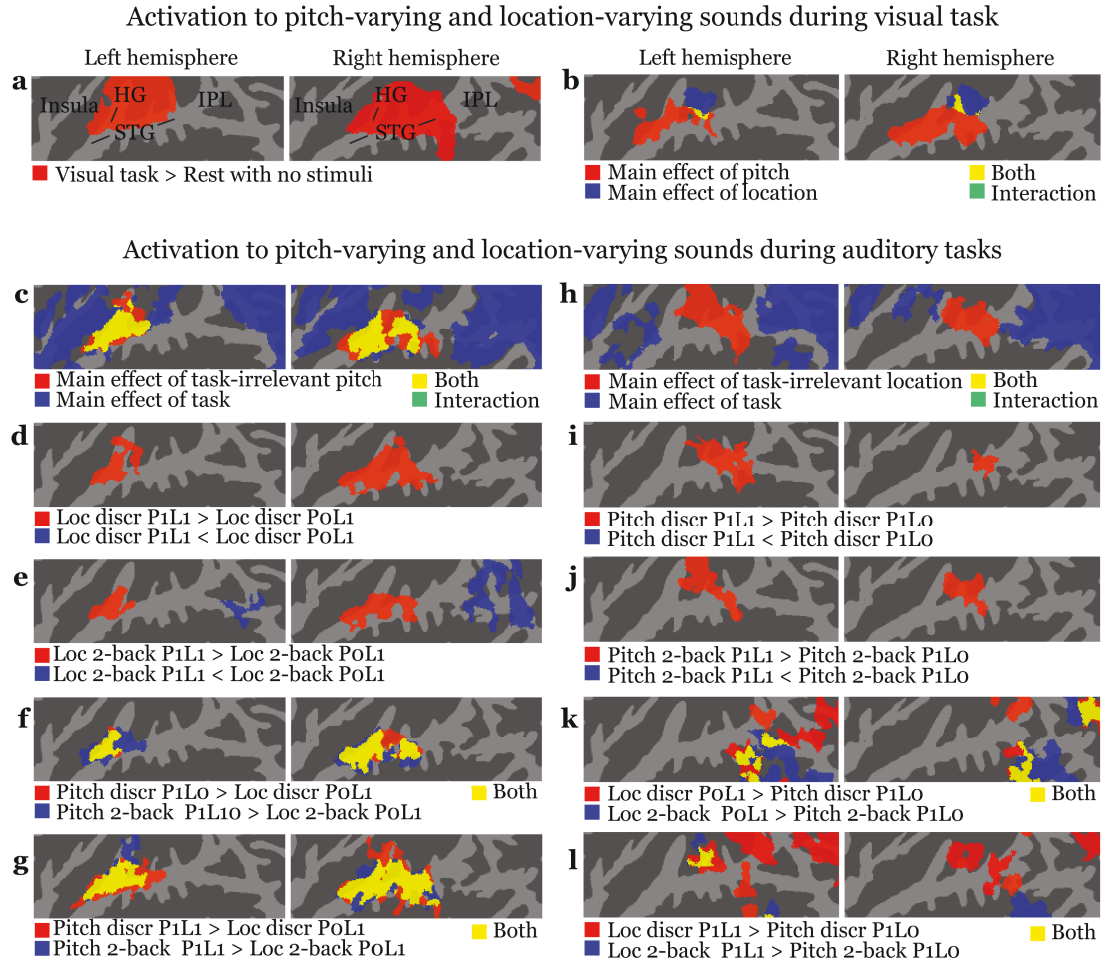


Figure 4 Activation to pitch-varying and location-varying sounds during visual (**a–b**) and auditory (**c–l**) tasks ($N = 22$; $p < 0.05$, cluster-corrected $Z > 2.3$). (**a**) Areas in red showed enhanced activation to pitch-varying and location-varying sounds presented during visual tasks (vs. silent rest). STG superior temporal gyrus, HG Heschl’s gyrus, IPL inferior parietal lobule. (**b**) ANOVA comparison of visual task activation with factors pitch (salient, diffuse) and location (salient, diffuse). (**c**) ANOVA comparison of auditory task activation with factors task-irrelevant pitch (salient, diffuse) and task (discrimination, 2-back). (**d**) Areas showing pitch sensitivity during location discrimination task. (**e**) Areas showing pitch sensitivity during location 2-back task. (**f**) Areas where activation was stronger during pitch P1Lo than location PoL1 tasks. (**g**) Areas where activation was stronger during pitch P1L1 than location PoL1 tasks. (**h–l**) The corresponding comparisons for location.

ANOVA analysis of pitch and location tasks performed on identical (P1L1) sounds revealed no main effect of task-relevant feature or interaction between task and task-relevant feature (**Figure 5**). Task-dependent activation during discrimination and *n*-back memory tasks was similar irrespective of whether the tasks were performed on sounds with or without salient and varying pitch and location (e.g. main effects in **Figure 5** and **Figure 4 c and h**). Therefore, even though selective attention to a specific pitch or location is known to affect AC representations (Fritz et al., 2010; Lee and Middlebrooks, 2010; Riecke et al., 2017), these results indicate that fMRI activation in AC during pitch and location tasks cannot be explained by enhanced stimulus-specific processing alone.

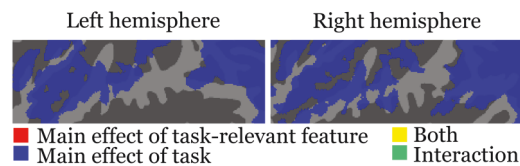


Figure 5 ANOVA comparison of pitch and location task activation with the same sounds ($N = 22$; $p < 0.05$, cluster-corrected $Z > 2.3$).

3.7 STUDY II. SOURCE ANALYSIS OF EVENT-RELATED POTENTIALS DURING PITCH DISCRIMINATION AND PITCH MEMORY TASKS

3.7.1 METHODS

The experiment tested whether source estimation of ERPs can be used to investigate task-dependent activation in AC during pitch discrimination and pitch *n*-back memory tasks. The experimental design was identical to that in the previous fMRI study (Rinne et al., 2009). Subjects were presented with 200-ms sound pairs and, in different blocks, performed pitch discrimination (easy, medium, hard), pitch category 1–3-back memory, and visual target detection tasks with the same sounds (**Table 3**).

EEG was recorded using 136 active scalp electrodes (sampling rate 512 Hz, bandwidth 128 Hz; ActiveTwo amplifier system, Biosemi, Amsterdam, The Netherlands). The locations of all electrodes were measured relative to the nasion and preauricular points using a 3D digitizer (Fastrack 3D, Polhemus, Colchester, VT, USA) for registration with MRI data.

Auditory ERPs and their sources were analyzed using the MNE software (martinos.org/mne). The data were re-referenced to common average, bandpass-filtered (0.5–40 Hz), divided into 900-ms epochs (–100 to 800 ms relative to tone onset), and baseline-corrected (–100 to 0 ms). First two epochs of each task block, epochs associated with a target in the pitch discrimination task, epochs associated with a button press (–300 to 1100 ms), and epochs with extensive extracerebral artefacts ($> 120 \mu\text{V}$ change) were excluded. Finally, the epochs were averaged separately for each condition.

ERP sources were analyzed using cortically distributed minimum-norm estimation (MNE; Hämäläinen and Ilmoniemi, 1984; Lin et al., 2006). A three-layer boundary-element model based on individual anatomical MRI images was used in the source analysis. Source space was defined by grid of white matter surface (ca. 5 mm spacing) with depth weighting and a loose orientation constraint. For group analysis, the individual cortical surfaces were normalized based on cortical folding patterns similar to **Study I** and MNEs were spatially smoothed (7 iterations) and baseline-corrected (–100 to 0 ms). Analysis was restricted to the STG–IPL region as in the previous study by Rinne et al. (2009).

Significance and latency of task-dependent effects was examined comparing the scalp potential distributions associated with different tasks using channel-wise tests and a topographic ANOVA (Manly, 1991). MNEs at these latencies were then investigated using repeated-measures *t* tests and repeated-measures ANOVAs at each time point. In all tests, the result was considered significant if $p < 0.05$ was found during at least 11 consecutive time points (21.5 ms). For illustration, the results of these statistical tests were collapsed into 50 ms bins so that each source point in a bin was assigned the most significant value during that bin.

3.7.2 RESULTS

Figure 6 shows scalp potential distributions to sounds during discrimination and 2-back memory tasks. Analysis of the distributions suggested that tasks modulated ERPs to sounds and that these modulations were different during pitch discrimination and pitch n -back memory tasks (ca. 200–700 ms from the onset of the sound pair).

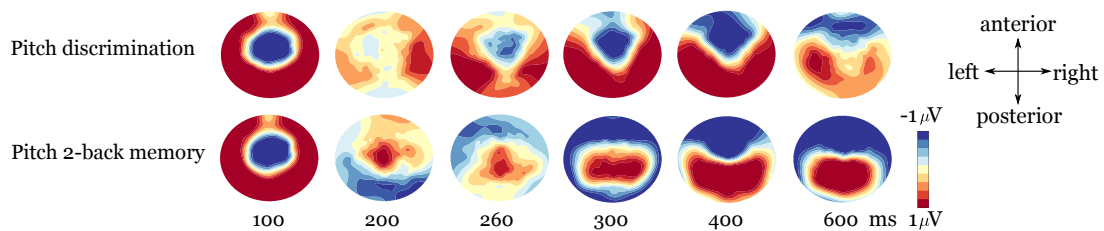


Figure 6 Scalp potential distributions (average reference) during pitch discrimination and pitch 2-back memory tasks at selected latencies 100–600 ms from the onset of the sound pair.

