

Department of Psychology and Logopedics  
Faculty of Medicine  
University of Helsinki

# **THE EFFECT OF LISTENING TASKS AND MOTOR RESPONDING ON ACTIVATION IN THE AUDITORY CORTEX**

**Patrik Wikman**

Doctoral Program of Brain and Mind

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine at  
the University of Helsinki, for public examination in lecture room 2,  
Biomedicum, on 5 June 2019, at 12 o'clock.

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

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

**Reviewers** Dr. Juha Salmitaival  
Department of Psychology  
University of Turku, Finland

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

**Opponent** Professor Erich Schröger  
Institute of Psychology  
University of Leipzig  
Leipzig, Germany

ISBN 978-951-51-5225-1 (pbk.)  
ISBN 978-951-51-5226-8 (PDF)

Unigrafia  
Helsinki 2019

# CONTENTS

Abstract .....	5
Abstrakt.....	7
Acknowledgments.....	9
List of original publications .....	10
Abbreviations .....	11
1 Introduction .....	12
1.1 The auditory cortex .....	12
1.2 Attention- and task-related activation in the auditory cortex .....	14
1.3 Motor effects in the auditory cortex .....	16
1.3.1 Auditory-motor integration .....	16
1.3.2 Suppression during motor execution .....	17
1.3.3 The effect of manual grip types.....	19
1.3.4 Relationship between task and motor effects.....	19
1.4 Reward incentive cues as a means to facilitate behavioral training of monkeys .....	20
2 Aims of the present thesis .....	23
3 Methods and results.....	25
3.1 General methods in studies I and II .....	25
3.1.1 Human subjects .....	25
3.1.2 Procedures.....	25
3.1.3 Stimuli .....	26
3.2 Methods in Study III .....	27
3.2.1 Stimuli and tasks.....	27
3.2.2 Procedures.....	29
3.3 Fmri data acquisition and analysis .....	30

3.4	Study I. The effect of precision and power grips on activation in human auditory cortex .....	32
3.4.1	Tasks.....	32
3.4.2	Responses.....	32
3.4.3	Results .....	33
3.5	Study II. Interaction of the effects associated with auditory-motor integration and attention-engaging listening tasks .....	35
3.5.1	Tasks.....	35
3.5.2	Responses.....	35
3.5.3	Results .....	36
3.6	Study III. Reward cues readily direct monkeys' auditory performance resulting in broad auditory cortex modulation and interaction with sites along cholinergic and dopaminergic pathways.....	39
3.6.1	Results .....	39
4	General discussion .....	44
4.1	Active listening strongly modulates activation in broad regions of human and monkey auditory cortex .....	44
4.2	Widespread regions of the auditory cortex are suppressed during motor responding.....	47
4.3	Auditory-Motor integration modulates activation in the auditory cortex .....	49
4.4	Does the hierarchical state feedback model explain the effects of motor responding in the auditory cortex?.....	50
4.5	Are task-related and motor-response-related activation modulations independent of each other?.....	52
4.6	Implications for theoretical models of the human auditory cortex.....	54
5	Conclusions .....	56
	References.....	57
	Original publications .....	71

# ABSTRACT

Previous human functional magnetic resonance imaging (fMRI) research has shown that activation in the auditory cortex (AC) is strongly modulated by motor influences. Other fMRI studies have indicated that the AC is also modulated by attention-engaging listening tasks. How these motor- and task-related activation modulations relate to each other has, however, not been previously studied. The current understanding of the functional organization of the human AC is strongly based on primate models. However, some authors have recently questioned the correspondence between the monkey and human cognitive systems, and whether the monkey AC can be used as a model for the human AC. Further, it is unknown whether active listening modulates activations similarly in the human and nonhuman primate AC. Thus, non-human primate fMRI studies are important. Yet, such fMRI studies have been previously impeded by the difficulty in teaching tasks to non-human primates. The present thesis consists of three studies in which fMRI was used both to investigate the relationship between the effects related to active listening and motor responding in the human AC and to investigate task-related activation modulations in the monkey AC. Study I investigated the effect of manual responding on activation in the human AC during auditory and visual tasks, whereas Study II focused on the question whether auditory-motor effects interact with those related to active listening tasks in the AC and adjacent regions. In Study III, a novel paradigm was developed and used during fMRI to investigate auditory task-dependent modulations in the monkey AC.

The results of Study I showed that activation in the AC in humans is strongly suppressed when subjects respond to targets using precision or power grips during both visual and auditory tasks. AC activation was also modulated by grip type during the auditory task but not during the visual task (with identical stimuli and motor responses). These manual-motor effects were distinct from general attention-related modulations revealed by comparing activation during auditory and visual tasks. Study II showed that activation in widespread regions in the AC and inferior parietal lobule (IPL) depends on whether subjects respond to target vowel pairs using vocal or manual responses. Furthermore, activation in the posterior AC and the IPL depends on whether subjects respond by overtly repeating the last vowel of a target pair or by producing a given response vowel. Discrimination tasks activated superior temporal gyrus (STG) regions more strongly than 2-back tasks, while the IPL was activated more strongly by 2-back tasks. These task-related (discrimination vs. 2-back) modulations were distinct from the response type effects in the AC. However, task and motor-response-type effects interacted in the IPL. Together the results of Studies I and II support the view that operations in the AC are shaped by its connections with motor

cortical regions and that regions in the posterior AC are important in auditory-motor integration. Furthermore, these studies also suggest that the task, motor-response-type and vocal-response-type effects are caused by independent mechanisms in the AC.

In Study III, a novel reward-cue paradigm was developed to teach macaque monkeys to perform an auditory task. Using this paradigm monkeys learned to perform an auditory task in a few weeks, whereas in previous studies auditory task training has required months or years of training. This new paradigm was then used during fMRI to measure activation in the monkey AC during active auditory task performance. The results showed that activation in the monkey AC is modulated during this task in a similar way as previously seen in human auditory attention studies. The findings of Study III provide an important step in bridging the gap between human and animal studies of the AC.

# ABSTRAKT

Tidigare forskning med funktionell magnetresonanstomografi (fMRI) har visat att aktiveringen i hörselhjärnbarken hos människor är starkt påverkad av motoriken. Andra fMRI-studier visar att aktiveringen i hörselhjärnbarken också påverkas av uppgifter som kräver aktivt lyssnande. Man vet ändå inte hur dessa motoriska och uppgiftsrelaterade effekter hänger ihop. Den nuvarande uppfattningen om hörselhjärnbarkens funktionella struktur hos människan är starkt påverkad av primatmodeller. Däremot har en del forskare nyligen ifrågasatt om apors kognitiva system motsvarar människans, och specifikt huruvida apans hörselhjärnbark kan användas som modell för människans. Dessutom vet man inte om aktivt lyssnande påverkar aktivering i hörselhjärnbarken hos apor på samma sätt som hos människor. Därför är fMRI-studier på apor viktiga. Sådana fMRI-studier har emellertid tidigare hindrats av svårigheten att lära apor att göra uppgifter. Denna doktorsavhandling utgörs av tre studier där man använde fMRI för att undersöka hur effekter som är relaterade till aktivt lyssnande och motorik förhåller sig till varandra i hörselhjärnbarken hos människan och hur aktiva uppgifter påverkar aktiveringar i hörselhjärnbarken hos apor. I Studie I undersöktes hur aktiveringen i hörselhjärnbarken hos människan påverkades medan försökspersonerna utförde auditiva och visuella uppgifter och gav sina svar manuellt. Studie II fokuserade på huruvida audiomotoriska effekter och effekter relaterade till aktiva hörseluppgifter samspelade i hörselhjärnbarken och dess omnejd. I Studie III utvecklades ett nytt försöksparadigm som sedermera användes för att undersöka auditiva uppgiftsrelaterade aktiveringar i hörselhjärnbarken hos apor.

Resultaten av Studie I visade att aktiveringen i hörselhjärnbarken dämpas starkt när försökspersonerna reagerar på målstimulus med precisions- och styrkegrepp både vid auditiva och visuella uppgifter. Aktivering i hörselhjärnbarken påverkas också av typen av grepp då försökspersonerna utförde auditiva uppgifter men inte då de utförde visuella uppgifter (med identiska stimuli och motoriska reaktioner). Dessa manuellt-motoriska effekter kunde särskiljas från allmänna uppmärksamhetsrelaterade effekter, vilka kom fram då man jämförde aktiveringen under auditiva och visuella uppgifter. Typen av motoriska reaktioner, dvs. hur försökspersonerna reagerade på målstimuli (genom att reagera med händerna eller att uttala ljud) påverkade aktiveringen i stora områden i hörselhjärnbarken och lobulus parietale inferior (IPL) i Studie II. Aktiveringen i den bakre delen av hörselhjärnbarken och IPL påverkades också av om försökspersonen upprepade målstimulusens sista vokal eller svarade genom att uttala en given responsvokal. Diskriminationsuppgifter aktiverade gyrus temporale superior mera än 2-back (minnes) -uppgifter, medan IPL aktiverades mera av 2-back -uppgifterna. Dessa

uppgiftsrelaterade (diskrimination vs. 2-back) påverkningar var oberoende av effekter som hade att göra med reaktionstypen i hörselhjärnbarken. Däremot fanns det ett samspel mellan uppgift och motoriska effekter i IPL. Tillsammans stärker resultaten från Studie I och II uppfattningen att funktioner inom hörselhjärnbarken är starkt beroende av dess sammankoppling med den motoriska hjärnbarken, och att bakre delarna av hörselhjärnbarken är viktiga för audiomotorisk integration. Dessa studier visar därtill att uppgiftsrelaterade, motoriska och uttalsrelaterade effekter produceras av oberoende mekanismer i hörselhjärnbarken.

I Studie III utvecklades ett nytt försöksparadigm som var baserat på belönings signaler. Med detta försöksparadigm lärdes makakapor att utföra en auditiv uppgift. I Studie III lärde sig makakaporna uppgiften inom ett par veckor, medan inläringen av auditiva uppgifter i tidigare studier har tagit upp till flera år. Detta paradig användes sedan med hjälp av fMRI för att mäta aktivering inom hörselhjärnbarken hos apor, medan aporna utförde aktiva auditiva uppgifter. Resultaten visar att aktiveringen i hörselhjärnbarken hos apor påverkas av uppgifter på liknande sätt som man tidigare har visat i människoforskning. Fynden i Studie II är ett viktigt framsteg för att kunna överbygga gapet mellan människostudier och djurstudier gällande hörselhjärnbarken.



# ACKNOWLEDGMENTS

This work was conducted at the Department of Psychology and Logopedics, University of Helsinki. The human fMRI measurements were carried out at the Advanced Magnetic Imaging Centre (AMI), Aalto University School of Science and the monkey fMRI measurements were conducted at the Comparative Biology Centre (CBC), Institute of Neuroscience in Newcastle upon Tyne (UK). My work was funded by the Academy of Finland, the Finnish Cultural foundation, the Alfred Kordelin Foundation, and the Doctoral Program of Brain and Mind.

First, I want to thank my supervisor Dr. Teemu Rinne for giving me the outstanding opportunity to work with such inspiring topics. Due to the strenuous efforts of my supervisor, I believe that now I have learned to appreciate the importance of crystal-clear argumentation. I also want to thank my supervisor professor Chris Stecker for his insightful comments on this thesis.

I want to thank the official reviewers Dr. Juha Salmitaival and Dr. Iiro Jääskeläinen for their thoughtful comments on this thesis and professor Erich Schröger for kindly agreeing to serve as my opponent. I thank Lari Vainio, who contributed to Study I. I especially want to thank professor Christopher Petkov for allowing me to work in his research group and his theoretical insights in Study III.

I also want to thank the excellent staff at AMI and CBC for providing the research facilities and the support needed to conduct these studies. A big thank you to the radiographer Marita Kattelus at AMI. I wish I had your warm calmness even when everything goes amok. Thank you, Dr. Heather Slater and Dr. Ross Muers for teaching me how to work in the lab in Newcastle. A special thank you to the animal technicians Ash Waddle, Steve O'Keefe and Carrie Todd, you always lighted up my day, making days in the lab with the monkeys a joy.

I want to thank all my coworkers at the Department of Psychology and Logopedics. A special thank you to my peer doctoral students. Thank you for sharing the peaks and valleys that go with starting an academic career, and especially thank you for laughing with (or at) me even when I might not have been the most cheerful of beings. A sincere thank you also to friends outside academia for support over the years, and Ilkka Järvinen for proof reading this thesis.

Finally, I want to thank my family and especially my mother for enduring late-night calls and helping me with everything ranging from grant applications to reviewer comments and spotting my frequent spelling mistakes.

Patrik Wikman  
Helsinki, January 2019

# LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

- Study I      Wikman P., Vainio L., and Rinne T. (2015). The effect of precision and power grips on activations in human auditory cortex. *Frontiers in Neuroscience*, 9.
- Study II     Wikman P. and Rinne T. (2019). Interaction of the effects associated with auditory-motor integration and attention-engaging listening tasks. *Neuropsychologia*, 124, 322-336.
- Study III    Wikman P., Rinne T. and Petkov C. I. (2019). Reward cues readily direct monkeys' auditory performance resulting in broad auditory cortex modulation and interaction with sites along cholinergic and dopaminergic pathways. *Scientific Reports*, 3055.

# ABBREVIATIONS

AC	Auditory cortex
ANOVA	Analysis of variance
a/p/PT	Anterior/posterior/planum temporale
$A_{pr, po, no}$	Auditory precision, power, no-response
BOLD	Blood-oxygen-level-dependent
Bu	Button
C/CW	Counter/clockwise
Discr	Discrimination
EPI	Echo-planar imaging
ER	Early-response rate
Fo	Fundamental frequency
fMRI	Functional magnetic resonance imaging
FSL	FMRIB Software Library
FWER	Family-wise error rate
HG	Heschl's gyrus
HiRe	High reward
HR	Hit rate
HSF	Hierarchical state feedback
IFG	Inferior frontal gyrus
IPL	Inferior parietal lobule
IRN	Iterated rippled noise
LoRe	Low reward
MC	Motor cortex
M/EEG	Magneto/Electroencephalography
MR	Miss rate
NPh	Non-phonemic
PALM	Permutation analysis of linear models
Ph	Phonemic-response blocks
piPh	Pitch-modulated vowel
Pr	Production
Re	Repetition
ROI	Region of interest
RT	Reaction time
SMG	Supramarginal gyrus
Spt	Sylvian-parietal-temporal
STG	Superior temporal gyrus
TE	Echo time
TR	Time of relaxation
$V_{pr, po, no}$	Visual precision, power, no-response

# 1 INTRODUCTION

Current models of the functional organization of the auditory cortex (AC) are largely based on (invasive) neuronal-level studies conducted in non-human primates during passive conditions (Rauschecker et al., 1995; Recanzone and Cohen, 2010; Romanski et al., 1999). However, non-invasive brain imaging studies in humans have shown that activation in wide regions of the human AC is strongly modulated during active listening tasks (Alho et al., 2014; De Martino et al., 2015; Hall et al., 2000; Petkov et al., 2004; Riecke et al., 2018; Rinne, 2010; Rinne et al., 2005; Woods et al., 2009). These activation modulations during active listening cannot be predicted by the current models of the AC. Monkey studies using active auditory tasks during fMRI could provide the missing link between neurophysiological measurements in monkeys and human fMRI studies. However, systematic use of active conditions in animal studies has been impeded by the difficulty of training behavioral auditory tasks in non-human primates.

Studies in both humans and animals have also shown that input from the motor cortex strongly modulates activation and operations in the AC (Baumann et al., 2007; Buchsbaum et al., 2001; Chen et al., 2006; Hickok et al., 2003; Schneider et al., 2014; Schneider et al., 2018; Wise et al., 2001). Although task- and motor-related modulations are seen in overlapping regions in the AC, these effects have been investigated in separate studies and thus their relationship is unclear.

The present thesis used functional magnetic resonance imaging (fMRI) to systematically investigate the effects of active auditory tasks and motor responding in the human AC and adjacent regions. Also, a novel paradigm was developed to speed up auditory task training in monkeys, and this paradigm was used during fMRI to measure the effect of active listening tasks on the activation in the monkey AC. The work in the present thesis is important because, to further develop comprehensive models of the human AC, it is important to understand the correspondence between task effects in non-human animals and humans, and how task and motor effects are related in the AC.

## 1.1 THE AUDITORY CORTEX

Models of the human AC are strongly influenced by neurophysiological and anatomical studies on the functional organization of the non-human primate AC (Rauschecker and Romanski, 2011; Rauschecker and Scott, 2009). This work suggests that the organization of the monkey AC is based on three main principles: (1) the auditory cortex can be subdivided into functionally independent sub-regions, (2) these regions are connected to each other in a hierarchical fashion and (3) regions of the AC are connected to the motor

cortex through parallel processing streams. The monkey AC has been suggested to consist of primary core regions hierarchically connected to surrounding secondary belt regions, which in turn are connected to parabelt regions (Hackett et al., 2001; Kaas and Hackett, 2000; Rauschecker et al., 1995; Recanzone and Cohen, 2010; Romanski et al., 1999). Further, the anterior AC is connected via a ventral stream to the frontal and motor cortices, while the posterior AC is connected to these regions through a separate dorsal processing stream (Kaas and Hackett, 2000; Rauschecker and Tian, 2000; Tian et al., 2001).

The functional organization of the human AC is not well understood, but is generally believed to follow the same organizational principles as the monkey AC. Post-mortem anatomical studies (Rivier and Clarke, 1997) and fMRI studies (Berlot et al., 2018; Moerel et al., 2014; Wessinger et al., 2001; Woods and Alain, 2009; Woods et al., 2009; Woods et al., 2010) in humans support the notion that the core-belt-parabelt organization is present also in the human AC. According to the current understanding, regions in and near the human Heschl's gyrus (HG) support primary-like functions (core), while the planum temporale (PT) posterior to the HG and the superior temporal gyrus (STG) lateral to the HG have belt- or parabelt-like properties. Further, the human AC has also been suggested to be connected to the frontal and motor cortices in a similar fashion as in the non-human primate AC, with anterior parts of the AC projecting through a ventral stream to the frontal cortex, and the posterior part of the AC projecting through a dorsal stream via the inferior parietal lobule (IPL) to the frontal cortices (Rauschecker and Scott, 2009).

It is likely that the evolution of human speech and speech-related functions have shaped the functional organization of the human AC and its connections with the motor cortices (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Further, models of the human AC stress the role of strong connections between motor cortical regions and the AC (Formisano et al., 2015; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009; Schneider and Mooney, 2018). On the other hand, another line of studies has shown that activation in wide regions of the human AC is strongly modulated during active listening tasks (Alho et al., 2014; Hall et al., 2000; Petkov et al., 2004; Rinne, 2010; Rinne et al., 2005; Woods et al., 2009). The exact neural mechanisms underlying such task-related effects are currently poorly understood. Task-related modulations have also not been discussed in the context of the aforementioned models of the AC that focus on auditory-motor effects. This is partly due to the fact that task- and motor-related effects have not previously been investigated in the same study. Therefore, it is currently not known whether task influences are caused by the same processing streams as thought to underlie motor influences in the AC or by some other independent mechanism.

Could studies in non-human primates help to integrate attention- and task-related influences in theoretical models of the human AC? Currently,

non-human primate studies often probe stimulus-driven effects while the animal is passively listening to the sounds or under general anesthesia (Rauschecker et al., 1995; Recanzone and Cohen, 2010; Romanski et al., 1999). When active tasks are adopted in non-human animals they mostly focus on a select few neurons and the primary auditory cortex (Atiani et al., 2009; Atiani et al., 2014; Bagur et al., 2018; Briggs et al., 2013; Francis et al., 2018; Fritz et al., 2003; Fritz et al., 2005a; Fritz et al., 2007a; Fritz et al., 2007b; Kölsch et al., 2009; Reynolds and Heeger, 2009). Therefore, it is currently unknown whether attention modulates activity in the non-human AC to the same extent as in humans.

The lack of studies on attention- and task-related factors in non-human primates is at least partly due to the fact that auditory tasks, readily taught to human subjects, are notoriously laborious to train non-human primates on. Furthermore, when monkeys are taught tasks, the neural and behavioral effects do not always correspond to those observed in human subjects. This has led some authors to question whether the use of monkeys as a model to investigate human cortical functions is valid (Patel et al., 2015; Schulze et al., 2012; Scott et al., 2012). Other authors, however, suggest that also in non-human animals, task modulations of auditory processing is an important aspect of the function of the AC (Scheich et al., 2007) and that the overall function of the AC is to solve higher level auditory problems (Weinberger, 2011). Therefore, it would be important to conduct comparative studies using similar measures and auditory tasks in non-human primates as in humans to incorporate more cognitive aspects into models of the human AC. Such studies could help to bridge the gap between the vast neurophysiological literature in monkeys and the human fMRI literature, and to refine our understanding of the functional organization of the human AC.

## **1.2 ATTENTION- AND TASK-RELATED ACTIVATION IN THE AUDITORY CORTEX**

Human fMRI studies have shown that auditory attention-engaging tasks have a profound influence on activation in the AC. For example, in the study by Petkov and colleagues (Petkov et al., 2004), subjects were presented with auditory stimuli varying in pitch during an auditory discrimination task or during a visual task (i.e., no directed auditory attention). The authors observed that stimulus-dependent activation to sounds during the visual task (vs. visual task without sounds) were centered on the HG in the superior temporal cortex. Attention to sounds (auditory task vs. visual task with the same sounds) enhanced activation in these regions. However, attention to sounds was also associated with broad activation in STG regions that were not activated by the presentation of the sounds during the visual task. Similar attention-related activation modulations have been observed in a number of

other human fMRI studies (e.g., Alho et al., 2014; De Martino et al., 2015; Hall et al., 2000; Jäncke et al., 1999; Loose et al., 2003; Riecke et al., 2018; Rinne, 2010; Rinne et al., 2005; Salmi et al., 2009; Santangelo et al., 2010; Woods and Alain, 2009; Woods et al., 2009).

A few previous studies have also compared AC activation to different auditory tasks performed on the same auditory stimuli. For example, Rinne and colleagues (Rinne et al., 2009) compared activation in the AC to similar sounds during pitch discrimination, pitch n-back memory tasks and visual tasks. Their subjects were presented with pitch-varying tone pairs that were organized in three separate pitch categories (low, mid or high). In the discrimination task, subjects focused on within-pair pitch differences and indicated when the parts of a tone pair were identical in pitch. In the n-back memory task, subjects were required to indicate whether the pitch category of a sound pair matched the pitch category of the sound pair presented 1–3 trials before (depending on the n-back task difficulty level). Consistent with previous studies showing attention-related modulations in the AC (see above), comparisons between the auditory and visual tasks revealed enhanced activation during auditory tasks in wide regions of the AC. However, comparisons between the two auditory tasks revealed task-dependent activation differences in the AC. Activation in anterior–middle STG regions was higher during the discrimination task than during the n-back tasks. Activation in the IPL, in turn, was higher during the n-back tasks than discrimination tasks. The authors suggested that the enhanced STG activation during discrimination tasks was related to the pitch discrimination tasks demanding detailed sound processing, while the enhanced IPL activation during the n-back tasks was related to the fact that the n-back tasks required working memory and categorical processing. It is important to note that these task-related activation differences cannot be simply explained by enhanced stimulus level processing, as both tasks were performed on identical stimuli. Further, more recent studies have shown that similar activation differences between discrimination and n-back tasks are seen irrespective of whether these tasks are performed on pitch-varying sounds (Häkkinen and Rinne, 2018; Häkkinen et al., 2015; Rinne et al., 2009; Talja et al., 2015), spatially varying sounds (Häkkinen and Rinne, 2018; Rinne et al., 2012; Rinne et al., 2014; Talja et al., 2015) or vowels (Harinen and Rinne, 2013; Harinen and Rinne, 2014).

Together these studies show that (1) attention-engaging auditory tasks modulate activation in wide regions of the human AC, (2) these modulations depend on the characteristics of the listening task, and (3) activation patterns during discrimination and n-back tasks cannot be explained by enhanced stimulus-level processing as similar task-dependent modulations are seen irrespective of the stimulus type.

Neuronal-level studies in animals have shown that attention-engaging tasks increase the sharpness of neuronal responses, or temporarily change the receptive fields of auditory neurons in both the primary and secondary

AC (Atiani et al., 2009; Atiani et al., 2014; Bagur et al., 2018; Briggs et al., 2013; Francis et al., 2018; Fritz et al., 2003; Fritz et al., 2005a; Fritz et al., 2007a; Fritz et al., 2007b; Kölsch et al., 2009; Reynolds and Heeger, 2009). How these neuronal-level effects relate to the attention-related modulations observed in wide AC regions in human fMRI studies is, however, unclear. This is partly because the exact relationship between fMRI measures and neuronal-level measures is currently unknown (Logothetis, 2008). In fMRI, the blood-oxygen-level-dependent (BOLD) MRI signal is measured, which is an indirect measure of neuronal activity. In neurophysiological studies, the electrical activity in single or multiple neurons is measured directly. These techniques also differ in their spatial resolution: neuronal studies can have the resolution of one specific neuron, while the smallest unit in fMRI (i.e., the voxel) contains thousands of neurons. On the other hand, fMRI can be used to measure activity in the whole brain while neurophysiological studies need to focus on a select neuronal site.

Monkey fMRI studies could provide the missing link between neurophysiological measurements in monkeys and human fMRI studies. However, auditory attention studies in actively behaving monkeys are rare. This is at least partly due to challenges in training non-human primates to perform auditory attention tasks. While humans easily learn auditory attention tasks during one training session, training monkeys to perform such tasks often requires hundreds of training sessions over weeks or months (Fritz et al., 2005b; Rinne et al., 2017). Furthermore, even after extensive training, monkeys have frequent lapses in auditory attention which affect both neuronal responses (Lakatos et al., 2016) and activation in the AC (Rinne et al., 2017). Thus, it is of paramount importance to develop such auditory tasks that can be quickly and easily taught to non-human primates.

## **1.3 MOTOR EFFECTS IN THE AUDITORY CORTEX**

In addition to task effects (see 1.2), activation in the AC is also strongly modulated by motor responding and effects related to auditory-motor integration (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). However, the relationship between task- and motor-related modulations in the AC is currently unknown.

### **1.3.1 AUDITORY-MOTOR INTEGRATION**

Speech production requires integration of auditory and motor information. The posterior parts of the AC have been highlighted as an important hub for such functions (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Most previous fMRI studies on auditory-motor integration have focused on the role of the PT during speech. Early studies found that the PT is activated both during listening to speech and covert speech production. For example,



in the study by Buchsbaum and colleagues (Buchsbaum et al., 2001), subjects listened to and covertly repeated speech sounds. The results revealed enhanced activation during both listening and covert rehearsal of speech. Based on this result, the authors suggested that the PT is important for both sensory and motor aspects of speech. Consistent with this view, the PT is also involved in a range of other speech production tasks, such as overt speech repetition and overt speech production (Peschke et al., 2009; Peschke et al., 2012; Shuster and Lemieux, 2005; Simmonds et al., 2014a; Simmonds et al., 2014b). Further, damage to the left PT is associated with conduction aphasia (Baldo et al., 2008; Buchsbaum et al., 2011; Northam et al., 2018; Rogalsky et al., 2015). In conduction aphasia, patients have intact speech perception and speech production skills but a specific problem in repeating words. Enhanced PT activation is, however, also observed during non-speech vocalization tasks such as humming of melodies. Thus, the effects in the PT observed during speech production tasks might not be specific to speech production per se, but rather the PT might support auditory-motor integration in general (Hickok et al., 2003).

In addition to the PT, effects related to auditory-motor integration have been reported elsewhere in the AC. For example, studies using real-time pitch shifting of one's own voice, which results in articulatory changes in the opposite direction to compensate for the artificial shift, have shown activation in the primary auditory cortex (Burnett et al., 1998; Purcell and Munhall, 2006; Tourville et al., 2008). It has also been shown that auditory-motor interactions in the AC are not restricted to vocal effectors, but that AC activation is also modulated during manual auditory-motor tasks, such as playing the piano (Baumann et al., 2007; Pa and Hickok, 2008) or tapping to musical rhythms (Chen et al., 2006; Chen et al., 2008a; Chen et al., 2008b; Chen et al., 2009). The role of auditory-motor integration outside the general framework of speech and music has, however, received less attention. Theoretically, it could be possible that strong motor influences on the AC are exclusive for vocal and musical sounds because of the inseparability of auditory perception and motor production of these sounds. Therefore, human fMRI studies investigating effects of both vocal and manual motor responding on processing of sounds outside the framework of speech and music are needed to understand the exact function of the connections between the auditory and motor cortex.

### **1.3.2 SUPPRESSION DURING MOTOR EXECUTION**

A large number of studies in humans and animals have reported that AC responses to the individuals' own voice are suppressed during overt and covert vocalization (Agnew et al., 2013; Christoffels et al., 2007; Curio et al., 2000; Eliades and Wang, 2003; Eliades and Wang, 2017; Flinker et al., 2010; Greenlee et al., 2011; Houde et al., 2002). This suppression is generally thought to be caused by modulatory signals (corollary discharge) from motor

areas providing predictive information on the expected auditory input (Christoffels et al., 2007; Reznik et al., 2014). However, this interpretation has been challenged by some authors. For example, similar motor suppression effects have been reported during manual responding (Schröger et al., 2015), suggesting that motor suppression is not specific to hearing one's own vocalizations.

The effects of manual motor processing on auditory processing have been extensively investigated using electroencephalography (EEG). In the widely used N1-suppression paradigm, subjects press a button to elicit a sound with a short (0–100 ms) or long (e.g., 1 s) delay. When the sound is presented immediately after a button press, subjects generally perceive that the button press triggered the sound. Using this paradigm, Schafer and colleagues (Schafer and Marcus, 1973) showed that the amplitude of the N1 component of the auditory evoked potential is smaller in response to sounds perceived to be self-administered than to those perceived to be computer-delivered. Most N1-suppression studies have interpreted the results to suggest that because the subjects perceive the sounds as self-caused, the sounds are fully predictable and therefore the processing of these self-caused sounds is suppressed (e.g., Aliu et al., 2009; Bäss et al., 2008; Bäss et al., 2009; Bäss et al., 2011; Martikainen et al., 2005; SanMiguel et al., 2013; Timm et al., 2013). However, it is still debated whether and to what extent the N1 suppression reflects predictive processes rather than some form of general suppression of auditory responses during motor behavior (Schröger et al., 2015). For example, in the study by Horváth and colleagues (Horváth et al., 2012), it was shown that N1 suppression is also observed when subjects do not perceive themselves as producing the sounds and the sounds just happen to randomly coincide with the manual response. Based on this result, the authors suggested that the N1-suppression effect might not be due to motor prediction but due to some form of general suppression of auditory responses during movement (motor-gating hypothesis, see also Kauramäki et al., 2010). In contrast, Timm and colleagues (Timm et al., 2014) showed that motor intention influences the N1-suppression effect. In their study, a sound was presented immediately after the subject either voluntarily or involuntarily moved his finger. Involuntary finger movements were triggered using transcranial magnetic stimulation of the motor cortex. The results showed that only those sounds that were triggered by voluntary movements caused N1 suppression. This supports the general idea that the N1-suppression effect can be caused by predictive mechanisms.