Figure 7 shows results of source analysis of the scalp-recorded ERPs. Mean MNEs during the 100–700 ms period showed enhanced activity in the STG–IPL region during both tasks relative to the visual task. The specific patterns were different during pitch discrimination and pitch n -back memory tasks. During the discrimination task (left), there was enhanced activation in the bilateral posterior STG and left IPL (200–400 ms), in the left anterior STG and insula (350–700 ms), in the right anterior STG (300–450 ms), and in the right posterior STG and IPL (450–700 ms). In the n -back memory task (right), activation first decreased relative to the visual task in the left anterior STG and HG (200–300 ms) and then increased in the left IPL (500–650 ms) and in the right IPL (250–700 ms). AC activation during the n -back memory task was modulated by task difficulty so that the more difficult n -back tasks were associated with weaker activation in STG and insula (150–250 and 550–700 ms) and stronger activation in the IPL (e.g. at 550–650 ms).

In the n -back memory tasks, the decreased STG activation during n -back memory task occurred at 150–300 ms at a latency consistent with the hypothesis that spectrotemporal analysis is actively halted as soon as

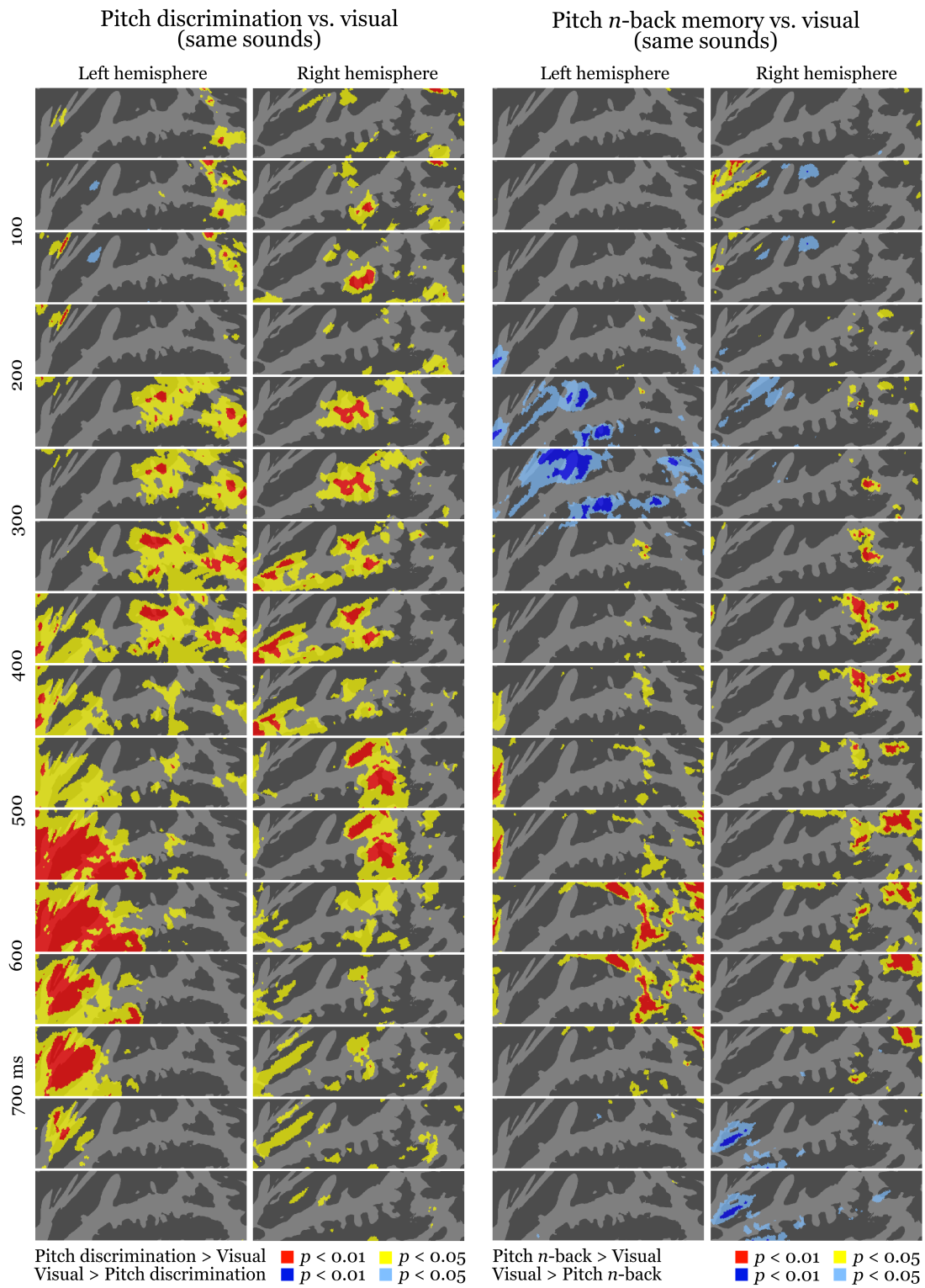


Figure 7 Temporal dynamics of MNEs in AC ($N = 17$, $p < 0.05$ for at least 21.5 ms). Comparison of MNEs to the same tones presented during the pitch discrimination and visual tasks (left) and the pitch memory and visual tasks (right).

category information has been obtained. Similarly, the enhanced IPL activation during *n*-back memory tasks at 200–700 ms suggests that this activation was related to operations on pitch categories after pitch analysis was completed (within 200 ms from sound onset; Butler and Trainor, 2012; Krumbholz et al., 2003; Massaro et al., 1976). The enhanced activation in STG during the discrimination task was also rather late (300–700 ms) and probably not associated with enhanced stimulus-specific processing of pitch.

In sum, source analysis showed enhanced AC activation during auditory tasks, distinct activation patterns during the pitch discrimination and pitch *n*-back memory tasks, and a systematic modulation of activation in AC as a function of task difficulty during the pitch *n*-back memory task. The spatial pattern and sign of effects were remarkably similar to the activation patterns in the previous fMRI study. This suggests that ERP source analysis can be used to complement fMRI to investigate these task-dependent activation patterns in human AC. It is important to note, however, that a multitude of factors contribute to the effects observed in EEG and fMRI during active listening and that the relationship between the fMRI signal and neural activation is not well understood (Logothetis, 2008; Singh, 2012).

3.8 STUDY III. INTRINSIC, STIMULUS-DRIVEN AND TASK-DEPENDENT CONNECTIVITY IN HUMAN AUDITORY CORTEX.

3.8.1 METHODS

Study III investigated whether (1) functional connectivity patterns in human AC reveal a modular structure consistent with the primate model, (2) AC connectivity patterns are task-dependent during discrimination and *n*-back memory tasks, and (3) the relationship between operations in STG and IPL is reciprocal during *n*-back memory tasks. Subjects were presented with identical pitch-varying sounds during pitch discrimination, pitch category 1–3-back, pitch direction 1–3-back, and visual target detection tasks (**Table 3**).

MRI imaging was carried out as in **Study I**. Connectivity analysis was performed using the CONN toolbox (www.nitrc.org/projects/conn). Data were slice-timing corrected, motion-corrected, and spatially smoothed (5 mm FWHM in volume). Data were detrended and high-pass filtered (0.008 Hz). Before the computation of connectivity measures, motion outliers were excluded and white matter average signal, signal averaged over the ventricles, main effects of run and task and their first temporal derivatives, and motion effects (movement and rotation along three orthogonal axes) were regressed out from the BOLD timeseries.

Network nodes in AC and adjacent regions in both hemispheres were defined based on FreeSurfer's IC3 mesh (101 nodes in each hemisphere). Networks were estimated using Pearson's correlation and generalized psychophysiological interactions (gPPIs). To investigate network topology during rest and task conditions, the correlation matrices were Fisher's z transformed, averaged across subjects, and binarized. Modular network structure was estimated using the InfoMap algorithm (Rosvall and Bergstrom, 2008). Normalized mutual information (Meilă, 2007) was used to compare the modular structure associated with different task conditions. The network nodes were further characterized using the participation coefficient (Guimerà and Nunes Amaral, 2005) to define areas of high or low inter-modular connectivity and the global variability coefficient (GVC; Cole et al., 2013) to detect areas showing the strongest task-dependent connectivity modulation.

Multivariate pattern analysis (MVPA) was used to evaluate whether the functional connectivity patterns contain task-specific information. Similarity among the connectivity patterns (gPPI beta values) was measured based on Spearman's rank correlation. Pattern classification was then used to separately compare each pair of conditions. Linear support vector machines were trained and tested on gPPI data using leave-one-subject-out cross-validation (PyMVPA; www.pymvpa.org). Standard z normalization was applied to the data (separately for each connection). In each fold of the cross-validation, test data were used to select the most informative connections (top 5% F scores) to be included in the classification. Significance of

classification accuracy was assessed using permutation testing (1000 permutations of random label orders; FDR-correction).

Finally, node-to-node gPPI analysis investigated the specific effects of task performance on the node-to-node gPPI connectivity patterns. In this analysis, results of the repeated-measures two-tailed t tests were FWER-corrected using the network-based statistics method (NBS; based on intensity values, primary threshold $p < 0.001$; Zalesky et al., 2010).

Activation analysis was carried out similar to **Study I**.

3.8.2 RESULTS

Analysis of connectivity patterns revealed a modular organization in both hemispheres (**Figure 8 a**). The estimated modular structure was quite consistent across hemispheres and during rest, visual, and auditory tasks (0.68–0.93 normalized mutual information at 0.15 density). During rest, inter-modular connectivity was high in regions surrounding the STP module and low in the STP and IPL modules (**b**). By contrast, during auditory and visual tasks (with identical stimuli), nodes with higher inter-modular connectivity than rest were observed bilaterally in STP, aSTG and IPL and in the right pSTG (**c**). High GVC across all tasks (i.e. high variability of the overall connectivity pattern across all tasks, independent of modular structure) was found particularly in bilateral aSTG and IPL (**d**). Thus, in support of the idea that functional connectivity is informative of the organization of human AC, connectivity patterns characterized a modular organization in the STG–IPL region that is well in line with primate models.