Motor suppression effects have also been demonstrated in intracellular AC recordings in mice. Schneider and colleagues (Schneider et al., 2014) showed that excitatory neurons in the mouse AC are suppressed before and during a wide range of natural movements that are not related to vocalization, such as locomotion and head movements. This suggests that AC cells are generally suppressed during movement. However, in concordance with the results of human studies using the N1-suppression paradigm, a

follow-up study by the same group showed that suppression effects in mouse AC neurons are stronger when the sound following the movement is predicted than when it is random (Schneider et al., 2018).

Together the results using the N1-suppression paradigm in humans and intracellular recordings in mice suggest that motor suppression in the AC consists of general motor-gating mechanisms and additional suppression related to motor prediction. In addition, the results of human fMRI studies show that motor suppression effects during vocalization can be observed in wide AC regions (Agnew et al., 2013; Christoffels et al., 2007; Curio et al., 2000; Flinker et al., 2010; Greenlee et al., 2011; Houde et al., 2002).

### **1.3.3 THE EFFECT OF MANUAL GRIP TYPES**

Manual grips in humans can be subdivided into the two general categories of precision and power grips. Precision grips are used to manipulate small objects such as a pencil by placing it between the thumb and fingertips, whereas power grips involving the whole hand are used to grasp bigger objects such as a screwdriver (Ehrsson et al., 2000). These grip types are supported by separate neural networks and they influence the processing of sensory information in distinct ways (Ehrsson et al., 2000; Grézes et al., 2003). For example, in the visual modality, it has been found that when subjects prepare to use a precision grip, the perception of small objects is facilitated, and when subjects prepare to use a power grip, the perception of large objects is facilitated (Symes et al., 2008). Other studies have shown that the size of a viewed object also interacts with the execution of precision and power grips. That is, people respond to smaller objects more quickly when using precision grips than power grips (Makris et al., 2013; Tucker and Ellis, 2001). Similar grip-type effects have also been reported in the auditory modality. In the study of Vainio and colleagues (Vainio et al., 2014), subjects prepared to use a precision or a power grip to respond to syllable targets. The syllables were of either high or low pitch, which was irrelevant for the task at hand. However, the authors found that the pitch of the syllables interacted with the grip types. That is, high-pitched syllables facilitated responses with precision grips, while low-pitched syllables facilitated responses with power grips. Together these results show that, at least at the behavioral level, manual grip type influences sensory perception and vice versa. However, it is currently unknown which brain regions and neural mechanisms support these auditory-motor interactions.

### **1.3.4 RELATIONSHIP BETWEEN TASK AND MOTOR EFFECTS**

The dual-stream model by Hickok and colleagues (Hickok, 2009, Hickok, 2012; Hickok, 2016; Hickok and Poeppel, 2007; Hickok et al., 2011), has been developed to account for auditory-motor-integration-related findings in

relation to speech processing. In this model, a dorsal stream serves speech production by forming a feedback loop between the posterior PT, IPL, motor cortical areas and inferior temporal gyrus. Specifically, the posterior PT serves as an interface between auditory functions in the AC and the motor cortex. This interface is particularly important for actions that are novel and non-automatic, such as repetition of vocalizations made by other individuals or learning how to produce novel sounds. Thus, in the model, the PT is an important hub for auditory-motor integration, and more specifically, in translating auditory input into motor programs and vice versa (Hickok, 2012; Hickok, 2016).

The model accounts for most of the aforementioned auditory-motor effects in the AC. However, previous studies have also shown that both auditory attention and auditory tasks modulate activation in the AC (cf. 1.2), including the posterior PT where most auditory-motor integration effects have been recorded (e.g., Harinen and Rinne, 2013; Harinen and Rinne, 2014; Häkkinen and Rinne, 2018; Häkkinen et al., 2015; Rinne et al., 2009; Rinne et al., 2012; Talja et al., 2015). Such attention- and task-related effects could easily have confounded motor-related effects in these regions in previous studies focusing only on auditory-motor integration effects. Furthermore, Hickok's model relies on the interpretation that the increased activation in the AC during covert rehearsal found in several studies (e.g., Buchsbaum et al., 2001; Hickok et al., 2009) is due to auditory-motor interactions. It is, however, evident that activation during covert rehearsal could equally be caused by some uncontrolled task-related factor, such as auditory imagery (see e.g. Simmonds et al., 2014a). That is, covert rehearsal does not only demand covertly producing the heard sound stimuli, but also other task-related operations on the sounds, such as working memory and mental imagery. Therefore, direct comparison of motor and auditory task effects should be performed within the same study.

## **1.4 REWARD INCENTIVE CUES AS A MEANS TO FACILITATE BEHAVIORAL TRAINING OF MONKEYS**

Teaching auditory tasks to non-human animals has posed a significant challenge, due to the time and effort needed. For example, in the study of Rinne and colleagues (Rinne et al., 2017) two monkeys were taught to perform an audiovisual selective attention task during fMRI. In their study, monkeys were rewarded for attending to stimuli in one modality while ignoring those in the other. The tasks were also taught to human participants, for whom the tasks were entirely trivial, and the participants learned them in a couple of trials. Monkeys, however, required tens of thousands of trials to reach criterion performance on the tasks. Why are auditory tasks used in humans so notoriously difficult to translate to animal studies? Firstly, communicating task instructions to animals is labor

intensive, since it depends on non-language learning. Secondly, most human paradigms rely upon rule-based choice tasks. Choice tasks involve two steps: first a target must be perceptually distinguished from a non-target; thereafter the correct action to the stimuli must be selected from a repertoire of response possibilities (e.g., withhold response to non-target sound/respond to target sound). Previous studies suggest that action selection is heavily dependent on frontal cortices (Buckley et al., 2009; Hoshi et al., 2000; Rushworth et al., 1997), which are less developed in monkeys than humans. However, comparative studies on auditory attention in humans and non-human animals are direly needed as some authors question the correspondence between monkey and human cognitive systems, including the auditory cognitive system (Patel et al., 2015; Schulze et al., 2012; Scott et al., 2012).

Paradigms based on reward incentive cues could provide a novel way to train active listening tasks in monkeys. For example, Minamimoto and colleagues (Minamimoto et al., 2009; Minamimoto et al., 2010) have shown that monkeys quickly learn to use visual reward incentive cues to influence their performance on a simple visual task. In these studies, monkeys first learned to perform a simple visual task (withholding a response while a red dot was presented and responding to a green dot). After monkeys mastered this simple task (ca. 100 trials), reward cues were incorporated. Throughout the trial either high reward (HiRe; e.g., picture of a dog) or low reward (LoRe; cat) cues were presented. The HiRe cue indicated that the monkeys would receive a large and instantaneous reward upon correct performance, while the LoRe cue indicated that correct performance would lead to a small and delayed reward. The reward cues drastically manipulated the monkeys' performance. That is, monkeys made fewer errors and had faster reaction times in trials with HiRe than LoRe cues. Importantly, the results showed that the monkeys recognized the visual categories within a single testing session. Thus, reward incentive cue paradigms achieve good task performance in monkeys within only a couple of hundreds of trials. This might relate to the fact that these paradigms demand no motor response selection or abstract task instruction that have been shown to be difficult for monkeys to comprehend. The utility of this paradigm becomes evident when one compares the speed of behavioral training to traditional paradigms that often requires tens of thousands of trials over months to years to reach adequate task performance in monkeys (Fritz et al., 2005b; Rinne et al., 2017).

Reward incentive cues could be used to manipulate auditory attention in monkeys. In human studies, reward-related manipulations have been found to strongly influence visual attention (Anderson, 2016; Anderson, 2018; Chelazzi et al., 2013; Della Libera and Chelazzi, 2006; Engelmann and Pessoa, 2007; Engelmann et al., 2009; Krebs et al., 2011; Pessoa, 2015). For instance, in the visual study by Engelmann and colleagues (Engelmann et al., 2009), reward incentive cues were used to indicate whether a correct

response would yield a high or low monetary gain. Performance was significantly better in HiRe than LoRe trials. Further, the fMRI results showed that the activity in the visual cortex was stronger during the HiRe trials than the LoRe trials. Importantly, the reward cues modulated visual cortex activity during the task period, and not during the processing of the cues. This suggests that the enhanced activity in the visual cortex was not due to stronger activity to the HiRe visual cue per se, but due to the fact that the reward cues directed attentional resources to the task-relevant stimuli. The reward manipulations resulted in similar effects in the visual cortex as has been previously obtained in attentional paradigms without differential reward value (see e.g., Liu et al., 2005). Together these findings suggest that reward incentive cue paradigms could be used to speed up the training of auditory tasks in monkeys, and study the neural correlates of attention-engaging auditory tasks using fMRI.

## 2 AIMS OF THE PRESENT THESIS

The present thesis investigated the effects of auditory attention, active listening tasks, motor responding and their interactions on activation patterns in the AC. Previous studies have shown that auditory attention, auditory tasks and auditory-motor integration all strongly modulate activation in the AC. However, as these modulatory influences have not been investigated in the same study it is currently not known whether these effects interact with each other. Also, fMRI was used to investigate auditory-attention-dependent modulations of the macaque monkey AC. Although current models of the human AC strongly rely on neuronal level measurements in monkeys, it is not currently known whether auditory attention modulates AC activation in monkeys in a similar manner as auditory attention modulates activation in the human AC.

Study I investigated the effects of manual motor responding on AC activation during auditory pitch discrimination and visual discrimination tasks. During fMRI, human subjects focused on either auditory or visual stimuli and reported the relative number of targets at the end of each task block. They also responded to each target either by using a precision grip, a power grip, or gave no overt responses. It was hypothesized that (1) activation in the human AC is stronger during auditory than visual tasks, (2) motor responding suppresses AC activity to sounds, and (3) AC activation is differentially modulated depending on whether subjects respond to targets using precision or power grips.

Study II used fMRI to investigate whether the effects related to auditory-motor integration and active listening task interact in the AC. Human subjects were presented with (Finnish) phonemic or nonphonemic vowels during auditory discrimination and 2-back tasks. They responded to targets by either overtly repeating the target vowel, by overtly producing a given response vowel or by pressing a response button. It was hypothesized that (1) auditory discrimination and 2-back tasks differently modulate activation in the AC and IPL, (2) vowel repetition is associated with stronger auditory-motor integration effects in the AC than vowel production, and (3) auditory-motor integration effects are stronger during repetition of nonphonemic than phonemic vowels as the requirements for auditory-motor integration are higher for non-phonemic vowels. In particular, it was hypothesized that (4) if auditory-motor and task-dependent effects interact in the AC, then auditory-motor effects could be at least partly related to changes in task demands rather than to auditory-motor integration per se.

Study III aimed to investigate attention-dependent activation modulations in the monkey AC using fMRI. To that end, first, a novel auditory paradigm was developed in order to facilitate and speed up behavioral task training. The paradigm was based on the general idea that monkeys would quickly learn to use incentive reward cues during an auditory

task. In particular, it was hypothesized that monkeys would be more motivated to actively process sounds during high- than low-reward trials and that this could be used to investigate the effects of active listening tasks on activation in the monkey AC.



## **3 METHODS AND RESULTS**

### **3.1 GENERAL METHODS IN STUDIES I AND II**

#### **3.1.1 HUMAN SUBJECTS**

In Study I (N = 16, 13 women, age 21–47 years, mean 25 years) and Study II (N = 20, 12 women; age 18–28, mean 24), subjects were healthy, normal hearing right-handed adults. All subjects provided informed consent. The ethical protocol was approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioral Sciences.

#### **3.1.2 PROCEDURES**

Subjects were presented with blocks of either concurrent but asynchronous auditory and visual stimuli (Study I) or auditory stimuli only (Study II). Each task block was followed by a rest block during which subjects focused on a fixation mark (+) presented in the middle of the screen. Graphic task instruction symbols were presented at the center of the screen a few seconds (4 s in Study I and 2.5 s in Study II) before the beginning of the next task block and remained on the screen throughout the task blocks. Before fMRI, subjects were carefully trained (1–3 h in total) to perform the demanding tasks.

During fMRI, auditory stimuli were delivered using Sensimetrics S14 insert earphones (Sensimetrics Corporation, Malden, USA). Scanner-noise was attenuated through the insert earphones, circumaural ear protectors (Bilsom Mach 1) and viscous foam pads attached to the sides of the head coil. All visual stimuli were presented in the middle of the screen via a mirror fixed to the head coil.

**Table 1.** *Stimuli and experimental design in human studies*

	<b>Study I</b>	<b>Study II</b>
<b>Auditory stimulation</b>		
Stimuli	Pairs of IRN bursts	Pairs of Ph, NPh and piPh vowels
Duration of sound pair	90 + 90 ms	90 + 90 ms
Sound pair onset-to-onset interval	800–1000 ms	1400–1900 ms
Between-pairs difference	Pitch, corresponding to 200–1400 Hz (200 levels)	Pitch, 77–156 Hz for male subjects and 122–254 Hz for female subjects (9 levels)
Within-pair difference	Pitch, corresponding to 9.5–95.5 mel	Pitch, 0.7 semitones
<b>Visual Stimulation</b>		
Stimuli	Gabor gratings	-
Duration	100 ms	-
Onset-to-onset interval	250–450 ms	-
Orientation	12 levels (180°)	-
Within-pair difference	Orientation, 14.5°	-
<b>Experiment</b>		
Conditions	6	18
Target to non-target ratio (auditory)	44–55%	28–42%
Target to non-target ratio (visual)	0.44–0.55	-
Blocks per condition	12	6
Block duration	12.5 s	12.5 s
Rest duration	12.5 s	10 s
Duration of experiment	34 min	68 min
Data acquisition	2013–2014	2014

### 3.1.3 STIMULI

In Study I, the auditory stimuli were pairs of iterated rippled noise (IRN) bursts varying in pitch and the visual stimuli were Gabor gratings varying in orientation (Table 1). In Study II, the stimuli consisted of Finnish phonemic

(Ph) and nonphonemic (NPh) vowels synthesized using the Praat software package (version 5.1.12, [www.praat.org](http://www.praat.org)) for a previous study (Harinen and Rinne, 2013; Fig. 1A). There were three Ph (/a/, /i/ and /u/) and three NPh (NPh1, NPh2, NPh3) vowel categories with nine vowels in each. In addition, pitch modulated vowel stimuli (piPh) were synthesized (Fig. 1B). Each of the three piPh categories (low, middle, high; separated by 4 semitones) contained three different vowel sounds with three pitch levels.

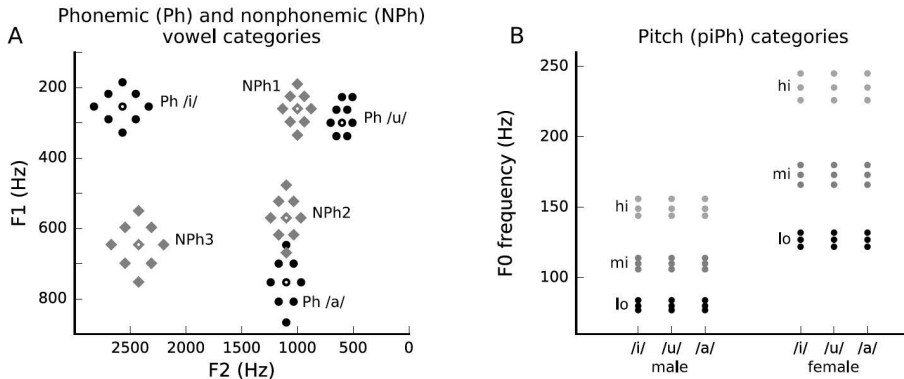


Figure 1 Stimuli used in Study II. (A) In the vowel task blocks, pairs of vowels (F0 150 Hz) from three phonemic (Ph, black circles) or three nonphonemic (NPh, gray diamonds) vowel categories were presented. Each category contained nine different vowels. The Ph categories were based on typical Finnish /i/, /u/ and /a/ phonemes. (B) In the pitch task blocks, pairs of pitch modulated vowels from three pitch categories (low, middle and high) were presented. Each category contained nine different sounds (three different vowels and three pitch levels). The pitch-modulated vowels were slightly different for male and female subjects.

## 3.2 METHODS IN STUDY III

Study III was conducted using three adult male (M1 6, M2 6 and M3 8 years of age) rhesus monkeys. All nonhuman animal work was performed at Newcastle University and was approved by the Animal Welfare and Ethical Review Body at Newcastle University and by the UK Home Office. The work complies with the Animal Scientific Procedures Act (1986) and with the European Directive on the protection of animals in research (2010/63/EU). All persons involved in animal handling and procedures were certified and their work was regulated by the UK Home Office.

### 3.2.1 STIMULI AND TASKS

In Study III, monkeys were first taught a simple auditory task. Thereafter, it was tested whether auditory (AudCue1, AudCue2 experiments) or visual

(VisCue experiment) reward cues could be used to influence auditory task performance in monkeys. As the visual reward cues caused stronger behavioral effects than the auditory reward cues, this paradigm was selected for the fMRI experiment.

In the auditory task, first a yellow dot (visual 'wait' signal) was presented on a grey background in the middle of a computer screen. The yellow dot remained on the screen until the end of the trial. After 500–1500 ms, an auditory 'go' signal (macaque 'coo', 400 ms in duration) sound was presented. If the monkey responded to the 'go' signal by pressing a response lever within 200–1300 ms (hit response), then the monkey received an immediate juice reward after which the next trial was initiated. Incorrect responses (early responses before the response window) or misses (no response during the window) were not rewarded and resulted in a 200 ms delay before the next trial. The monkeys mastered (i.e., above chance level performance) this auditory task quickly, i.e. within one training session (ca. 500 trials).

Next, auditory high (HiRe) and low (LoRe) reward incentive cues were introduced. In the AudCue1 experiment (Fig. 2A), the HiRe cue was a narrow-band noise burst (bandpass filter centered at 2 kHz, width 2 kHz, 3 Hz sinusoidal amplitude modulation, 90% depth) and the LoRe cue was a sinusoidal tone (2 kHz sinusoid, 8 Hz amplitude modulation). In the AudCue2, the HiRe cue was a high-pitched sinusoidal tone (2 kHz, 8 Hz amplitude modulation) and the LoRe cue was a low-pitched tone (200 Hz, 3 Hz amplitude modulation). A HiRe (50%) or LoRe cue was presented in each trial. HiRe cues predicted that a large reward (ca. 1 ml) would be delivered immediately after a correct response, whereas the LoRe cues predicted that a correct response would result in a small (ca. 0.1 ml) and delayed reward (7 s after a correct response). The cues were always presented from trial onset until the end of the trial (including the 7 s delay in LoRe hit trials).

When data collection was completed in the AudCue experiments, a visual reward cue (VisCue) experiment was conducted. In the VisCue experiment, HiRe and LoRe cues consisted of high and low spatial frequency vertical gratings, respectively (Fig. 2B), the auditory 'go' signal was a 4 kHz sinusoidal tone (duration 400 ms) and the visual 'wait' signal (yellow dot) was replaced by an auditory 'wait' signal (2 kHz tone, 8 Hz amplitude modulation).

For fMRI, the paradigm used in the VisCue experiment was slightly modified due to fMRI imaging timing constraints. The target was presented later than in the behavioral experiments (2300–3000 ms after trial start). In addition, the auditory 'wait' signal was either a low-pitch tone (0.2 kHz sinusoid, 3 Hz amplitude modulation; 50% of runs) or a high-pitch tone (2 kHz sinusoid, 8 Hz amplitude modulation; 50% of the runs). This sound was always played until the end of the MRI volume acquisition irrespective of the monkey's responses. Further, in HiRe hit trials, reward delivery started after volume acquisition (to avoid movement effects associated with juice

consumption). Early-response and miss trials were terminated after the completion of the volume acquisition. Otherwise the task, visual cues and auditory ‘go’ signals were identical to those used in the VisCue experiment (Fig. 2B).

### **3.2.2 PROCEDURES**

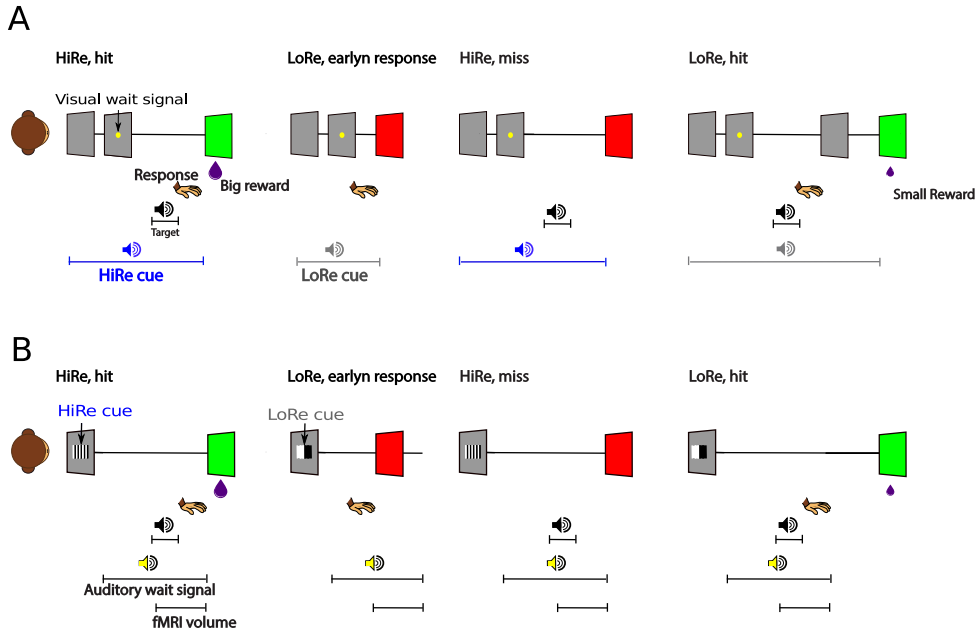
In the behavioral experiments, visual cues and visual ‘wait’ signals were presented in the middle of a computer screen in front of the monkey (distance 1 m). All sounds were presented from two loudspeakers (distance 1 m, 30° to the left and right from the center of the screen; 65 dB SPL at the monkey’s head).

During fMRI, an fMRI volume was acquired 2500 ms after the onset of the auditory wait signal. That is, the volume was acquired during the rising edge of the BOLD response to the auditory ‘wait’ signal (Baumann et al., 2010; Fig. 1B, Table 1), which was identical across all trials in each session.

During fMRI, sounds were presented via MRI-compatible headphones and the visual stimuli were projected to a screen that the monkeys could see in a mirror in front of them.

Monkeys M1 and M2 were already implanted with an MRI compatible head post for head immobilization. Monkey M3, in contrast, was previously trained to perform tasks wearing a head-immobilizing facemask and helmet (Slater et al., 2016). Prior to the experiments, all of the animals were acclimated to work within a primate testing-chair and to allow the required periods of head immobilization. No contrast agent such as monocrySTALLINE iron oxide nanoparticles (MION) were used in fMRI.

Animals were on a customized fluid control procedure to ensure motivation to work on the tasks. The fluid was not restricted on days when the animals were not being tested.



**Figure 2** In all conditions, monkeys responded to an auditory target in order to receive a juice reward. In HiRe trials, monkeys received a large reward (1 ml) immediately after a correct response. In LoRe trials, monkeys received a small reward (0.1 ml) upon correct performance after a 7 s delay. (A) Trials in the auditory reward cue conditions (AudCue1 and AudCue2). If the monkey responded to the target 200–1300 ms from its onset in a HiRe trial, then a big juice reward was immediately delivered and the screen turned green. During LoRe trials, the juice reward was delayed and small. Note that the LoRe cue was presented until the reward was delivered. Responses before the response window (early response) resulted in a red screen and trial termination. (B) Four exemplary trials in the fMRI experiment with visual reward cues.

### 3.3 FMRI DATA ACQUISITION AND ANALYSIS

In Study I and Study II, a high-resolution T1-weighted anatomical image was first acquired. Based on this image, the middle slice of the functional echo-planar imaging (EPI) image was aligned in the same orientation as the Sylvian fissures. At the end of the imaging session, a T2-weighted image using the same imaging slices as in the EPI-series but a denser in-plane resolution was acquired for coregistration purposes. In Study III (data acquisition 2015), two structural images (full-head EPI with extra slices and a high resolution MDEFT image) aligned with the functional volumes were acquired for coregistration purposes. Details of fMRI data acquisition are shown in Table 2.

In all studies, fMRI data analysis was performed using FMRIB Software Library (FSL; Jenkinson et al., 2012). The data were first corrected

for motion artifacts and high pass filtered. In Studies I and II, the data were thereafter resampled to the standard cortical surface (using Freesurfer, Dale et al., 1999) and spatially smoothed. General linear model was used for the first level global voxel-wise analysis (in surface space). In Study II, the data from the two runs were combined in a second-level fixed effects analysis. In Study III, the first level analysis was conducted in 3D EPI space and thereafter the data from each run were co-registered to a template monkey brain (McLaren et al., 2009; Petkov et al., 2015). Next, the contrast parameter estimates from the first level analysis were resampled to the cortical surface of a template monkey brain (McLaren et al., 2009; Petkov et al., 2015) and smoothed on the surface.

**Table 2.** *Details of fMRI data acquisition.*

	<b>Study I and II</b>	<b>Study III</b>
<b>Scanner</b>		
Type	MAGNETOM Skyra 3 tesla scanner	Bruker Vertical MRI 4.7 tesla scanner
Number of head-coil channels	20 channels	1 channel
<b>EPI parameters</b>		
TR	2.2 s	5–9 s
Time of acquisition	2.2 s	2 s
TE	30 ms	22 ms
Flip angle	78°	90°
Voxel matrix	96 × 96	96 × 96
Slice thickness	2 mm	2 mm
Field of View	18.9 cm	9.6 cm
Slices	29	20
In plane resolution	2.0 × 2.0 mm <sup>2</sup>	1.0 × 1.0 mm <sup>2</sup>
Imaging paradigm	Continuous	Sparse

In all studies, group analysis was performed using PALM (Permutation Analysis of Linear Models, version alpha26, Winkler et al., 2014; 10 000 permutations). Significance was addressed using permutation inference. Correction for multiple comparisons was performed using threshold-free cluster enhancement (Study I) or cluster mass correction (initial cluster forming threshold  $Z > 2.3$  in Study II and  $Z > 2.6$  in Study III).

## **3.4 STUDY I. THE EFFECT OF PRECISION AND POWER GRIPS ON ACTIVATION IN HUMAN AUDITORY CORTEX**

### **3.4.1 TASKS**

During all tasks, the auditory stimulation consisted of pairs of IRN bursts. The bursts were either equal (ca. 50% of the pairs) or slightly different in pitch. The visual stimulation consisted of Gabor gratings. The orientation of the Gabor gratings changed in ca. 50% of the cases. During the auditory task blocks, subjects were instructed to ignore the visual stimuli, focus on the sound pairs, and to respond to targets (pairs with a pitch change). In half of the task blocks, there were more (70–75%) targets with a pitch increase (second burst higher than the first burst), whereas in the other half of the blocks there were more targets with a pitch decrease. During visual task blocks, subjects ignored the auditory stimuli and responded when there was a change in the orientation of the Gabor gratings. In half of the blocks, there were more (70–75%) clockwise (CW) changes, whereas in the other half of the blocks there were more counterclockwise (CCW) orientation changes. After each auditory block, an arrow was presented for 2 s. The arrow pointed either up or down with equal probability indicating the question “there were more targets with pitch increase?” or “there were more targets with pitch decrease?”. Subjects were to answer this question by pressing a response button once (yes) or twice (no). After the visual task blocks, there was an identical task except that the arrow pointed either left or right (more targets with CCW or CW change, respectively).

### **3.4.2 RESPONSES**

Subjects responded either using precision grips, power grips or made no overt responses. The combination of task and response type yielded six conditions (auditory: precision,  $A_{pr}$ ; power,  $A_{po}$ ; no-response,  $A_{no}$  and visual: precision,  $V_{pr}$ ; power,  $V_{po}$ ; no-response,  $V_{no}$ ). Responses were given using a modified joystick for precision grips (Current Designs, USA), a grip force bar for power grips (Current Designs, USA), and a button on the joystick device for button presses. The joystick and the grip force bar were attached to a custom-made plastic frame that was placed on the subject’s torso.

It was reasoned that sensorimotor modulations would be stronger if the motor responses were executed with one out of two response alternatives. Therefore, subjects used slightly different grips (two or three-finger precision grip and two or five finger power grip) depending on the type of the target (i.e., rising pitch/falling pitch or clockwise/counterclockwise target).



### 3.4.3 RESULTS

In line with a large number of previous studies (e.g., Hall et al., 2000; Petkov et al., 2004; Rinne, 2010; Rinne et al., 2009; Rinne et al., 2012), activation in regions extending from the anterior to the posterior STG were strongly modulated by the active auditory task (i.e., activation to identical sounds were stronger during auditory than visual tasks; Fig. 3A).

Comparisons between the motor conditions and corresponding non-motor conditions revealed wide motor suppression effects in STG and IPL regions (Fig. 3B:  $A_{pr} < A_{no}$ ,  $A_{po} < A_{no}$ ; Fig. 3C:  $V_{pr} < V_{no}$ ,  $V_{po} < V_{no}$ ). Importantly, these effects were similarly observed during the auditory and visual tasks. As the motor suppression effect in the AC was observed also when task-irrelevant sounds were presented during the visual task (subjects responded to targets in the visual task), this effect is probably not related to auditory-motor integration, but to motor responding in general. This result is also in line with previous studies using EEG in human subjects and intracellular recordings in mice that show general suppressive modulations in the AC during movement (Horvath et al., 2012; Schneider et al., 2014). Importantly, the results of Study I show that such motor suppression effects are observed in broad regions extending from the anterior to the posterior STG and IPL.

During the auditory task, activation was stronger during the precision than power grip blocks in the left lateral HG, right anterior STG and temporal pole (Fig. 3D). The corresponding contrast for visual tasks ( $V_{pr} > V_{po}$ ) revealed no significant effects associated with the grip type.