While connectivity patterns during different tasks were overall very similar ($r_s > 0.76$ in all pair-wise comparisons), pattern classification analysis confirmed that the connectivity patterns were modulated during active auditory task performance (**Figure 9**). The binary classifications were successful between conditions with auditory stimulation (auditory and visual tasks) and (silent) rest (accuracy > 89%, FDR-corrected $p < 0.01$ for all tests), between auditory and visual tasks (accuracy > 79%, FDR-corrected

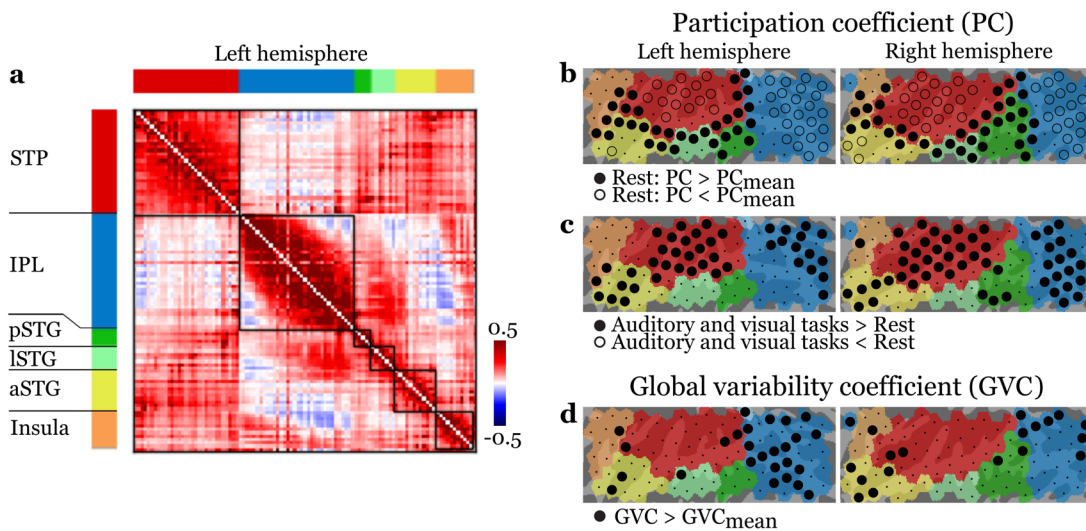


Figure 8 Connectivity-based parcellation and topography of AC. **(a)** Correlation matrix of left hemisphere nodes at rest ordered by parcellation at 0.15 density. Supratemporal plane (STP, red), anterior STG (aSTG, yellow), lateral STG (lSTG, light green), posterior STG (pSTG, dark green), insula (orange) and IPL (blue). For anatomical correspondence, see **(b–d)**. **(b)** Inter-modular connectivity measured using the participation coefficient during rest and task conditions ($N = 19$; FDR-corrected $p < 0.01$; modules estimated at 0.15 density). Nodes showing higher (black circles) or lower (open circles) inter-modular connectivity during rest as compared to the mean across all nodes. **(c)** Comparison of inter-modular connectivity during tasks and rest. **(d)** Global variability coefficient (GVC) computed across all auditory and visual tasks (with identical stimulation). Black circles show nodes with significantly higher than mean GVC across all nodes. Modular structure was not used in GVC calculations and is shown only for illustration.

$p < 0.01$ for all tests), and between some of the different auditory tasks. The node-to-node gPPI analysis revealed that connectivity was enhanced between IPL and other modules during visual and auditory tasks as compared with rest. During auditory tasks, the set of enhanced IPL connections was more substantial and there was decreased connectivity within several modules (STP, lateral STG modules, IPL). There were also significant effects in comparisons between the pitch discrimination and pitch category/direction 2-back tasks. These results showed that functional connectivity in the auditory STG–IPL network is dynamically modulated depending on the behavioral context. The results also highlight the important role that IPL

plays during the analysis of both task-irrelevant and task-relevant auditory information in humans.

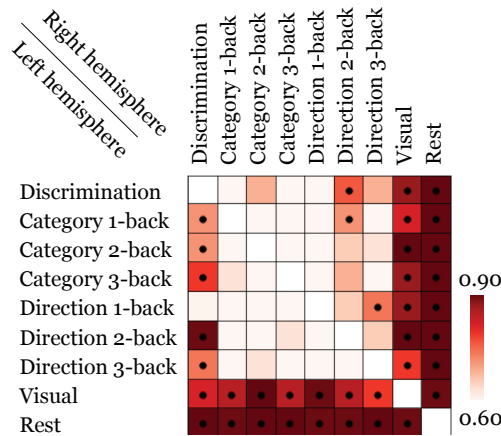


Figure 9 MVPA of task-dependent connectivity patterns. Lower and upper triangular show pattern classification results of pair-wise task comparisons in the left and right hemisphere, respectively. Color shows classification accuracy. Significant (FDR-corrected $p < 0.05$) classifications are indicated by black dots.

Consistent with Rinne et al. (2009), activation was enhanced in STG and IPL during pitch discrimination and pitch category 2-back memory tasks, respectively (**Figure 10 a**). Contrary to the hypothesis that STG and IPL are dynamically connected during active listening, however, STG activation was stronger during pitch discrimination than pitch category/direction 3-back tasks ($t_{18} > 4.7$, $p < 0.001$ for both tests) and there were no significant differences between activation during the pitch category and pitch direction n -back memory tasks ($t_{18} < 2.0$, $p > 0.06$ for all three tests; **b**). IPL activation was stronger during both the 3-back tasks as compared to discrimination ($t_{18} > 2.9$, $p < 0.01$ for both tests) and stronger during the pitch category than pitch direction 2-back and 3-back tasks ($t_{18} > 2.2$, $p < 0.05$ for both tests; **c**). Some task-dependent connectivity differences between STG–IPL were observed. This supports the hypothesis that STG and IPL operations are linked.

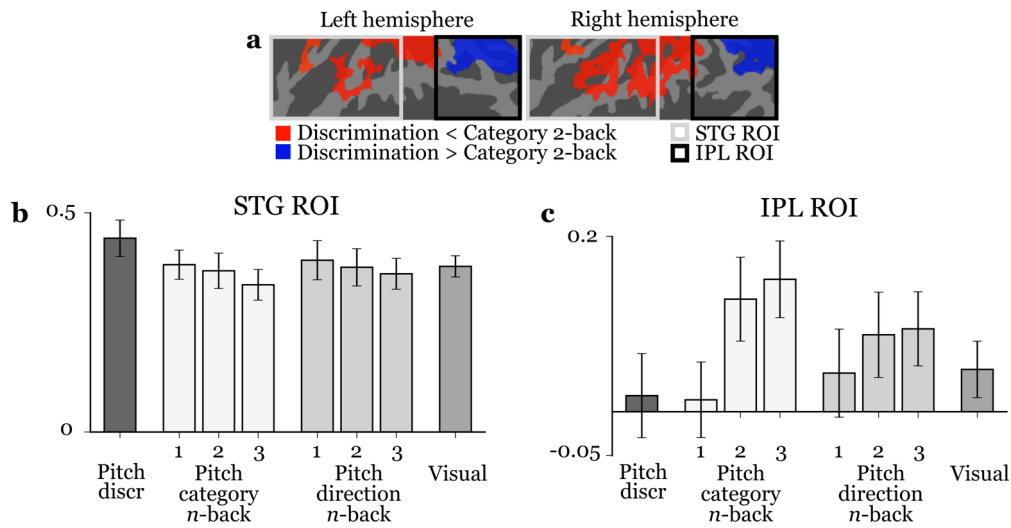


Figure 10 Activation results. **(a)** Comparison of activation during pitch discrimination and pitch category 2-back tasks with identical stimuli ($N = 19$; FWER-corrected $p < 0.05$, cluster-corrected $Z > 2.3$). STG and IPL ROIs are outlined in gray and black, respectively. **(b–c)** Mean ($N = 19$) percent signal change in the STG and IPL ROIs relative to silent rest.

4 DISCUSSION

Results of this thesis show that activation and functional connectivity in human AC are strongly modulated during active listening tasks. **Study I** showed that similar fMRI activation patterns are observed when analogous discrimination and *n*-back tasks are performed on sounds that vary in pitch, location or both, and that these task-dependent patterns are independent of stimulus-level processing of pitch and location. **Study II** showed that the task-dependent activation patterns during the pitch discrimination and *n*-back memory tasks can also be observed using EEG source localization, and that these effects occur relatively late (200–700 ms from sound onset) as compared to stimulus-specific pitch effects (0–200 ms in previous ERP studies). **Study III** found that functional connectivity patterns between STG and IPL are significantly modulated during the presentation of sounds and during active tasks.

4.1 IMAGING THE MODULAR ORGANIZATION OF HUMAN AC

According to the primate model, AC consists of anatomically and functionally distinct regions and areas. Consistently, parcellation based on functional connectivity in **Study III** revealed a modular network structure in the STG–IPL region. A similar modular structure was quite consistently observed during rest and task conditions in line with the idea that the intrinsic connectivity patterns reflect the underlying anatomy (Buckner et al., 2013; Cole et al., 2014; Honey et al., 2009; Krienen et al., 2014). The network structure consisted of six modules in supratemporal plane (STP), superior temporal gyrus (aSTG, lSTG and pSTG), insula, and IPL in both hemispheres. Based on previous anatomical and imaging studies (Baumann et al., 2013; Moerel et al., 2014), it can be estimated that the STP module likely contains core region and some or all of the belt areas. The STP module was surrounded by aSTG, lSTG and pSTG modules, which showed the highest inter-modular connectivity during rest. Thus, these modules might

correspond to the primate belt/parabelt regions that are more globally connected to both low-level and higher-level areas.

The fMRI activation results were consistent with the idea that the modules are functionally specialized (**Table 4**). Contrasts between conditions with pitch-varying and location-varying sounds showed pitch and location sensitivity within the STP module (i.e. putative core/belt areas). Contrasts between discrimination and *n*-back memory tasks, in turn, describe a difference between anterior/posterior STG and posterior STG/IPL (Rinne et al., 2009).