Additional analyses were conducted to evaluate whether more lenient thresholds would reveal additional grip effects. These analyses revealed that activation was (non-significantly) stronger in the IPL during the precision grips than power grips both for auditory and visual task conditions (Fig. 3E). These nonsignificant effects in the IPL could be due to higher task requirements for precision than power grips (i.e., higher demands for integration between motor programs and somatosensory feedback, see Ehrsson et al., 2000).

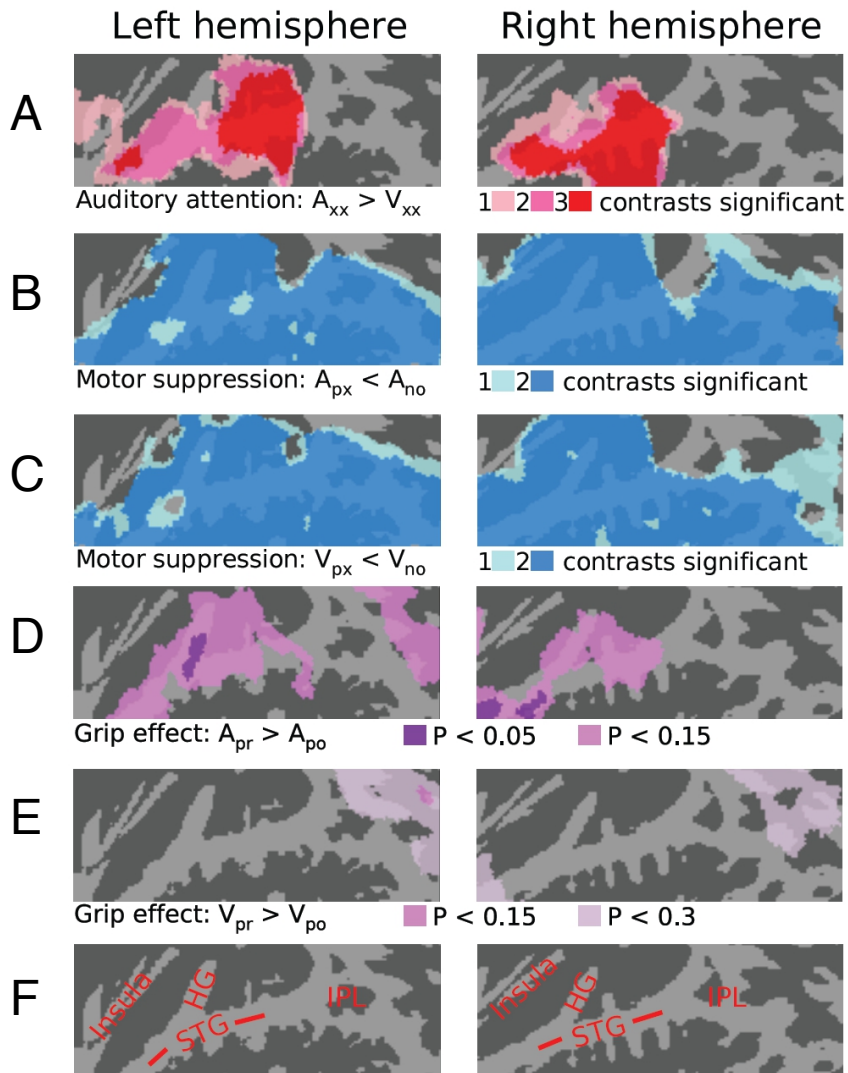


Figure 3 (A) Modulations associated with active listening were revealed by comparing activation during auditory and visual tasks with identical stimuli and motor responses. The results of three separate contrasts ( $x$  stands for the lower cases in  $A_{pr} > V_{pr}$ ,  $A_{po} > V_{po}$ ,  $A_{no} > V_{no}$ ) are shown so that areas where any one of the contrasts was significant are shown in pink and areas where all three contrasts were significant are shown in red. (B, C) Motor suppression during the auditory tasks and visual tasks. The results of two contrasts ( $A_{pr} < A_{no}$ ,  $A_{po} < A_{no}$  in B and  $V_{pr} < V_{no}$ ,  $V_{po} < V_{no}$  in C) are plotted so that areas where either of the contrasts was significant are shown in light blue and areas where both contrasts were significant are shown in darker blue. (D, E) Contrasts between  $A_{pr}$  and  $A_{po}$  (D) and  $V_{pr}$  and  $V_{po}$  blocks (E). The results are plotted with three thresholds (corrected  $P < 0.05$ , corrected  $P < 0.15$  and corrected  $P < 0.3$ ). (F) Anatomical labels: STG superior temporal gyrus, HG Heschl's gyrus, PT planum temporale, IPL inferior parietal lobe.

## 3.5 STUDY II. INTERACTION OF THE EFFECTS ASSOCIATED WITH AUDITORY-MOTOR INTEGRATION AND ATTENTION-ENGAGING LISTENING TASKS

### 3.5.1 TASKS

In Study II, subjects were presented with vowel pairs (phonemic, non-phonemic or pitch-modulated vowel pairs) during discrimination or 2-back tasks. In the discrimination tasks (Fig. 4A), subjects were to respond when the first and second part of the vowel pair were the same. In the 2-back tasks (Fig. 4B), they were to respond when the vowels in the pair belonged to the same vowel or pitch category (category /i/, /u/ or /a/ in the Ph blocks; category NPh1, NPh2 or NPh3 in the NPh blocks; category low, middle or high in the piPh blocks) as the vowels in the pair two trials before.

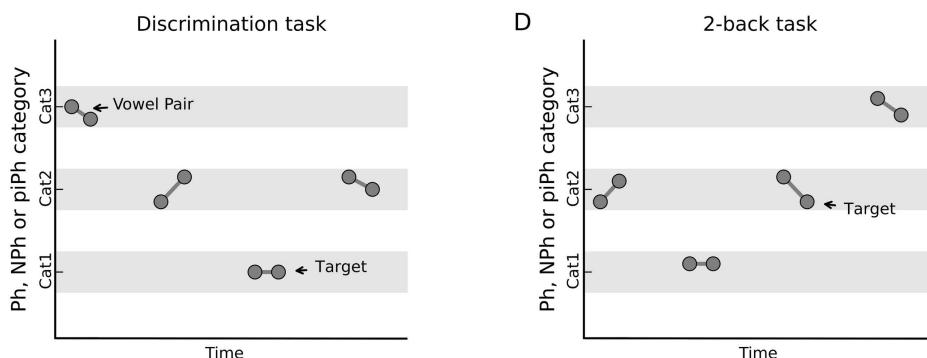


Figure 4 In the discrimination task, subjects were to indicate when the first and the second part of the pair were identical. In the 2-back task, subjects indicated when the pair belonged to the same vowel or pitch category as the one presented two trials before.

### 3.5.2 RESPONSES

In different blocks, subjects responded to targets either by pressing a response button (button response blocks; Bu), by overtly vocalizing the response vowel (/æ/, /o/ or /y/) displayed on the computer screen throughout the block (production response blocks, Pr) or by overtly vocalizing the last part of the target vowel pair (phoneme or nonphoneme repetition response blocks, PhRe or NPhRe). During pitch-modulated vowel repetition response blocks, subjects hummed the pitch of the second vowel of the target pair (pitch repetition response blocks, piPhRe). The task conditions are summarized in Table 3.

**Table 3.** *Experimental conditions: Phonemic (Ph), nonphonemic (NPh) or pitch-modulated phonemic (piPh) vowel pairs were presented during discrimination or 2-back blocks. Subjects responded to targets either by repeating the last part of a target pair (PhRe, NPhRe), humming its pitch (piPhRe), producing a given response vowel (Pr), or pressing a button (Bu).*

Task	Vowel	Response
Discrimination or 2-back	Ph	Re, Pr, Bu
	NPh	Re, Pr, Bu
	piPh	Re, Pr, Bu

### 3.5.3 RESULTS

Two-way analysis of variance (ANOVAs) were used to study the effect of task (discrimination vs. 2-back) and motor-response type (repetition vs. button or production vs. button; Re vs. Bu or Pr vs. Bu, respectively) in the AC and adjacent regions. The ANOVAs revealed that activation in wide regions extending from the insula to the STG and IPL depended on the task (Fig. 5A). These task-dependent modulations were due to higher activation in the insula and STG during discrimination than 2-back tasks, while the 2-back tasks were associated with stronger activation in the IPL. These task-dependent activation patterns are consistent with those observed in previous studies using similar stimulus and task conditions (Harinen and Rinne, 2013, 2014).

Unexpectedly, the ANOVAs also showed extensive main effects of motor-response type in STG and IPL regions (Fig. 5B) due to activation being stronger during the vocal-response (Re or Pr) blocks than button-response (Bu) blocks.

Task and motor-response-type effects interacted in the IPL (Fig. 5C), where activation was stronger during vocal-response blocks than during button-response blocks in the discrimination tasks but not in the 2-back tasks. However, no interactions were detected in the STG suggesting that task and response effects in the STG are caused by independent mechanisms.

## Task (Discr, 2-back) × Response (Re/Pr, Bu)

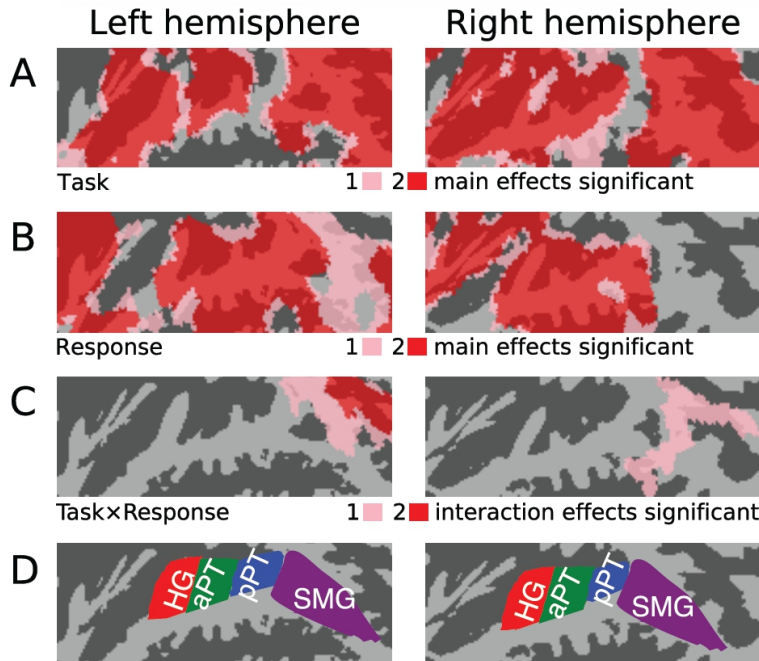


Figure 5 Task and motor-response-type effects in the AC and adjacent regions (Ph, NPh and piPh sound blocks collapsed together). Results of two task × response ANOVAs, where the task factor (Discr, 2-back) was the same but the response factor varied [(repetition, button) or (production, button); (Re, Bu) or (Pr, Bu)]. The results of the two separate ANOVAs are shown so that areas where one main effect (A, B) or interaction (C) was significant (N = 20, permutation inference, Family-wise error rate (FWE) corrected  $P < 0.05$ ) are shown in pink and areas where both effects were significant are plotted in red. (D) Areas used in region-of-interest (ROI) analysis: HG Heschl's gyrus, aPT anterior planum temporale, pPT posterior planum temporale, and SMG supramarginal gyrus.

Because repetition responses demanded translation of acoustic input to the corresponding motor program, these responses should be associated with more auditory-motor integration than the production responses. Consistently, activation in the left PT and IPL was stronger during vowel repetition than production blocks (Fig. 6A). However, although activation was stronger in STG regions during non-phonemic than phonemic button blocks (not shown), no significant differences were observed between activation in phonemic and nonphonemic repetition blocks. Further, the pitch task blocks with repetition responses (subjects responded to targets by humming the pitch of vowels) were not associated with stronger activation than pitch-task blocks with vowel-production responses (Fig. 6B). This was

unexpected because previous studies have shown auditory-motor integration effects in the left PT during both covert speech and humming tasks (Hickok et al., 2003). However, it is possible that the specific effects related to pitch repetition might have been affected by some other uncontrolled difference between humming and vowel responding. Subsequent studies on pitch-repetition should test this possibility directly using pitch-repetition and pitch-production responses with exactly the same motor requirements.

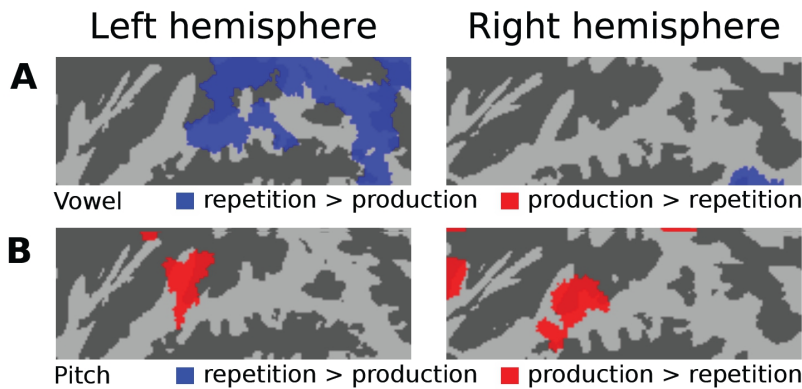


Figure 6 Direct comparisons between repetition- and production-response blocks (N = 20, permutation inference, FWER corrected  $P < 0.05$ ). (A) During vowel tasks (collapsed across Ph and NPh blocks), activation in the left PT and IPL was stronger when the tasks were performed with repetition than production responses (blue). (B) During pitch tasks, activation in HG and anterior PT regions was stronger during vowel-production (Ph) than during humming of the pitch of the target (red).

Region-of-interest (ROI) analyses were performed to investigate the motor- and vocal-response-type effects in more detail. Mean signal magnitudes for each condition in four ROIs (defined anatomically, Fig. 5D) were extracted and collapsed across discrimination and 2-back tasks. The results of three- and two-way ANOVAs testing these effects in each ROI are shown in Table 4. In line with the effects shown in Fig 5, significant main effect of motor-response type (Re, Bu) was present in all ROIs. Further, a significant interaction between hemisphere and vocal-response type (Re, Pr) was observed in the pPT and supramarginal gyrus (SMG) during the vowel tasks (Re > Pr in the left hemisphere, cf. Fig. 6A), whereas during the pitch tasks there was a vocal-response-type main effect in the HG and aPT (Pr > Re, cf. Fig. 6B). Importantly, no significant interactions were observed between vowel type (Ph, NPh) and vocal-response type (Re, Pr), indicating that auditory-motor interactions were similar with both vowel types and not stronger with the non-phonemic vowels.