Table 4. A summary of potentially useful functional contrasts to investigate the modules in STG and IPL

Contrast	Module
Sound > Silence	STP
Pitch during auditory and visual tasks	STP (anterior)
Location during auditory and visual tasks	STP (posterior)
Task-irrelevant pitch during auditory tasks (decrease)	IPL
Location task (no task-irrelevant pitch) > Pitch task (no task-irrelevant location)	IPL
Discrimination > <i>n</i> -back memory task	STP, aSTG, mSTG, pSTG (anterior)
Discrimination < <i>n</i> -back memory task	pSTG (posterior), IPL
Category > Direction <i>n</i> -back memory task	IPL
<i>n</i> -back difficulty (linear increase)	IPL, Insula (anterior)
<i>n</i> -back difficulty (linear decrease)	STP, aSTG, mSTG

The group-level functional connectivity and activation effects in **Studies I–III** are clearly informative about the functional organization of human AC. Studies on the functional organization of human AC would benefit from a multimodal approach in which various measures are combined for parcellation (e.g. tonotopy, myelin maps, cortical thickness, task activation contrasts, functional and structural connectivity; Glasser et al., 2016; Parisot et al., 2017). Attention and task contrasts could be especially useful for the parcellation of regions outside the core, as the functional organization of these areas is not well understood in humans.

4.2 THE EFFECT OF ACTIVE LISTENING ON STIMULUS-LEVEL ACTIVATION

The results of **Study I** showed distinct stimulus-specific activation to pitch in anterior–middle STG and lateral HG and to location in middle–posterior STG and PT. However, pitch and location tasks were not associated with a significant activation difference when the tasks were performed on identical sounds varying in both pitch and location. Thus, the results are not consistent with the general idea that pitch and location tasks enhance the processing of task-relevant stimulus dimensions. Rather, the results suggest that fMRI activation in AC during pitch and location tasks is mainly due to (obligatory) stimulus-level processing and distinct task-dependent effects, whereas any effects associated with enhanced processing of task-relevant stimulus information are negligible.

Stimulus-specific activation in **Study I** was not, however, fully independent of the listening task. First, PT showed sensitivity to pitch during auditory but not visual tasks. A number of previous studies have reported sensitivity to both spatial and nonspatial features in PT during active listening (Da Costa et al., 2015; Griffiths and Warren, 2002; Smith et al., 2010; Zatorre et al., 2002). Based on the results of these studies, it has been suggested that PT is involved in the segregation of sound sources (objects) by combining different acoustical cues and previously learnt patterns. Attention biases how an auditory scene is segregated into objects during complex behavioral conditions (Alain and Bernstein, 2008; Carlyon, 2004; Winkler et al., 2009). Thus, the sensitivity to pitch in PT observed in **Study I** might be due to operations related to sound segregation. Second, although the results of **Study I** replicated the common finding that activation in IPL is stronger during spatial tasks performed on spatially varying sounds than during pitch tasks performed on pitch-varying sounds (Arnott et al., 2004; Degerman et al., 2006; Weeks et al., 1999), **Study I** also showed that IPL activation during location tasks was decreased when the sounds varied in pitch. However, the direct comparison between pitch and location tasks performed on identical sounds did not show significant effects in IPL. These results suggest that these IPL activation patterns could be due to effects associated

with the processing of (task-irrelevant) pitch rather than location. This interpretation is also supported by the behavioral results in **Study I** showing that performance in location tasks was lower when the sounds also varied in pitch. Previous behavioral studies have shown that location processing and location memory are affected by task-irrelevant pitch (Delogu et al., 2014; Joseph et al., 2015). Thus, IPL might be involved in these processes. Taken together, the results of **Study I** suggest that selective attention to both pitch and location modulate activation along the posterior “where” pathway. Thus, either operations in the anterior (“what”, pitch) and posterior (“where”, spatial) streams are not fully independent of each other or the posterior stream is involved also in the processing of pitch information. It is important to note that activation within the regions of the “what” and “where” streams is not determined only by stimulus-level information but is also strongly modulated by the requirements of the task at hand.

According to the reverse-hierarchy theory (Hochstein and Ahissar, 2002; Nelken and Ahissar, 2006), the specific requirements of the listening task (e.g. level of abstraction) determine whether task processing involves (detailed, low-level) stimulus-specific representations or whether it can be accomplished based on default high-level abstractions. It could be speculated that the categorical *n*-back memory task was carried out at higher levels of abstraction than the discrimination task, and that the discrimination task required access to acoustical representations and could not be performed based on abstract representations. However, in **Study I** the comparisons between the pitch and location discrimination tasks performed on identical sounds did not show specific modulations in pitch and location sensitive regions. This result is inconsistent with the hypothesis that the required level of abstraction determines whether task processing involves stimulus-specific representations or not. By contrast, the results of **Study I** suggest that the processing of pitch and location information is largely independent of task requirements.

It is important to note that **Study I** investigated the hypothesis that the processing of pitch and location engage distinct areas of AC. The experimental (blocked design) and analysis (univariate GLM) approach of

Study I are well-suited for this purpose (Davis et al., 2014). However, multi-voxel pattern analysis could be more sensitive to whether an area is generally sensitive to the difference between pitch and location task blocks (Formisano and Kriegeskorte, 2012). Further, selective attention to a target frequency modulates activation in tonotopic AC in a frequency-specific manner (Da Costa et al., 2013). The experimental designs (blocked design, no target frequency) used in **Studies I–III** are not optimal for investigating such effects.

4.3 TEMPORAL DYNAMICS OF TASK-DEPENDENT ACTIVATION

The results of **Study II** showed that task-dependent modulation in AC can be investigated using EEG source analysis. Importantly, unlike fMRI, EEG source analysis makes it possible to investigate the temporal dynamics of activation in AC associated with individual sound events at the millisecond scale. **Study II** showed enhanced AC activation during the discrimination task at 300–700 ms, enhanced IPL activation during the *n*-back memory task at 200–700 ms, and decreased AC activation during the *n*-back memory task at 150–300 ms from sound onset.

Previous studies suggest that the initial pitch and location processing in AC is completed within the first 200 ms from sound onset (Bourquin et al., 2013; Butler and Trainor, 2012; De Santis et al., 2007; Griffiths et al., 2010; Gutschalk et al., 2004; Krumbholz et al., 2003; Massaro et al., 1976; Näätänen and Winkler, 1999). Based on these results, it can be estimated that the enhanced STG activation in **Study II** during the pitch discrimination (300–700 ms from sound onset; i.e. 200–600 ms from the second sound part) and enhanced IPL activation during the pitch *n*-back memory tasks (200–700 ms) occur after the completion of the initial pitch analysis. This account is consistent with the results of **Study I** showing that the task-dependent effects associated with pitch and location tasks cannot be explained by enhanced stimulus-level processing. This distinction between stimulus-specific and task-dependent activation effects is not consistent with the early selection theories which predict that attention (i.e. attention-

engaging tasks) directly modulates acoustical analysis based on the relevance of individual features in the current task. The results (task effects occur later than and are independent of stimulus processing) are more in line with late attentional selection in object-based theories of auditory attention (see Section 4.6). However, it should be noted that the discrimination and *n*-back memory tasks used in **Study II** were not designed to test the assumptions of early and late selection theories.

4.4 HIERARCHICAL ORGANIZATION OF HUMAN AC

It is generally assumed that auditory representations become more complex and integrative from lower to higher levels in the hierarchy. Consistently, **Studies I** and **III** showed that stimulus-dependent activation (during the visual task) occurred mainly in areas near HG and attention-related modulation was strongest in the areas outside the HG in lateral STG. However, the results are not fully consistent with previous hierarchical models. First, task-irrelevant pitch modulated activation in PT and IPL during location tasks. This finding is in contrast with the idea that pitch and location are processed in separate “what” and “where” streams that are fully independent of each other. Second, task-dependent modulation during pitch and location tasks was independent of the stimulus-specific effects to pitch and location. Thus, task-dependent effects cannot be explained by a simple model in which pitch tasks enhance pitch processing and location tasks enhance location processing. Third, the presence of stimulus-sensitivity in IPL is inconsistent with models in which processing propagates from stimulus-specific representations in areas near HG to stimulus-independent abstract representations in IPL. Functional connectivity between wide regions of STG and IPL was also enhanced by sound presentation and modulated by auditory attention and task requirements. These results suggest that processing hierarchies are modulated dynamically depending on the requirements of the task at hand, and that IPL is strongly involved in operations in STG.

Anatomical data in primates shows that AC regions are connected with each other via both feedforward and feedback connections (e.g. **Figure 1**). These connections could constitute several ascending and descending processing hierarchies. Results of the present thesis show that IPL plays a strong role in AC operations both during the processing of task-irrelevant sounds and during active listening. In addition, it is likely that AC modulates processing in subcortical auditory nuclei. Together, the results of **Studies I–III** suggest that AC operations should be studied in a wider context including IPL as well as other cortical and subcortical structures.

4.5 FUNCTIONAL SIGNIFICANCE OF TASK-DEPENDENT ACTIVATION IN STG AND IPL

In this thesis, wide regions in anterior and mid-STG showed stronger activation during the auditory discrimination task than when the same sounds were presented during the visual task (no directed auditory attention). Enhanced activation in AC during auditory tasks is often interpreted to reflect enhanced stimulus-specific processing. However, as discussed above, enhanced activation in STG was observed similarly during analogous pitch and location discrimination tasks performed on identical stimuli (**Study I**). Therefore, enhanced activation in AC during discrimination task cannot be explained by enhancement of stimulus-specific processing of pitch and location but is more likely to be related to the requirements of the discrimination task (comparison between the two parts of a sound pair). This interpretation is also supported by the relatively late onset of the effect (**Study II**).

Previous studies have implicated IPL in tasks requiring working memory (e.g. Gaab et al., 2006; Koelsch et al., 2009; Leung and Alain, 2010). Consistently, in the present thesis and in previous studies using similar tasks, activation enhancement in IPL has been observed irrespective of whether the *n*-back task is performed on pitch, location or vowel categories (**Study I**; Harinen and Rinne, 2013; Rinne et al., 2009; 2012). Previous studies have also reported that IPL is activated during a two-vowel discrimination task when the vowels are discriminated based on their vowel category

membership instead of acoustical information (Harinen and Rinne, 2014). This suggests that IPL activation is more related to general operations on categorical representations than to working memory. In **Study III**, IPL activation was stronger during pitch category than pitch direction *n*-back tasks. It is possible these two versions of the *n*-back task required slightly different categorical processing or that the tasks differed in terms of task complexity or mental imagery (Bzdok et al., 2013; Corbetta and Shulman, 2002; Leung and Alain, 2010).

Rinne et al. (2009) hypothesized that the decreased activation in STG and enhanced activation in IPL during *n*-back memory tasks are dynamically linked with each other. According to this hypothesis, STG deactivation in the *n*-back memory task is due to active suppression of processing in STG in favor of categorical processing in IPL. **Study III** explicitly tested this hypothesis by comparing activation during pitch discrimination and pitch category *n*-back memory tasks to that during a novel task in which subjects first had to discriminate the pitch of sounds in a pair (rising pitch, falling pitch, no pitch change) and then to perform the *n*-back task using the discrimination result. As the pitch direction *n*-back memory task required a full acoustical analysis of the sounds, it was expected to show enhanced IPL activation but no decreased STG activation. However, it was found that both *n*-back tasks were associated with a similar decrease in STG activation irrespective of whether the *n*-back task required detailed acoustical analysis of all sounds or not. Thus, the results do not support the hypothesis that STG activation is suppressed when detailed acoustical analysis in STG is halted due to the requirements of categorical *n*-back tasks.

In a slightly different account, Huang et al. (2013) suggested that the decrease in STG activation during *n*-back tasks is due to the suppression or interruption of the processing of task-irrelevant aspects of the sound sequence. The present results are not consistent with this account either as the decrease in STG activation was present in the pitch direction *n*-back task even though all sounds were task-relevant and had to be processed in detail.

If active suppression of acoustical analysis in STG does not play a role, then what might be the explanation for the decreased STG activation during

demanding *n*-back memory tasks? STG and IPL regions could be hemodynamically linked (both are supplied by the middle cerebral artery; Kim and Ogawa, 2012; Leech et al., 2014) such that an increase in IPL activation is associated with a decrease in STG activation. However, **Study II** showed decreased activation in anterior–middle STG during demanding *n*-back tasks also using EEG source analysis. As EEG is a direct measure of electrical activity of neurons, this result suggests that if dynamic linking of STG and IPL contributes to activation patterns during discrimination and *n*-back memory tasks, then this linking occurs at the neuronal level.

Auditory cortical processing involves both excitatory and inhibitory neuronal mechanisms, which could be associated with either increased or decreased metabolic demands (and fMRI signal). It has been suggested that inhibition mechanisms are central in the selection of relevant information during auditory selective attention, categorization, and working memory (e.g. Desimone and Duncan, 1995; Engell et al., 2016; Linke et al., 2011; Luo et al., 2005). Thus, it could be possible that the STG deactivation during the *n*-back memory task is due to such task-related selection process. Specifically, based on fMRI results, it has been suggested that the representations of tones compete with each other during the maintenance period of an auditory short-term memory task and that this competition results in deactivation in STG (Kumar et al., 2016; Linke et al., 2011). Thus, the activation decrement in STG in the present *n*-back memory tasks could be because more acoustical representations are maintained during the demanding *n*-back memory tasks than during 1-back or discrimination tasks. In **Study III**, the STG activation decrease was similarly observed during both the pitch category and pitch direction *n*-back tasks even though the processing requirements in these tasks were vastly different. This suggests that if the decreased STG activation during *n*-back memory tasks is due to the competition of memory representations, then these competing representations contain relatively low-level acoustical information.

4.6 A TENTATIVE MODEL OF TASK-DEPENDENT ACTIVATION IN HUMAN AUDITORY CORTEX

Figure 11 shows a tentative model of task-dependent processing in human AC. Similar to the object-based models of auditory attention (Alain and Arnott, 2000; Shinn-Cunningham, 2008), in the proposed model, auditory tasks operate on auditory objects that are based on, but independent of, stimulus-specific auditory processing. A key idea in the model is that the difference between discrimination and *n*-back memory tasks is related to whether the task requires high or low abstraction.

Behavioral studies suggest that pitch and location are not processed completely separately in the auditory system (Delogu et al., 2014; Dyson and Quinlan, 2004; Joseph et al., 2015; Mondor et al., 1998). The results of these studies suggest that during initial processing these features are integrated, and thus they cannot be directly accessed by selective attention. Rather, auditory tasks are performed based on objects. According to Bizley and Cohen (2013), auditory objects are “the perceptual consequence of the auditory system’s interpretation of acoustic events and happenings”. The formation of such object representation is strongly affected by both stimulus properties and attention (Alain and Bernstein, 2008; Carlyon,

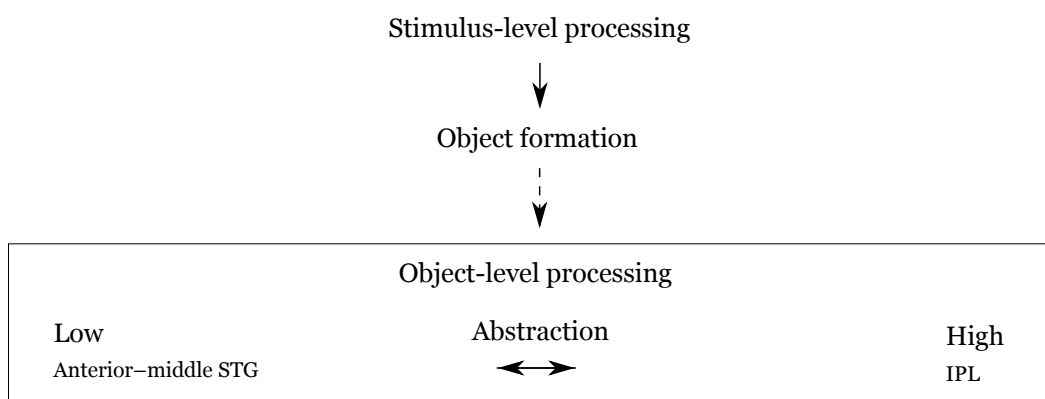


Figure 11 Model of task-dependent processing in human AC. AC is involved in stimulus-level processing (cannot be accessed by task-specific processes), object formation (affected by both stimulus-level processing and the goals of behavior), and object-level processing (task-dependent). Modulation of activation in AC during discrimination and *n*-back memory tasks is because the tasks operate on low or high abstraction, respectively.

2004; Shinn-Cunningham, 2008; Winkler et al., 2009). It is generally thought that the brain areas that process stimulus-specific information are also involved in representing auditory objects. Higher-level regions of AC (e.g. anterolateral STG, PT and STS) enable the selection of object representations via feedback connections (from hierarchically higher to lower regions of AC; Ahveninen et al., 2016; Bizley and Cohen, 2013; Buffalo et al., 2010). The proposed model also assumes that AC operations during listening tasks depend on whether the task operates on low or high abstraction level. Auditory information (objects) cannot be simultaneously processed at a low and high abstraction level. A distinction between low and high abstraction tasks is supported by behavioral evidence. For example, the cost of dual-tasking is higher when the dual tasks consist of two low-abstraction or two high-abstraction tasks than one low-abstraction and one high-abstraction task (Ahissar et al., 2009; Gallun et al., 2007; Hafter et al., 1998; Semal and Demany, 2006). Previous lesion, ERP, and fMRI studies suggest that tasks requiring low or high abstraction are associated with activation in different regions of the temporal lobe (Harinen and Rinne, 2014; Johnsrude et al., 2000; Nahum et al., 2010). In particular, Harinen and Rinne (2014) showed that non-categorical discrimination (low abstraction) tasks are associated with stronger fMRI activation in STG than categorical discrimination (high abstraction) tasks, whereas the categorical task showed stronger activation in IPL. Based on these results, the model assumes that the representation of auditory objects is strongly task-dependent. In low-abstraction tasks, the object representation involves AC areas that encode the stimulus-level information, whereas the more abstract information in high-abstraction tasks is encoded in IPL. During active tasks, activation is enhanced in the areas encoding object information (STG in low-abstraction and IPL in high-abstraction tasks) and in the higher-order regions that control object formation (e.g. lateral STG).

The proposed model explains the lack of strong interactions between task-dependent and stimulus-specific activation (attending to an object modulates representations of all features that define the object; **Study I**) and the relatively late timing of task effects (task-dependent processing can begin

only after stimulus-level processing and object formation; **Study II**). The model is also consistent with the idea that operations in STG and IPL are linked reciprocally so that both cannot occur at the same time. In **Studies I–III**, task-dependent activation in IPL was enhanced during *n*-back memory tasks but decreased in STG and, further, with increasing *n*-back difficulty activation increased in IPL and decreased in STG. In **Study III**, activation was higher in IPL during the pitch category than pitch direction *n*-back task. This could be because the pitch direction *n*-back task required both low and high abstraction processing that cannot occur simultaneously. Finally, functional connectivity between STG areas and IPL was strongly and dynamically modulated during the presentation of sounds (during visual task) and active auditory tasks (**Study III**). These effects might at least partially reflect dynamics between regions associated with low and high abstraction tasks.

In comparison to previous functional models, the suggested model is fully compatible with accounts that describe stimulus-specific activation in low-level AC. Several previous studies have also reported that the stimulus-specific activation is modulated during active listening (e.g. Ahveninen et al., 2006; Da Costa et al., 2013; Fritz et al., 2003). These results seem to suggest that attention modulates stimulus-level processing, i.e. attention effects are not restricted to the object level. It is unclear how these results can be combined with the object-based model of attention. It might be possible, for example, that during some tasks an object is defined by a very limited set of features (e.g. a certain frequency). In such a case, attention-related modulation might be limited to areas that process these features (e.g. a frequency-specific area; Alain and Arnott, 2000). However, the object-based model of auditory attention predicts that, even during such tasks, attention also modulates activation in the higher-level object formation regions.