**Table 4.** *The results of ANOVAs conducted on the ROI data (discrimination and 2-back tasks collapsed together). For vowel tasks, the ANOVAs tested the effects of motor-response type (Re, Bu), vocal-response type (Re, Pr), and vowel type (Ph, NPh). For pitch tasks (piPh stimuli), the ANOVAs tested the effects of motor-response type (Re, Bu) and vocal-response type (Re, Pr). The ANOVAs were conducted separately for each ROI. Only significant (false detection rate corrected within and across ANOVAs,  $P < 0.05$ ) effects are reported.*

<b>Vowel tasks: hemisphere <math>\times</math> vowel type (Ph, NPh) <math>\times</math> motor-response type (Re, Bu)</b>			
<b>ROI</b>	<b>Significant effect</b>	<b>F (1,19)</b>	<b>P</b>
HG	Vowel type	13	.008
	Motor-response type	63	.001
	Hemisphere $\times$ motor-response type	13	.008
aPT	Vowel type	7.4	.037
	Motor-response type	65	.001
pPT	Motor-response type	34	.001
	Hemisphere $\times$ motor-response type	10	.016
SMG	Vowel type	21	.002
	Motor-response type	10	.016
	Hemisphere $\times$ motor-response type	30	.001
<b>Vowel tasks: hemisphere <math>\times</math> vowel type (Ph, NPh) <math>\times</math> vocal-response type (Re, Pr)</b>			
pPT	Hemisphere $\times$ vocal-response type	18	.002
SMG	Hemisphere $\times$ vocal-response type	9.0	.024
<b>Pitch tasks: hemisphere <math>\times</math> motor-response type (Re, Bu)</b>			
HG	Hemisphere $\times$ motor-response type	10	.013
SMG	Hemisphere $\times$ motor-response type	20	.002
<b>Pitch tasks: hemisphere <math>\times</math> vocal-response type (Re, Pr)</b>			
HG	Vocal-response type	13	.008
aPT	Vocal-response type	7.0	.042

### 3.6 STUDY III. REWARD CUES READILY DIRECT MONKEYS' AUDITORY PERFORMANCE RESULTING IN BROAD AUDITORY CORTEX MODULATION AND INTERACTION WITH SITES ALONG CHOLINERGIC AND DOPAMINERGIC PATHWAYS

#### 3.6.1 RESULTS

First, it was determined whether the reward cues influenced the monkey's motivation to perform the simple auditory task. Mean performance across each run and animal in the AudCue1 and VisCue experiments is shown in Fig. 7. Linear mixed models with the repeated measures factor reward cue (HiRe, LoRe) and fixed factors experiment (AudCue1, AudCue2) and monkey (M1, M2, M3) were used to test whether the reward manipulation (HiRe vs. LoRe)

modulated performance in the AudCue1 and AudCue2 experiments. Each performance parameter (hit rate, HR; early-response rate, ER; miss rate, MR; and reaction time, RT) was analyzed using separate linear models. All models included intercept for run. These models revealed significant reward cue main effects: HR was higher, MR lower and RT faster in the HiRe than LoRe trials (HR:  $F_{1,64}=125$ ,  $p < .001$ , MR:  $F_{1,61}=162$ ,  $p < .001$ , RT:  $F_{1,75}=42$ ,  $p < .001$ ). That is, the monkeys' performance was significantly influenced by the reward cues. However, the linear models also revealed significant reward cue  $\times$  monkey (HR:  $F_{2,62}=62$ ,  $p < .001$ , ER:  $F_{2,114}=6$ ,  $p < .01$ , MR:  $F_{2,62}=54$ ,  $p < .001$ , RT:  $F_{2,75}=21$ ,  $p < .001$ ) and reward  $\times$  experiment  $\times$  monkey (HR:  $F_{2,78}=7.2$ ,  $p < .001$ , ER:  $F_{2,110}=12$ ,  $p < .001$ , MR:  $F_{2,73}=8.3$ ,  $p < .001$ ) interactions, indicating that the effects associated with the reward cue manipulation varied between the two experiments and the three monkeys. In the AudCue paradigm, M3 showed nearly categorical preference for the HiRe cue trials, while M1 and M2 showed more subtle effects. In all monkeys, there were, however, significant effects for one or more performance parameter.

Correspondingly, performance in the VisCue experiment was analyzed using linear mixed models with repeated measures factor reward cue (HiRe, LoRe) and fixed factor monkey (M1, M2). These analyses showed significant reward cue main effects for HR ( $F_{1,46}=315$ ,  $P=3.3 \times 10^{-22}$ ; HiRe > LoRe), MR ( $F_{1,46}=144$ ,  $P=9.1 \times 10^{-16}$ ; HiRe < LoRe) and RT ( $F_{1,46}=33$ ,  $P=6.9 \times 10^{-7}$ ; HiRe < LoRe). As can be seen in Fig. 7, monkeys (M1 and M2) showed nearly categorical preference for the HiRe trials over the LoRe trials. Performance during fMRI was similar to that of the VisCue experiment.

Taken together, the behavioral effects indicated that although two of the monkeys showed stronger effects in the VisCue paradigm than in the AudCue paradigm, performance in both paradigms was significantly better in HiRe trials than LoRe trials. Importantly, these performance effects were achieved within only 13–30 training runs ( $\approx 10$  training days). In previous studies using traditional paradigms, task training typically has required months or even years of systematic training (Downer et al., 2015; Fritz et al., 2005b; Niwa et al., 2015; Rinne et al., 2017). Further, the reward incentive cue paradigm in Study III required no motor response selection, abstract task instruction or other demanding training components, which have proven difficult to teach to monkeys (Minamimoto et al., 2009; Minamimoto et al., 2010).

Finally, monkeys M1 and M2 were trained to perform a slightly modified version of the VisCue paradigm during fMRI. It was hypothesized that the monkeys would focus more on the auditory wait signals during HiRe than LoRe trials and thus stronger AC activation would be observed during HiRe trials. Consistently, the results revealed stronger activation in STG regions bilaterally in HiRe than LoRe trials and no regions showed higher activation during LoRe than HiRe trials (Fig. 8).



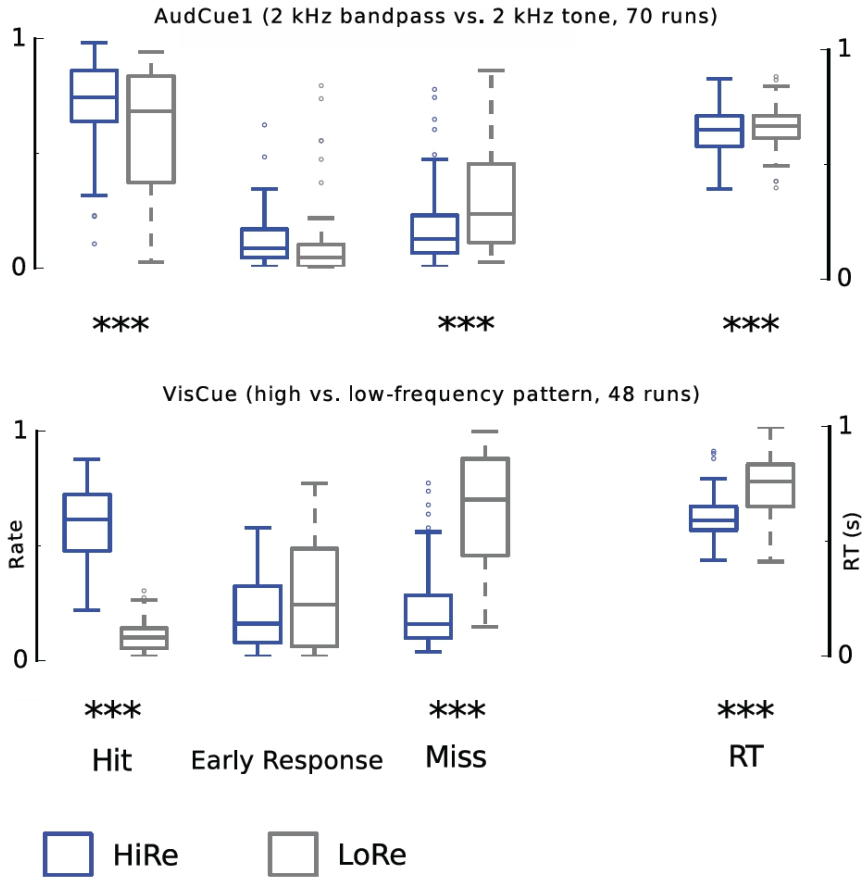
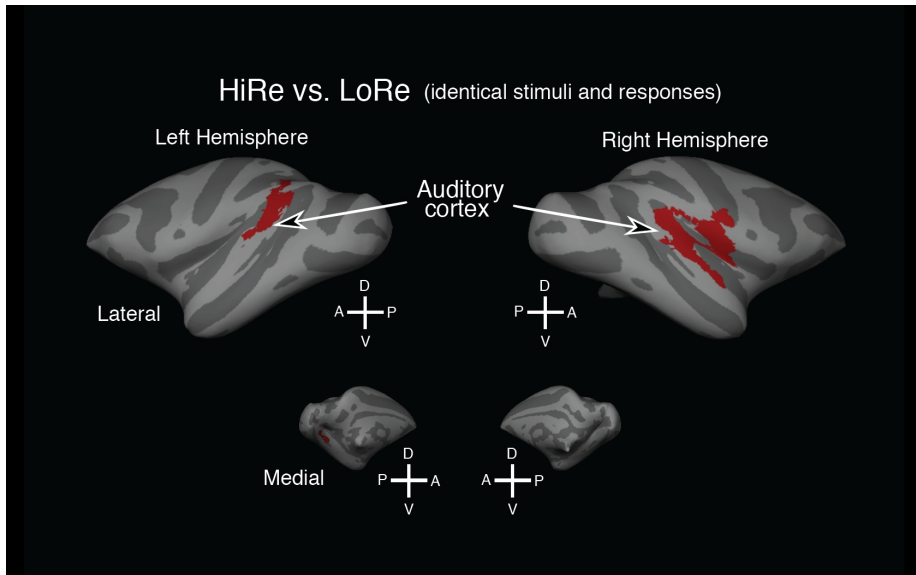


Figure 7 Performance during behavioral AudCue1 (top) and VisCue (bottom) experiments. The scale for reaction times (RTs) is given at the right. Asterisks indicate significant differences between HiRe and LoRe trials [i.e., main effect of reward, \*  $p < 0.05$ , \*\*  $p < 0.01$  and \*\*\*  $p < 0.001$ ; AudCue: linear mixed model with factors reward (HiRe, LoRe), monkey (M1, M2, M3) and experiment (AudCue1, AudCue2); VisCue: linear mixed model with factors reward (HiRe, LoRe) and monkey (M1, M2)].



**Figure 8** Brain areas showing stronger activation during HiRe than LoRe trials. Results are shown on inflated cortical surfaces (gyri: light gray; sulci: dark gray). The comparisons (Welch's *v* test) were performed in surface space across 1<sup>st</sup>-level contrast parameter estimates and permutation inference was used to assess statistical significance (19 HiRe vs. baseline and 19 LoRe vs. baseline contrast parameter estimates; the runs of each monkey were treated as a permutation and variance group to accommodate heteroscedasticity; initial cluster-forming *Z* threshold was 2.6, cluster-corrected *P* < 0.05). Abbreviations: D, dorsal; V, ventral; A, anterior; P, posterior.

To better understand the source of the reward cue effects on AC activation, further analyses were conducted using ROIs constructed by dividing the STG into four segments in the anterior-posterior direction. Mean signal magnitudes were computed separately for each ROI and hemisphere, and the extreme 15% of the values were excluded from the analysis. Fig. 9 shows mean signal magnitudes in each ROI during HiRe and LoRe hit, early response and miss trials. This figure also summarizes the results of separate tests comparing signal magnitude between HiRe and LoRe trials in each ROI.

Significant reward cue differences were observed in hit and early-response trials, but not in miss trials. These results are consistent with the hypothesis that the monkeys more actively processed the auditory wait signals during the HiRe trials than the LoRe trials. Especially since no consistent HiRe vs. LoRe differences were observed during miss trials when the monkeys were likely paying less attention to the sounds irrespective of the cue type.

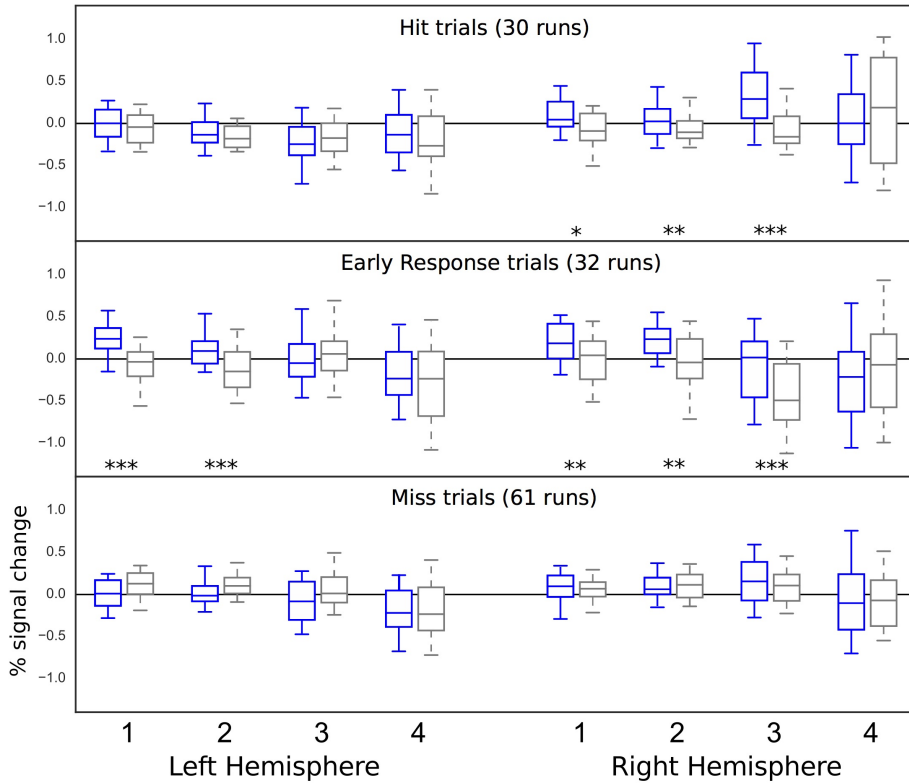


Figure 9 Mean signal magnitude in differences between HiRe trials (blue) and LoRe trials (grey) in each anatomically defined STG ROI. To remove outliers, the extreme 15% of the values were discarded. Asterisks indicate significant pair-wise tests comparing signal magnitude between HiRe and LoRe trials in each ROI (permutation-based significance testing using Welch's  $v$  tests, two-sided, 10 000 permutations, FWER corrected across all pair-wise comparisons, \*  $p < 0.05$ , \*\*  $p < 0.01$  and \*\*\*  $p < 0.001$ ). The difference between HiRe and LoRe trials was significant only during the hit and early-response trials.

## **4 GENERAL DISCUSSION**

Previous human studies show that attention-engaging auditory tasks are associated with enhanced activation in wide regions of the AC (Alho et al., 2014; De Martino et al., 2015; Hall et al., 2000; Petkov et al., 2004; Riecke et al., 2018; Rinne, 2010; Rinne et al., 2005; Woods and Alain, 2009). Previous imaging studies using traditional paradigms have, however, been unable to replicate this result in non-human primates, probably because even after laborious training, monkeys have lapses in their auditory attention that alter the AC activation patterns (Rinne et al., 2017). The results of the present thesis indicate that active listening strongly modulates activation in both the human and monkey AC. Importantly, the results of Study III show that when the task is specifically designed for monkeys (i.e., reward cues are used to direct attention to sounds), active listening tasks are associated with similar broad modulations of sound-related activation in the monkey AC as in humans. The results of the present thesis also show that motor responding strongly influences the activation pattern of the human AC. First, motor responding was associated with widespread suppression in the AC. This suppression was seen both when subjects performed an auditory task and when they performed a visual task designed to divert attention from the sounds. Second, auditory-motor integration modulated activation in the AC. This was revealed as stronger AC activation during auditory discrimination tasks with precision grips than power grips in Study I and stronger PT and IPL activation during vowel repetition than vowel production responses in Study II. Importantly, the results of the present thesis suggest that although task, motor and auditory-motor interaction effects all strongly modulate AC activation, they do not interact with each other, and thus are caused by independent neural mechanisms in the AC. The results of the present thesis lay the groundwork for studying the effect of manual and vocal responding in the human AC. Also, the paradigm designed for monkey training could be further developed to study the effect of selective attention in monkeys in a similar manner as they have been studied in humans, helping to bridge the current gap between the extensive neurophysiological literature in animals and fMRI literature in humans regarding the functional architecture of the AC.

### **4.1 ACTIVE LISTENING STRONGLY MODULATES ACTIVATION IN BROAD REGIONS OF HUMAN AND MONKEY AUDITORY CORTEX**

In Study I, subjects performed either an auditory pitch discrimination task (i.e., attended to the sounds) or a visual discrimination task (i.e., ignored the sounds). The activation was stronger in the AC when comparing the auditory

task to the visual task with the same auditory stimuli. This is in line with several previous studies that have shown attention-related modulations in the human AC using similar paradigms (Alho et al., 2014; Grady et al., 1997; Hall et al., 2000; Petkov et al., 2004). It is also important to note that in Study I these attention-related modulations were seen in the AC irrespective of whether subjects overtly responded to targets or not, suggesting that attention-related modulations in the AC are independent of effects related to motor responding (cf. 4.2 and 4.3). Thus, the results of Study I suggest that modulations related to auditory attention in the AC are cognitive in nature and not related to motor demands as such.

In Study III, an auditory-attention paradigm was specifically designed to quickly teach monkeys to perform an auditory task to be used in fMRI. In this paradigm, monkeys performed a very simple auditory task in which they were only required to make a response after a target sound occurred in order to receive a juice reward. Visual reward cues were then used to either motivate the monkey (HiRe cues, large reward) or not (LoRe cues, small reward) to focus its attention on the sounds. The fMRI results in Study III revealed that the reward incentive cues modulated activation in the monkey AC in a similar manner as attention has been shown to modulate AC activation in previous human fMRI studies. That is, activation was stronger in the AC during the HiRe trials in comparison to the LoRe trials. Importantly, this is the first monkey fMRI study to show such broad activation modulations in the AC in relation to active attention-engaging tasks.

As the results of Study III were obtained using a drastically different task, paradigm and a different species (monkeys instead of humans) than traditionally used in fMRI studies on auditory attention, it is important to consider whether the results in Study III were actually related to attention and not to some other uncontrolled effect such as visual stimulation, motor responding or reward expectancy. Although a different visual cue was used in the HiRe than the LoRe trials, the activation difference between the HiRe and LoRe trials in the AC is unlikely to be a visual effect. This is because a significant HiRe vs. LoRe effect was only observed during hit and early-response trials and not during miss trials. All trial types, however, contained exactly the same visual difference between the cues. Motor responding does not explain the results in Study III either, as all HiRe vs. LoRe comparisons were conducted across trials with identical performance (e.g., hit HiRe vs. hit LoRe trials; Figure 9). Further, the results of Studies I and II show that motor responding is associated with decreased activation in the AC (see also 4.2). Thus, if motor responding had affected the HiRe vs. LoRe comparisons in Study III, the effect should have been reverse (i.e., LoRe > HiRe).

Previous studies have shown that reward expectancy modulates activity in the AC (Brosch et al., 2011; Scheich et al., 2007; Weis et al., 2013). It is important to note, however, that comparing the results of previous studies on reward expectancy in the AC to those of the Study III is

complicated by methodological differences between the respective studies. For example, in the paradigm used by Brosch and colleagues (Brosch et al., 2011), reward expectancy was modulated based on the performance of the animals on previous trials. That is, unlike Study III, reward cues were not used in their study. Further, in their study, reward expectancy effects on behavior were subtle in comparison to those of Study III. Other studies, such as that by Weis and colleagues (Weis et al., 2013), have used specific auditory stimulus–reward expectancy associations. In their study, fMRI activation to sounds that predicted upcoming reward was compared to sounds that did not predict reward. By contrast, in Study III the auditory wait signal was associated with both the HiRe and LoRe cues, allowing no auditory stimulus–reward expectancy links to be made.

It is also important to note that reward expectancy is an integral component of all auditory tasks and focused listening, and in fact, the results of most studies using active auditory tasks might be also affected by reward-expectation-related effects (Maunsell, 2004; Peck and Salzman, 2014; Seitz and Dinse, 2007; Seitz and Watanabe, 2005). That is, the specific effects of reward expectancy and auditory tasks are difficult to segregate. Thus, one is left with the conclusion that the AC modulations in Study III were related to either reward expectancy or auditory attention or a conjunction of both. The results of Study III are, however, entirely consistent with the hypothesis that reward cues direct monkeys' attention to the sounds, which causes widespread modulation of AC activation. Furthermore, similar paradigms as the one used in Study III have previously been used in human studies to investigate reward-driven attention in the visual modality (Anderson, 2016; Anderson, 2018; Chelazzi et al., 2013; Della Libera and Chelazzi, 2006; Engelmann and Pessoa, 2007; Engelmann et al., 2009; Hopf et al., 2015; Krebs et al., 2011; Pessoa, 2015). Using such paradigms, it has been found that visual reward cues modulate activation in the visual cortex similarly to how attention modulates visual-cortex activation in more standard visual-attention studies (Engelmann et al., 2009). The results of Study III show that in monkeys visual-reward-cue-driven auditory attention might similarly modulate AC activation in the same manner as in more standard auditory attention studies.

Recently, some authors have raised concerns about whether findings from non-human primate studies can be used in models of human auditory cognition (Schulze et al., 2012; Scott et al., 2012). These concerns are partly related to the fact that while humans easily make long-term associations to sound stimuli, monkeys have been notoriously difficult to teach tasks that demand long-term memory for auditory stimuli. The results from Study III alleviate these concerns in two ways: First, the behavioral results of the AudCue paradigms show that if the auditory task is incentivized, monkeys can learn to memorize that a certain sound correlates with large rewards and another with small rewards, within a few hundred trials. Second, the fMRI

results of Study III suggest that active listening might modulate AC activation similarly in monkeys as in humans.

The findings of Study III open up new possibilities for using monkey studies to develop better models of the human AC. First, subsequent studies could use different types of auditory wait-signal features associated with specific reward cues to manipulate selective attention rather than general attentive listening as in Study III. Second, the auditory-cue paradigm could easily be manipulated into a dichotic-listening task, by presenting the cue sounds in one ear and an irrelevant sound in the other. In a dichotic task, the same sound could then be used in different runs as either a cue sound (i.e., relevant to the task) or as the irrelevant sound, allowing one to study selective attention to a specific sound among other competing irrelevant sounds. Further, in monkeys it is much easier to use fMRI results to guide subsequent neurophysiological recordings. Therefore, using the same task paradigm in monkey fMRI and monkey neurophysiology will likely bridge the gap between the vast neurophysiology-based monkey literature and the fMRI-based human literature. Lastly, although the anatomical and functional organization of the monkey AC has been established, the exact anatomical details of the human AC remain elusive (see for example Rauschecker and Romanski, 2011, for a review). One reason for this has been the difficulty translating animal results to human studies due to differences in research methods, tasks and other study design factors. Thus, the result of Study III—that macro-anatomically similar regions are modulated by attention in monkeys as in humans—will aid in translating monkey findings to refine the anatomical details of models of the human AC.

## **4.2 WIDESPREAD REGIONS OF THE AUDITORY CORTEX ARE SUPPRESSED DURING MOTOR RESPONDING**

Studies I and II revealed extensive activation modulations in the STG during motor responding. In Study I, STG, superior temporal plane and IPL activation decreased both during the auditory and visual tasks when subjects responded to targets using manual grips. Similar decreased activation in the STG has previously been observed during overt vocalizations (Agnew et al., 2013; Curio et al., 2000; Eliades and Wang, 2003; Eliades and Wang, 2017; Flinker et al., 2010; Greenlee et al., 2011; Houde et al., 2002) and during button responses using the N1-suppression paradigm (Aliu et al., 2009; Bässl et al., 2008; Bässl et al., 2009; Bässl et al., 2011; Martikainen et al., 2005; SanMiguel et al., 2013; Timm et al., 2013). Traditionally, such suppression effects have been interpreted to be related to modulatory signals from motor cortices suppressing auditory input because the sounds caused by motor acts are fully predictable (Christoffels et al., 2007; Kauramäki et al., 2010; Reznik et al., 2014). However, in Study I, the motor responses did not cause or

trigger the presentation of the sounds. Furthermore, suppression effects were seen equally during the visual task where subjects ignored the sounds and responded to visual targets. Previous N1-suppression studies (Horváth et al., 2012) and animal studies conducted in freely behaving mice (Nelson et al., 2013; Schneider and Mooney, 2018; Schneider et al., 2014; Schneider et al., 2018) have shown that neurons in the AC are suppressed before, during and after motor execution in situations where motor execution does not cause or is not perceived to cause sounds. Together these previous results and the results of Study I suggest that there is a general mechanism that suppresses auditory processing during movement irrespective of whether the movement is associated with distinct sounds or not. However, although the present motor suppression effects were most parsimoniously explained by similar motor-gating effects as have previously been revealed in freely moving mice and certain EEG studies, this does not mean that the AC could not also be suppressed in a predictive manner by motor responses. For example, a study by Schneider and colleagues (Schneider et al., 2018) recently showed that in mice there is evidence for both general motor-gating mechanisms and motor-predictive suppression of AC neurons. That is, if there is a coupling between the movement and the sound, there is additional suppression of neurons that code for that particular sound (i.e., on top of the general suppression of AC neurons caused by all sorts of movements). This suggests that the specific situations where motor responses cause general vs. predictive suppression should also be investigated in more detail in humans. It is also important to note that the exact timing of the motor suppression effects in Study I cannot be determined because of the poor temporal resolution of fMRI. However, the paradigm used in Study I could easily be converted to be used with methods with higher temporal resolution, such as invasive neurophysiological recordings or M/EEG. However, when converting the present paradigm to EEG, additional motor (control) conditions should be added, because in contrast to fMRI, in EEG, signals from different parts of the brain get intermixed (i.e. the motor control conditions would be used to account for signals originating in the motor cortex).

Unexpectedly, Study II found that activation in extensive regions of the AC and IPL was stronger when subjects responded to targets by (overtly) uttering a vowel (repetition and production responses) than when they performed the same tasks but responded with button presses. These activation modulations were observed in similar regions as the motor suppression effects in Study I. At first glance it may seem that these effects are obviously due to stimulus-dependent activation to the heard sounds in the vocal-response-type blocks. However, stronger activation during vowel-response than button-response blocks was in fact seen throughout the block, even in the beginning of the block before any responses were made. Thus, this effect is unlikely to be due to stimulus-dependent activation to the self-produced vowels. Rather, the most parsimonious explanation is that the activation difference between vocal- and button-response blocks was due to



stronger suppression of activation during manual (button) than vocal responding. Although motor suppression has been reported for both vocal and manual responses (e.g., Agnew et al., 2013; Aliu et al., 2009), it can be argued that vocal responses were not as strongly suppressed as the button responses, because the auditory correlates of the motor responses were relevant when subjects responded using vocal responses but irrelevant and nonexistent when they made button response. However, it is important to note that there was no specific hypothesis in the present thesis regarding the vocal vs. manual responses in Study II, as no previous study has compared vocal and manual responses. Rather, the results of Study II could be used to form specific testable hypotheses in further studies addressing this issue.

### **4.3 AUDITORY-MOTOR INTEGRATION MODULATES ACTIVATION IN THE AUDITORY CORTEX**

In Studies I and II, it was reasoned that associating auditory targets with specific motor responses would induce auditory-motor integration. Therefore, it was hypothesized in Study I that AC activation would depend on whether responses during the auditory task were made using precision or power grips. It was reasoned that such effects should not be seen during the visual task, when the responses were not related to the sound stimuli and thus auditory-motor associations were not important for the task. It should also be noted that to enhance auditory-motor integration effects in Study I, subjects also used slightly different grips (two- or three-finger precision grip and two- or five-finger power grasp) depending on the type of the target (i.e., rising pitch/falling pitch; cf. 3.3.2). Accordingly, in Study I, AC activation was stronger when subjects performed the auditory task using precision grips than when they used power grips. Further, there were no significant differences between the grip types during the visual task when the responses were not associated with the auditory stimuli and therefore auditory-motor integration was presumably not needed. Study I is the first fMRI study to show manual auditory-motor interaction effects in the AC outside the context of music (Zatorre et al., 2007). The finding that grip type modulates AC activation during an active auditory task supports the general idea that there exist selective modulatory connections between the motor and auditory cortices (i.e., connections between a specific sound type and a specific motor program) that are used to monitor and fine-tune auditory processing during motor responding (Hickok, 2010; Hickok, 2016; Hickok and Poeppel, 2007; Hickok et al., 2011; Warren et al., 2005). Furthermore, the result suggests that manual motor behavior may modulate AC activation not only in a general (see previous chapter) but also in a task-specific manner. That is, the behaviorally relevant auditory-motor links formed by the sounds and their corresponding motor responses modulate sound processing in the AC.

In Study II, it was hypothesized that comparing vowel repetition (requiring direct auditory-motor translation) and production responses (no need for auditory-motor translation) would reveal speech-related auditory-motor integration effects in the PT and IPL. Consistently, activation was stronger during repetition- than production-response blocks in left hemisphere regions extending from the mid-STG to IPL. This is in line with previous studies on speech repetition effects in the AC (Carey et al., 2017; Parker Jones et al., 2014; Simmonds et al., 2011; Simmonds et al., 2014a; Simmonds et al., 2014b). However, the hypothesis that repetition of NPh would show stronger auditory-motor effects than repetition of Ph vowels did not receive support. This is, however, in line with the results of a previous study in which STG activation did not significantly differ when subjects repeated words or pseudo words (Parker Jones et al., 2014). It could be argued that comparisons between repetition of phonemic and non-phonemic vowels and between words and pseudowords are not as strongly associated with auditory-motor integration because auditory-to-motor translation is required in all cases. The contrast between repetition and production blocks may be more strongly associated with auditory-to-motor translation as it is required during repetition but not during production responses. However, results from previous studies (e.g., Harinen and Rinne, 2013; Harinen and Rinne, 2014) and Study II show that phonemic and non-phonemic vowels are associated with distinct activation patterns in the AC during both visual and auditory tasks. Thus, although speech-sound information is present in the AC, it does not seem to strongly affect auditory-motor integration in the AC.

#### **4.4 DOES THE HIERARCHICAL STATE FEEDBACK MODEL EXPLAIN THE EFFECTS OF MOTOR RESPONDING IN THE AUDITORY CORTEX?**

Hickok has developed the hierarchical state feedback (HSF) model to explain the dynamics of auditory-motor integration during speech production (Figure 10). This model emphasizes the role of the Sylvian-parietal-temporal (Spt) area in the left hemisphere as an important hub for auditory-motor integration (Hickok, 2012; Hickok, 2016; Hickok et al., 2011). Some of the results of the present thesis are compatible with this model, as the left posterior PT (i.e., approximately the left Spt) was more strongly activated during vowel repetition responses than production responses in Study II.