5 CONCLUSIONS

The results of this thesis show that activation in human AC is strongly modulated by the requirements of the listening task. The task-dependent activation patterns cannot be explained by enhanced processing of stimulus-specific information. The prevalent primate models of AC focus on stimulus-specific processing and are not able to predict such task-dependent modulation. The results support the view that auditory tasks operate at the level of auditory objects that are based on but are independent of stimulus-specific auditory processing and that activation in AC during auditory tasks depends on whether the task requires low or high abstraction. The results also highlight the role of IPL in AC operations. IPL is dynamically connected with STG during sound presentation and during active listening. In subsequent studies, the functional significance of task-dependent modulation in AC should be investigated in the context of wider network of brain areas including IPL.

REFERENCES

- Ahissar, M., Nahum, M., Nelken, I., and Hochstein, S. (2009) Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 364, 285–299.
- Ahveninen, J., Huang, S., Ahlfors, S. P., Hämäläinen, M., Rossi, S., Sams, M., and Jääskeläinen, I. P. (2016) Interacting parallel pathways associate sounds with visual identity in auditory cortices. *NeuroImage*, 124, 858–868.
- Ahveninen, J., Jääskeläinen, I. P., Rajj, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F. H., Sams, M., and Shinn-Cunningham, B. G. (2006) Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences*, 103, 14608–14613.
- Ahveninen, J., Kopčo, N., and Jääskeläinen, I. P. (2014) Psychophysics and neuronal bases of sound localization in humans. *Hearing Research*, 307, 86–97.
- Alain, C., and Arnott, S. R. (2000) Selectively attending to auditory objects. *Frontiers in Bioscience*, 5, D202–D212.
- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., and Grady, C. L. (2001) "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 12301–12306.
- Alain, C., and Bernstein, L. J. (2008) From sounds to meaning: the role of attention during auditory scene analysis. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 16, 485–489.
- Alain, C., McDonald, K. L., Kovacevic, N., and McIntosh, A. R. (2009) Spatiotemporal analysis of auditory “what” and “where” working memory. *Cerebral Cortex*, 19, 305–214.
- Alho, K., Rinne, T., Herron, T. J., and Woods, D. L. (2014) Stimulus-dependent activations and attention-related modulations in the auditory cortex: a meta-analysis of fMRI studies. *Hearing Research*, 307, 29–41.

- Altmann, C. F., Bledowski, C., Wibral, M., and Kaiser, J. (2007) Processing of location and pattern changes of natural sounds in the human auditory cortex. *NeuroImage*, 35, 1192–1200.
- Anourova, I., Nikouline, V. V., Ilmoniemi, R. J., Hotta, J., Aronen, H. J., and Carlson, S. (2001) Evidence for dissociation of spatial and nonspatial auditory information processing. *NeuroImage*, 14, 1268–1277.
- Arnott, S. R., Binns, M. A., Grady, C. L., and Alain, C. (2004) Assessing the auditory dual-pathway model in humans. *NeuroImage*, 22, 401–408.
- Arthur, D. L., Lewis, P. S., Medvick, P. A., and Flynn, E. R. (1991) A neuromagnetic study of selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, 78, 348–360.
- Bajo, V. M., Nodal, F. R., Moore, D. R., and King, A. J. (2010) The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience*, 13, 253–260.
- Barker, D., Plack, C. J., and Hall, D. A. (2012) Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. *Cerebral Cortex*, 22, 745–753.
- Barker, D., Plack, C. J., and Hall, D. A. (2013) Representations of pitch and slow modulation in auditory cortex. *Frontiers in Systems Neuroscience*, 7, 62.
- Barrett, D. J. K., and Hall, D. A. (2006) Response preferences for “what” and “where” in human non-primary auditory cortex. *NeuroImage*, 32, 968–977.
- Baumann, S., Petkov, C. I., and Griffiths, T. D. (2013) A unified framework for the organization of the primate auditory cortex. *Frontiers in Systems Neuroscience*, 7, 1–8.
- Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P. E., Giard, M. H., and Bertrand, O. (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *Journal of Neuroscience*, 27, 9252–9261.
- Bizley, J. K., and Cohen, Y. E. (2013) The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, 14, 693–707.

- Bourquin, N. M. P., Murray, M. M., and Clarke, S. (2013) Location-independent and location-linked representations of sound objects. *NeuroImage*, 73, 40–49.
- Bregman, A. S. (1990) *Auditory Scene Analysis*. MIT Press, Cambridge, MA.
- Broadbent, D. (1958) The selective nature of learning. In: *Perception and communication*. Pergamon Press, Elmsford, NY.
- Buckner, R. L., Krienen, F. M., and Yeo, B. T. T. (2013) Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience*, 16, 832–837.
- Budinger, E., and Scheich, H. (2009) Anatomical connections suitable for the direct processing of neuronal information of different modalities via the rodent primary auditory cortex. *Hearing Research*, 258, 16–27.
- Buffalo, E. A., Fries, P., Landman, R., Liang, H., and Desimone, R. (2010) A backward progression of attentional effects in the ventral stream. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 361–365.
- Butler, B. E., and Trainor, L. J. (2012) Sequencing the cortical processing of pitch-evoking stimuli using EEG analysis and source estimation. *Frontiers in Psychology*, 3, 180.
- Bzdok, D., Langner, R., Schilbach, L., Jakobs, O., Roski, C., Caspers, S., Laird, A. R., Fox, P. T., Zilles, K., and Eickhoff, S. B. (2013) Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. *NeuroImage*, 81, 381–392.
- Carlyon, R. P. (2004) How the brain separates sounds. *Trends in Cognitive Sciences*, 8, 465–471.
- Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., and Petersen, S. E. (2014) Intrinsic and task-evoked network architectures of the human brain. *Neuron*, 83, 238–251.
- Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., and Braver, T. S. (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, 16, 1348–1355.

- Corbetta, M., and Shulman, G. L. (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201–215.
- Culling, J. F., Colburn, H. S., and Spurchise, M. (2001) Interaural correlation sensitivity. *The Journal of the Acoustical Society of America*, 110, 1020–1029.
- Da Costa, S., Bourquin, N. M. P., Knebel, J. F., Saenz, M., Van der Zwaag, W., and Clarke, S. (2015) Representation of sound objects within early-stage auditory areas: a repetition effect study using 7T fMRI. *PLoS ONE*, 10, e0124072.
- Da Costa, S., van der Zwaag, W., Marques, J. P., Frackowiak, R. S. J., Clarke, S., and Saenz, M. (2011) Human primary auditory cortex follows the shape of Heschl's gyrus. *Journal of Neuroscience*, 31, 14067–14075.
- Da Costa, S., van der Zwaag, W., Miller, L. M., Clarke, S., and Saenz, M. (2013) Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. *Journal of Neuroscience*, 33, 1858–1863.
- David, S. V. (2018) Incorporating behavioral and sensory context into spectro-temporal models of auditory encoding. *Hearing Research*, 360, 107–123.
- Davis, T., LaRocque, K. F., Mumford, J. A., Norman, K. A., Wagner, A. D., and Poldrack, R. A. (2014). What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *NeuroImage*, 97, 271–283.
- De Santis, L., Clarke, S., and Murray, M. M. (2007) Automatic and intrinsic auditory “what” and “where” processing in humans revealed by electrical neuroimaging. *Cerebral Cortex*, 17, 9–17.
- Desimone, R., and Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- Degerman, A., Rinne, T., Salmi, J., Salonen, O., and Alho, K. (2006) Selective attention to sound location or pitch studied with fMRI. *Brain Research*, 1077, 123–134.

- Degerman, A., Rinne, T., Särkkä, A. K., Salmi, J., and 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.
- Delogu, F., Gravina, M., Nijboer, T., and Postma, A. (2014) Binding "what" and "where" in auditory working memory: An asymmetrical association between sound identity and sound location. *Journal of Cognitive Psychology*, 26, 788–798.
- Deutsch, J. A., and Deutsch, D. (1963) Attention: some theoretical considerations. *Psychological Review*, 70, 80–90.
- Dick, F., Tierney, A. T., Lutti, A., Josephs, O., Sereno, M. I., and Weiskopf, N. (2012) *In vivo* functional and myeloarchitectonic mapping of human primary auditory areas. *Journal of Neuroscience*, 32, 16095–16105.
- Donchin, E., and Coles, M. G. H. (1988) Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11, 357–374.
- Dyson, B. J., and Quinlan, P. T. (2004) Stimulus processing constraints in audition. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 1117–1131.
- Engell, A., Junghöfer, M., Stein, A., Lau, P., Wunderlich, R., Wollbrink, A., and Pantev, C. (2016) Modulatory effects of attention on lateral inhibition in the human auditory cortex. *PLoS ONE*, 11, e0149933.
- Formisano, E., and Kriegeskorte, N. (2012) Seeing patterns through the hemodynamic veil – the future of pattern-information fMRI. *NeuroImage*, 62, 1249–1256.
- Frey, S., Campbell, J. S. W., Pike, G. B., and Petrides, M. (2008) Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *Journal of Neuroscience*, 28, 11435–11444.
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, 6, 1216–1223.

- Fritz, J. B., David, S. V., Radtke-Schuller, S., Yin, P., and Shamma, S. A. (2010) Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nature Neuroscience*, 13, 1011–1019.
- Fullerton, B. C., and Pandya, D. N. (2007) Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *Journal of Comparative Neurology*, 504, 470–498.
- Gaab, N., Gaser, C., and Schlaug, G. (2006) Improvement-related functional plasticity following pitch memory training. *NeuroImage*, 31, 255–263.
- Gallun, F. J., Mason, C. R., and Kidd, G. (2007) Task-dependent costs in processing two simultaneous auditory stimuli. *Perception & Psychophysics*, 69, 757–771.
- Giard, M. H., Perrin, F., Pernier, J., and Peronnet, F. (1988) Several attention-related waveforms in auditory areas: a topographic study. *Electroencephalography and Clinical Neurophysiology*, 69, 371–384.
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., and Van Essen, D. C. (2016) A multi-modal parcellation of human cerebral cortex. *Nature*, 536, 171–178.
- Glasser, M. F., and Van Essen, D. C. (2011) Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *Journal of Neuroscience*, 31, 11597–11616.
- Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P., and Liegeois-Chauvel, C. (2001) Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clinical Neurophysiology*, 112, 1850–1859.
- Goodale, M. A., and Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neurosciences*, 15, 20–25.
- Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., and Rauschecker, J. P. (1997) Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport*, 8, 2511–2516.

- Griffiths, T. D., Buchel, C., Frackowiak, R. S. J., and Patterson, R. D. (1998) Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, 1, 422–427.
- Griffiths, T. D., Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., Patterson, R. D., Brugge, J. F., and Howard, M. A. (2010) Direct recordings of pitch responses from human auditory cortex. *Current Biology*, 20, 1128–1132.
- Griffiths, T. D., and Warren, J. D. (2002) The planum temporale as a computational hub. *Trends in Neurosciences*, 25, 348–353.
- Guimerà, R., and Nunes Amaral, L. A. (2005) Functional cartography of complex metabolic networks. *Nature*, 433, 895–900.
- Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., and Rupp, A. (2004) Temporal dynamics of pitch in human auditory cortex. *NeuroImage*, 22, 755–766.
- Hackett, T. A. (2015) Anatomic organization of the auditory cortex. In: Celesta, G. G., and Hickok, G. (eds.), *Handbook of clinical neurology*. Elsevier, London.
- Hackett, T. A., de la Mothe, L. A., Camalier, C. R., Falchier, A., Lakatos, P., Kajikawa, Y., and Schroeder, C. E. (2014) Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. *Frontiers in Neuroscience*, 8, 72.
- Hackett, T. A., Preuss, T. M., and Kaas, J. H. (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *Journal of Comparative Neurology*, 441, 9–222.
- Hackett, T. A., Stepniewska, I., and Kaas, J. H. (1998) Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *Journal of Comparative Neurology*, 394, 475–495.
- Haftor, E. R., Bonnel, A. M., Gallun, E., and Cohen, E. (1998) A role for memory in divided attention between two independent stimuli. In: Palmer, A. R., Rees, A., Summerfield, A. Q., and Meddis, R. (eds.), *Psychophysical and physiological advances in hearing*. Whurr, London.

- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Summerfield, A. Q., Palmer, A. R., Elliott, M. R., and Bowtell, R. W. (2000) Modulation and task effects in auditory processing measured using fMRI. *Human Brain Mapping*, 10, 107–119.
- Hämäläinen, M. S., and Ilmoniemi, R. J. (1984) Interpreting measured magnetic fields of the brain: estimates of current distributions. Technical Report TKK-F-A559, Helsinki University of Technology, Espoo, Finland.
- Hansen, J. C., and Hillyard, S. A. (1980) Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, 49, 277–290.
- Hari, R., Hämäläinen, M., Kaukoranta, E., Mäkelä, J., Joutsiniemi, S. L., and Tiihonen, J. (1989) Selective listening modifies activity of the human auditory cortex. *Experimental Brain Research*, 74, 463–470.
- Harinen, K., and Rinne, T. (2013) Activations of human auditory cortex to phonemic and nonphonemic vowels during discrimination and memory tasks. *NeuroImage*, 77, 279–287.
- Harinen, K., and Rinne, T. (2014) Acoustical and categorical tasks differently modulate activations of human auditory cortex to vowels. *Brain and Language*, 138, 71–79.
- Herrmann, C. S., and Knight, R. T. (2001) Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews*, 25, 465–476.
- Hickok, G., and Saberi, K. (2012) Redefining the functional organization of the planum temporale region: space, objects, and sensory–motor integration In: Poeppel, D., Overath, T., Popper, A. N., and Fay, R. R. (eds.), *The Human Auditory Cortex*. Springer, NY.
- Higgins, N. C., McLaughlin, S. A., Rinne, T., and Stecker, G. C. (2017) Evidence for cue-independent spatial representation in the human auditory cortex during active listening. *Proceedings of the National Academy of Sciences of the United States of America*, 114, E7602–E7611.
- Hill, K. T., and Miller, L. M. (2010) Auditory attentional control and selection during cocktail party listening. *Cerebral Cortex*, 20, 583–590.

- Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973) Electrical signs of selective attention in the human brain. *Science*, 182, 177–180.
- Hochstein, S., and Ahissar, M. (2002) View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron*, 36, 791–804.
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., and Hagmann, P. (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 2035–2040.
- Huang, S., Seidman, L. J., Rossi, S., and Ahveninen, J. (2013) Distinct cortical networks activated by auditory attention and working memory load. *NeuroImage*, 83, 1098–1108.
- Irvine, D. R. F. (2017) Plasticity in the auditory system. *Hearing Research*, 362, 61–73.
- Jemel, B., Oades, R. D., Oknina, L., Achenbach, C., and Röpcke, B. (2003) Frontal and temporal lobe sources for a marker of controlled auditory attention: the negative difference (Nd) event-related potential. *Brain Topography*, 15, 249–262.
- Johnsrude, I. S., Penhune, V. B., and Zatorre, R. J. (2000) Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, 123, 155–163.
- Joseph, S., Kumar, S., Husain, M., and Griffiths, T. D. (2015) Auditory working memory for objects vs. features. *Frontiers in Neuroscience*, 9, 13.
- Kaas, J. H., and Hackett, T. A. (2000) Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, 97, 11793–11799.
- Kauramäki, J., Jääskeläinen, I. P., Hänninen, J. L., Auranen, T., Nummenmaa, A., Lampinen, J., and Sams, M. (2012) Two-stage processing of sounds explains behavioral performance variations due to changes in stimulus contrast and selective attention: an MEG study. *PLoS ONE*, 7, e46872.

- Kim, S. G., and Ogawa, S. (2012) Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. *Journal of Cerebral Blood Flow and Metabolism*, 32, 1188–1206.
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., and Gruber, O. (2009) Functional architecture of verbal and tonal working memory: an fMRI study. *Human Brain Mapping*, 30, 859–873.
- Kosaki, H., Hashikawa, T., He, J., and Jones, E. G. (1997) Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *Journal of Comparative Neurology*, 386, 304–316.
- Krienen, F. M., Yeo, B. T. T., and Buckner, R. L. (2014) Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369, 20130526.
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., and Lütkenhöner, B. (2003) Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, 13, 765–772.
- Kumar, S., Joseph, S., Gander, P. E., Barascud, N., Halpern, A. R., and Griffiths, T. D. (2016) A brain system for auditory working memory. *Journal of Neuroscience*, 36, 4492–4505.
- Lakatos, P., Pincze, Z., Fu, K. M. G., Javitt, D. C., Karmos, G., and Schroeder, C. E. (2005) Timing of pure tone and noise-evoked responses in macaque auditory cortex. *Neuroreport*, 16, 933–937.
- Lee, C. C., and Middlebrooks, J. C. (2010) Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, 14, 108–114.
- Leech, R., Scott, G., Carhart-Harris, R., Turkheimer, F., Taylor-Robinson, S. D., and Sharp, D. J. (2014) Spatial dependencies between large-scale brain networks. *PLoS ONE*, 9, e98500.
- Leung, A. W. S., and Alain, C. (2010) Working memory load modulates the auditory "What" and "Where" neural networks. *NeuroImage*, 55, 1260–1269.

- Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., and Chauvel, P. (1994) Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology*, 92, 204–214.
- Lin, F. H., Belliveau, J. W., Dale, A. M., and Hämäläinen, M. S. (2006) Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, 27, 1–13.
- Linden, D. E. J. (2005) The P300: where in the brain is it produced and what does it tell us? *Neuroscientist*, 11, 563–576.
- Linke, A. C., Vicente-Grabovetsky, A., and Cusack, R. (2011) Stimulus-specific suppression preserves information in auditory short-term memory. *Proceedings of the National Academy of Sciences*, 108, 12961–12966.
- Logothetis, N. K. (2008) What we can do and what we cannot do with fMRI. *Nature*, 453, 869–878.
- Loveless, N., Vasama, J. P., Mäkelä, J., and Hari, R. (1994) Human auditory cortical mechanisms of sound lateralisation: III. Monaural and binaural shift responses. *Hearing Research*, 81, 91–99.
- Luo, H., Husain, F. T., Horwitz, B., and Poeppel, D. (2005) Discrimination and categorization of speech and non-speech sounds in an MEG delayed-match-to-sample study. *NeuroImage*, 28, 59–71.
- Manly, B. F. J. (1991) *Randomization and Monte Carlo Methods in Biology*. Chapman & Hall, London.
- Massaro, D. W., Cohen, M. M., and Idson, W. L. (1976) Recognition masking of auditory lateralization and pitch judgments. *Journal of the Acoustical Society of America*, 59, 434–441.
- Meilă, M. (2007) Comparing clusterings—an information based distance. *Journal of Multivariate Analysis*, 98, 873–895.
- Mesgarani, N., and Chang, E. F. (2012) Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485, 233–236.

- Mishkin, M., and Ungerleider, L. G. (1982) Contribution of the striate inputs to the visuospatial functions of the parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, 6, 57–77.
- Moerel, M., De Martino, F., and Formisano, E. (2014) An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, 8, 225.
- Mondor, T. A., Zatorre, R. J., and Terrio, N. A. (1998) Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 66–79.
- Murphy, S., Spence, C., and Dalton, P. (2017) Auditory perceptual load: a review. *Hearing Research*, 352, 40–48.
- Näätänen, R. (1990) The role of attention in auditory information-processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13, 201–232.
- Näätänen, R., and Winkler, I. (1999) The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125, 826–859.
- Nahum, M., Daikhin, L., Lubin, Y., Cohen, Y., and Ahissar, M. (2010) From comparison to classification: a cortical tool for boosting perception. *Journal of Neuroscience*, 30, 1128–1136.
- Neelon, M. F., Williams, J., and Garell, P. C. (2011) Elastic attention: enhanced, then sharpened response to auditory input as attentional load increases. *Frontiers in Human Neuroscience*, 5, 41.
- Nelken, I., and Ahissar, M. (2006) High-level and low-level processing in the auditory system: the role of primary auditory cortex. In: Divenyi, P., Greenberg, S., and Meyer, G. (eds.), *Dynamics of Speech Production and Perception*. IOS Press, Amsterdam.
- Neubert, F. X., Mars, R. B., Thomas, A. G., Sallet, J., and Rushworth, M. F. S. (2014) Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron*, 81, 700–713.

- Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005) Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, 131, 510–532.
- Overath, T., Cusack, R., Kumar, S., Von Kriegstein, K., Warren, J. D., Grube, M., Carlyon, R. P., and Griffiths, T. D. (2007) An information theoretic characterisation of auditory encoding. *PLoS Biology*, 5, 2723–2732.
- Parisot, S., Glocker, B., Ktena, S. I., Arslan, S., Schirmer, M. D., and Rueckert, D. (2017) A flexible graphical model for multi-modal parcellation of the cortex. *NeuroImage*, 162, 226–248.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (2002) The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36, 767–776.
- Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., and Woods, D. L. (2004) Attentional modulation of human auditory cortex. *Nature Neuroscience*, 7, 658–663.
- Picton, T. W. (1992) The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, 9, 456–479.
- Polich, J. (2007) Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148.
- Puschmann, S., Uppenkamp, S., Kollmeier, B., and Thiel, C. M. (2010) Dichotic pitch activates pitch processing centre in Heschl's gyrus. *NeuroImage*, 49, 1641–1649.
- Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H. J., and Zilles, K. (2001) Probabilistic mapping and volume measurement of human primary auditory cortex. *NeuroImage*, 13, 669–683.
- Rauschecker, J. P., and Tian, B. (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 11800–11806.
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, 268, 111–114.

- Recanzone, G. H., and Cohen, Y. E. (2010) Serial and parallel processing in the primate auditory cortex revisited. *Behavioural Brain Research*, 206, 1–7.
- Recanzone, G. H., Guard, D. C., Phan, M. L., and Su, T. I. K. (2000) Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *Journal of Neurophysiology*, 83, 2723–2739.
- Riecke, L., Peters, J. C., Valente, G., Kemper, V. G., Formisano, E., and Sorger, B. (2017) Frequency-selective attention in auditory scenes recruits frequency representations throughout human superior temporal cortex. *Cerebral Cortex*, 27, 3002–3014.
- Rif, J., Hari, R., Hämäläinen, M. S., and Sams, M. (1991) Auditory attention affects two different areas in the human supratemporal cortex. *Electroencephalography and Clinical Neurophysiology*, 79, 464–472.
- Rinne, T., Balk, M. H., Koistinen, S., Autti, T., Alho, K., and Sams, M. (2008) Auditory selective attention modulates activation of human inferior colliculus. *Journal of Neurophysiology*, 100, 3323–3327.
- Rinne, T., Koistinen, S., Salonen, O., and Alho, K. (2009) Task-dependent activations of human auditory cortex during pitch discrimination and pitch memory tasks. *Journal of Neuroscience*, 29, 13338–13343.
- Rinne, T., Koistinen, S., Talja, S., Wikman, P., and Salonen, O. (2012) Task-dependent activations of human auditory cortex during spatial discrimination and spatial memory tasks. *NeuroImage*, 59, 4126–4131.
- Rinne, T., Muers, R. S., Salo, E., Slater, H., and Petkov, C. I. (2017) Functional imaging of audio-visual selective attention in monkeys and humans: how do lapses in monkey performance affect cross-species correspondences? *Cerebral Cortex*, 27, 3471–3484.
- Rivier, F., and Clarke, S. (1997) Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *NeuroImage*, 6, 288–304.

- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*, 2, 1131–1136.
- Ross, B., Hillyard, S. A., and Picton, T. W. (2010) Temporal dynamics of selective attention during dichotic listening. *Cerebral Cortex*, 20, 1360–1371.
- Rosvall, M., and Bergstrom, C. T. (2008) Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105, 1118–1123.
- Ruschel, M., Knösche, T. R., Friederici, A. D., Turner, R., Geyer, S., and Anwander, A. (2014) Connectivity architecture and subdivision of the human inferior parietal cortex revealed by diffusion MRI. *Cerebral Cortex*, 24, 2436–2448.
- Saenz, M., and Langers, D. R. M. (2014) Tonotopic mapping of human auditory cortex. *Hearing Research*, 307, 42–52.
- Schadwinkel, S., and Gutschalk, A. (2010) Activity associated with stream segregation in human auditory cortex is similar for spatial and pitch cues. *Cerebral Cortex*, 20, 2863–2873.
- Scheich, H., Brechmann, A., Brosch, M., Budinger, E., and Ohl, F. W. (2007) The cognitive auditory cortex: task-specificity of stimulus representations. *Hearing Research*, 229, 213–224.
- Semal, C., and Demany, L. (2006) Individual differences in the sensitivity to pitch direction. *Journal of the Acoustical Society of America*, 120, 3907–3915.
- Shinn-Cunningham, B. G. (2008) Object-based auditory and visual attention. *Trends in Cognitive Sciences*, 12, 182–186.
- Singh, K. D. (2012) Which “neural activity” do you mean? fMRI, MEG, oscillations and neurotransmitters. *NeuroImage*, 62, 1121–1130.
- Slee, S. J., and David, S. V. (2015) Rapid task-related plasticity of spectrotemporal receptive fields in the auditory midbrain. *Journal of Neuroscience*, 35, 13090–13102.

- Smith, K. R., Hsieh, I. H., Saberi, K., and Hickok, G. (2010) Auditory spatial and object processing in the human planum temporale: no evidence for selectivity. *Journal of Cognitive Neuroscience*, 22, 632–639.
- Stewart, L., Overath, T., Warren, J. D., Foxton, J. M., and Griffiths, T. D. (2008) fMRI evidence for a cortical hierarchy of pitch pattern processing. *PLoS ONE*, 3, e1470.
- Suga, N., and Ma, X. F. (2003) Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Reviews Neuroscience*, 4, 783–794.
- Suzuki, M., Kitano, H., Kitanishi, T., Itou, R., Shiino, A., Nishida, Y., Yazawa, Y., Ogawa, F., and Kitajima, K. (2002) Cortical and subcortical activation with monaural monosyllabic stimulation by functional MRI. *Hearing Research*, 163, 37–45.
- Sweet, R. A., Dorph-Petersen, K. A., and Lewis, D. A. (2005) Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *Journal of Comparative Neurology*, 491, 270–289.
- Tian, B., Reser, D., Durham, A., Kustov, A., and Rauschecker, J. P. (2001) Functional specialization in rhesus monkey auditory cortex. *Science*, 292, 290–293.
- Tiitinen, H., Salminen, N. H., Palomäki, K. J., Mäkinen, V. T., Alku, P., and May, P. J. C. (2006) Neuromagnetic recordings reveal the temporal dynamics of auditory spatial processing in the human cortex. *Neuroscience Letters*, 396, 17–22.
- van der Zwaag, W., Gentile, G., Gruetter, R., Spierer, L., and Clarke, S. (2011) Where sound position influences sound object representations: a 7-T fMRI study. *NeuroImage*, 54, 1803–1811.
- Walker, K. M. M., and King, A. J. (2011) Auditory neuroscience: temporal anticipation enhances cortical processing. *Current Biology*, 21, R251–R253.
- Wallace, M. N., Cronin, M. J., Bowtell, R. W., Scott, I. S., Palmer, A. R., and Gowland, P. A. (2016) Histological basis of laminar MRI patterns in high resolution images of fixed human auditory cortex. *Frontiers in Neuroscience*, 10, 455.

- Warren, J. D., and Griffiths, T. D. (2003) Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience*, 23, 5799–5804.
- Weeks, R. A., Aziz-Sultan, A., Bushara, K. O., Tian, B., Wessinger, C. M., Dang, N., Rauschecker, J. P., and Hallett, M. (1999) A PET study of human auditory spatial processing. *Neuroscience Letters*, 262, 155–158.
- Weinberger, N. M. (2011) Reconceptualizing the primary auditory cortex: learning, memory and specific plasticity. In: Winer, J. A., and Schreiner, C. E. (eds.), *The Auditory Cortex*. Springer, NY.
- Wightman, F. L., Kistler, D. J., and Perkins, M. E. (1987) A new approach to the study of human sound localization. In: Yost, W. A., and Gourevitch, G. (eds.), *Directional Hearing*. Springer, NY.
- Winer, J. A. (2005) Decoding the auditory corticofugal systems. *Hearing Research*, 207, 1–9.
- Winer, J. A., and Lee, C. C. (2007) The distributed auditory cortex. *Hearing Research*, 229, 3–13.
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., and Nichols, T. E. (2014) Permutation inference for the general linear model. *NeuroImage*, 92, 381–397.
- Winkler, I., Denham, S. L., and Nelken, I. (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, 13, 532–540.
- Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., and Bloom, F. E. (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 8722–8726.
- Woldorff, M. G., and Hillyard, S. A. (1991) Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, 79, 170–191.
- Woods, D. L., and Alain, C. (2001) Conjoining three auditory features: an event-related brain potential study. *Journal of Cognitive Neuroscience*, 13, 492–509.

- Woods, D. L., Alho, K., and Algazi, A. (1992) Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography and Clinical Neurophysiology*, 82, 341–355.
- Woods, D. L., Alho, K., and Algazi, A. (1994) Stages of auditory feature conjunction: an event-related brain potential study. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 80–94.
- Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., Liao, I., and Kang, X. (2009) Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS ONE*, 4, e5183.
- Yost, W. A. (1996) Pitch of iterated rippled noise. *Journal of the Acoustical Society of America*, 100, 511–518.
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010) Network-based statistic: identifying differences in brain networks. *NeuroImage*, 53, 1197–1207.
- Zatorre, R. J., Bouffard, M., Ahad, P., and Belin, P. (2002) Where is 'where' in the human auditory cortex? *Nature Neuroscience*, 5, 905–909.

ORIGINAL PUBLICATIONS