According to the HSF model (Figure 10), auditory-motor interactions are generated by bidirectional information transfer between the AC and the motor cortex through the Spt. An internal model of a sound is activated in the AC when one intends to produce a sound. The motor execution of the sound generates an efference copy of the motor program, which is sent to the Spt. In the Spt this efference copy is translated into an auditory prediction of what the to-be-produced sound would sound like. This prediction signal is

transferred to the AC to be compared with internal model of the intended sound. If these two differ, then an error signal is sent back to the motor cortices via the Spt to fine-tune the motor output nodes for this particular speech sound. A representation of the produced sound is also compared to the efference copy of the to-be-produced vocalization and used to fine-tune motor commands. According to the model, this feedback loop is particularly important for actions that are novel and non-automatic such as repetition of vocalizations made by other individuals or learning how to produce novel sounds. Based on this model, it was hypothesized in Study II that repeating novel non-phonemic vowels would more strongly activate the left posterior PT (i.e., approximately the Spt) than repeating familiar phonemic vowels. However, in Study II activation did not depend on whether subjects repeated Ph or NPh vowels, in the PT or any other region studied. Thus, in contrast to one of the core assumptions of the HSF model, the difficulty of producing single vowel sounds does not affect auditory-motor translation in the posterior PT. This seems surprising as the HSF model explicitly states that learning new sound-motor correspondences should involve activity specifically in the Spt.

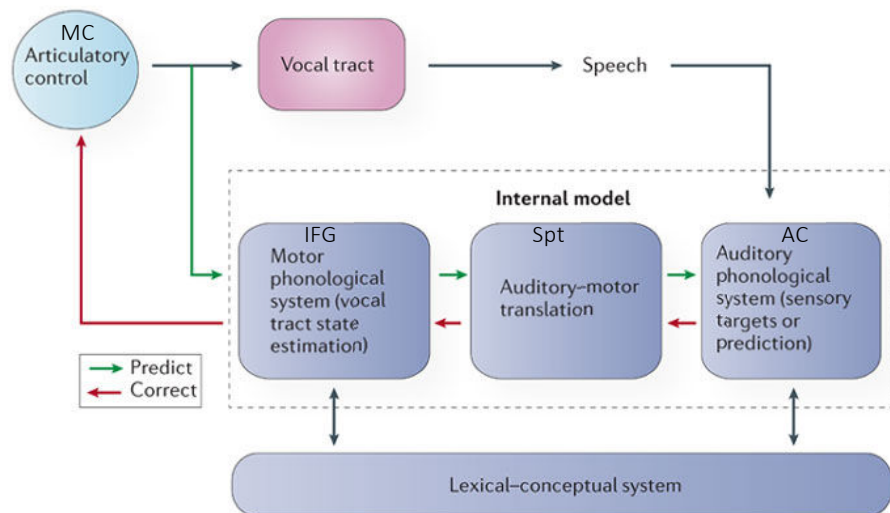


Figure 10 The HSF model postulates that when a sound is produced, the articulatory control units (in the motor cortex, MC) produce an efference copy of the intended sound, which is sent to the motor phonological system in the inferior frontal gyrus (IFG), which generates predictions about the state of the vocal tract, as well as predictions of the sensory consequences of the motor action in the auditory system (AC). Communication between the auditory and motor systems is achieved by an auditory-motor translation system (neuroanatomically the Sylvian-parietal-temporal region; Spt). Figure is reproduced, and modified with permission, from Nature Reviews (2011) Elsevier.

Other findings in the present thesis are also not entirely compatible with the HSF model. As discussed in 4.2, motor responding was found in Study I to suppress AC activation irrespective of whether the responses were made to auditory stimuli or asynchronous visual stimuli while subjects did not focus on the sounds. That is, in Study I widespread suppression was observed in the AC although it is highly improbable that the subjects heard their manual responses, nor did the responses in any other way cause sound stimulation. In contrast, in the HSF model, motor suppression in the AC is modeled as an error signal arising from a mismatch between the sound input (of the sound generated by a motor response such as an utterance) and a prediction of this sound based on a corollary discharge from the motor cortex. Thus, the HSF model does not predict that motor responding would suppress the AC in situations where the responses cause no sounds and have no temporal relationship with the sounds. Finally, the results of Study II are generally consistent with the idea presented in the HSF model that area Spt is more selective for vocal than manual responding. However, this effect was not specific to the Spt, as vocal responding was associated with enhanced activation in wide STG and IPL regions bilaterally. This suggests that the difference between vocal and manual responding is a more general motor effect and that it is not specifically related to auditory-motor integration in the posterior STG as such.

Taken together, not all the results of the present thesis are comprehensively predicted by the HSF model. Firstly, the HSF model suggests that production of novel sounds requires more auditory-motor integration in the Spt. The model does not specify, however what type of information is translated into a motor code in the Spt. For example, is the distinction between native and non-native phonemes (as in Study II) substantial enough to induce stronger auditory-motor-integration effects in the Spt? Further, the predictive coding explanation for motor suppression in the AC suggested in the HSF model appears insufficient. More specifically, the HSF model should be able to explain motor suppression effects in the AC in cases when the motor acts are not associated with a sound. Thus, to be able to comprehensively test predictions based on the HSF model it would be important for the model to specify the exact mechanism behind the auditory-motor translation in the Spt.

#### **4.5 ARE TASK-RELATED AND MOTOR-RESPONSE-RELATED ACTIVATION MODULATIONS INDEPENDENT OF EACH OTHER?**

A key motivation for Studies I and II was to assess whether auditory-attention, task and motor modulations are caused by similar or different mechanisms in the AC. Although all of these effects have been studied extensively in previous studies, these effects have not been examined within

one single study. This is a significant gap in the literature, because task-related factors might modulate or even be the cause of auditory-motor integration effects. Furthermore, motor responding (e.g., to targets) is an integral part of nearly all auditory tasks used in previous auditory-attention fMRI studies (cf. 1.2). The results of Study I showed that auditory-attention-related effects are observed in the AC irrespective of whether subjects overtly respond to targets or not. Also, in Study II, effects related to task and motor responding did not interact in regions of the AC. This lack of interactions between auditory-task- and motor-related effects suggests that in the STG, task and response effects are caused by independent mechanisms. On the other hand, discrimination tasks and n-back memory tasks were associated with different motor-response activation bilaterally in the IPL. What might be the source of such an interaction effect in the IPL? Previous studies have implicated the IPL in working-memory tasks (Gaab et al., 2003; Gaab et al., 2006; Kölsch et al., 2009; Leung and Alain, 2010; Rimmele et al., 2019; Uluc et al., 2018) and categorical processing (Harinen and Rinne, 2014; Husain et al., 2006; Raizada and Poldrack, 2007). Therefore, it could be argued that the interaction of task and motor-response type in the IPL seen in Study II were due to the discrimination task requiring more working memory (which vocalization to make) or categorical processing (maintenance of vowel categories) when subjects made vocal than button responses, whereas in the (categorical) 2-back task working memory task, categorical processing was essential irrespective of motor-response type.

The effects related to motor responding and task also differed in other ways. In previous studies using similar auditory discrimination and n-back tasks as in Study II, it has been observed that task-related modulations in the STG and the IPL seldom occur together. Also, stimulus-level features (pitch, location, etc.) seldom interact strongly with task-dependent activation (Harinen and Rinne, 2013; Harinen and Rinne, 2014; Häkkinen and Rinne, 2018; Häkkinen et al., 2015; Rinne et al., 2009; Rinne et al., 2012; Rinne et al., 2014; Talja et al., 2015;). Based on this it has been suggested that operations in the STG and IPL are linked reciprocally so that both cannot occur at the same time. That is, task manipulations that enhance activation in the STG do not modulate activation in the IPL, and vice versa (e.g., Häkkinen et al., 2015). Consistently, in Study II discrimination tasks activated regions of the STG more strongly than the 2-back tasks while the 2-back tasks activated the IPL more strongly. Further, there were no interactions between stimulus-dependent effects (phonemes and non-phonemes) and task-dependent effects. However, in the present thesis effects related to motor responding were observed both in STG and IPL regions simultaneously. In Study I, motor suppression was observed both in the STG and the IPL, and in Study II, both effects related to motor-response and vocal-response type occurred in regions of the STG and IPL in the same comparisons. Thus, it seems that the motor influences in the present thesis show different dynamics from those related to task processing, strengthening

the argument that effects related to task and motor responding are due to independent neural mechanisms.

Current models of the human AC emphasize that the AC is embedded within functional streams ultimately connecting to the motor cortex (Hickok, 2009; Hickok, 2010; Hickok, 2016; Hickok and Poeppel, 2007; Rauschecker, 2010; Rauschecker and Romanski, 2011; Rauschecker and Scott, 2009). However, as argued above, the previous and present results show that the AC is also strongly modulated by attention and task influences. Furthermore, as these task influences were observed independently of effects related to motor responding in the AC, it seems unlikely that they are caused by feedback projections from streams connecting the AC to the motor cortex. If task influences were not caused by the functional streams connecting the AC to the motor cortex, what might be the origin of such influences? One intriguing possibility is that task influences are related to modulatory connections between basal forebrain regions and the AC. In microstimulation studies conducted in mice it has been shown that neuromodulatory input from the nucleus basalis regulates auditory cortical functions during learning, memory and attention (Kilgard, 2012). Such pathways have, however, been difficult to study in humans, as the basal forebrain is situated in the base of the skull and therefore cannot be non-invasively stimulated in humans. However, the new paradigm developed in SIII opens up the possibility to investigate the role of such neuromodulatory input using microstimulation (Kilgard, 2012) of basal forebrain during auditory task performance in monkeys.

## **4.6 IMPLICATIONS FOR THEORETICAL MODELS OF THE HUMAN AUDITORY CORTEX**

The present thesis provides novel results that will help to refine models of the human AC. These models are currently strongly based on (invasive) neuronal-level studies conducted in non-human primates during passive presentation of sounds (Rauschecker, 2010; Rauschecker and Romanski, 2011; Rauschecker and Scott, 2009). Because of this emphasis, current models of the AC mainly account for the processing of different sound features provided by ascending (thalamic) pathways to the AC. However, it is known that the majority of the input to the primary AC comes from other cortical areas and not via the ascending inputs from the thalamus (Scheich et al., 2007). Consistently, in the current thesis, it was shown that sound processing in the AC is modulated by factors such as attention, task, manual-motor responding, vocal-motor responding and reward-related influences, all of which are unlikely to stem from the ascending thalamic pathway. Furthermore, these strong modulations were observed independent of each other and also independent of stimulus level effects in the AC. Thus, the results of the present thesis support the conclusion that it is probably misleading to view the primary AC as a low-level sound analyzer, while

higher (more cognitive) functions and operations occur in other cortical areas outside the AC. Rather, the function of the AC should be seen as embedded within strong modulatory influences from different brain regions outside the ascending auditory pathway, which independently shape the function of the AC (see for example Scheich et al., 2007 and Weinberger, 2011).

## 5 CONCLUSIONS

The results of the present thesis are consistent with influential theoretical models that highlight auditory-motor links as integral to the functional organization of the AC (Hickock and Poeppel, 2007; Rauschecker and Romanski, 2011; Rauschecker and Scott, 2009; Zatorre et al., 2007). That is, auditory-motor integration was found to strongly influence activation in the human AC during both manual and vocal responding: (1) Activity during auditory tasks (but not visual tasks with identical stimuli) was stronger in the AC when auditory targets were responded to using precision (vs. power) grips, and (2) activation was stronger in the left posterior PT during vowel repetition demanding auditory-motor translation processes as compared with production responses. The results of the present thesis also indicate that motor responding strongly suppresses AC activation when there is no behaviorally relevant auditory-motor link. These motor suppression effects are consistent with the view that there is a general motor-gating mechanism that suppresses auditory processing during movement. Such mechanisms are, however, currently inadequately accounted for in theoretical models of motor influences in the AC that currently focus strongly on auditory-motor integration. The results of all three studies of the present thesis support the notion that task influences are integral to the function of the AC in both humans and non-human primates (Scheich et al., 2007; Weinberger, 2011). Importantly, the results of the present thesis show that such task influences and demands related to motor responding and auditory-motor integration do not interact in the AC. Thus, these effects seem to be caused by independent mechanisms in the AC. Importantly, the results of Study III imply that a substantial portion of the human auditory neurocognitive system may be evolutionarily conserved. This has been difficult to demonstrate in previous studies because it is difficult to use the same research methods in monkeys and humans. Thus, the present findings help bridge the gap between the extensive neurophysiological literature on auditory cortical processing in monkeys and the results of auditory fMRI studies using attention-engaging tasks in humans.



## REFERENCES

Agnew, Z.K., McGettigan, C., Banks, B., Scott, S.K., 2013. Articulatory movements modulate auditory responses to speech. *Neuroimage* 73, 191-199.

Alho, K., Rinne, T., Herron, T.J., 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.

Aliu, S.O., Houde, J.F., Nagarajan, S.S., 2009. Motor-induced suppression of the auditory cortex. *Journal of cognitive neuroscience* 21, 791-802.

Anderson, B.A., 2016. Value-driven attentional capture in the auditory domain. *Attention Perception and Psychophysics* 78, 242-250.

Anderson, B.A., 2018. Neurobiology of value-driven attention. *Current opinion in psychology* 29, 27-33

Atiani, S., David, S.V., Elgueda, D., Locastro, M., Radtke-Schuller, S., Shamma, S.A., Fritz, J.B., 2014. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82, 486-499.

Atiani, S., Elhilali, M., David, S.V., Fritz, J.B., Shamma, S.A., 2009. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61, 467-480.

Baldo, J.V., Klostermann, E.C., Dronkers, N.F., 2008. It's either a cook or a baker: Patients with conduction aphasia get the gist but lose the trace. *Brain and Language* 105, 134-140.

Bagur, S., Averseng, M., Elgueda, D., David, S., Fritz, J., Yin, P., Shamma, S., Boubenec, Y., Ostojic, S., 2018. Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex. *Nature Communications* 9, 2529-2540.

Baumann, S., Griffiths, T.D., Rees, A., Hunter, D., Sun, L., Thiele, A., 2010. Characterisation of the BOLD response time course at different levels of the auditory pathway in non-human primates. *Neuroimage* 50, 1099-1108.

Baumann, S., Köneke, S., Schmidt, C.F., Meyer, M., Lutz, K., Jancke, L., 2007. A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research* 1161, 65-78.

Berlot, E., Formisano, E., De Martino, F., 2018. Mapping frequency-specific tone predictions in the human auditory cortex at high spatial resolution. *Journal of Neuroscience* 38, 4934-4942.

Briggs, F., Mangun, G.R., Usrey, W.M., 2013. Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits. *Nature* 499, 476-480.  
Brosch, M., Selezneva, E., Scheich, H., 2011. Representation of reward feedback in primate auditory cortex. *Frontiers in Systems Neuroscience* 5, 5.

Buchsbaum, B.R., Hickok, G., Humphries, C., 2001. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science* 25, 663-678.

Buchsbaum, B.R., Baldo, J., Okada, K., Berman, K.F., Dronkers, N., D'Esposito, M., Hickok, G., 2011. Conduction aphasia, sensory-motor integration, and phonological short-term memory – An aggregate analysis of lesion and fMRI data. *Brain and Language* 119, 119-128.

Buckley, M.J., Mansouri, F.A., Hoda, H., Mahboubi, M., Browning, P.G.F., Kwok, S.C., Phillips, A., Tanaka, K., 2009. Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science* 325, 52-58.

Burnett, T.A., Freedland, M.B., Larson, C.R., Hain, T.C., 1998. Voice F0 responses to manipulations in pitch feedback. *Journal of the Acoustical Society of America* 103, 3153-3161.

Bäss, P., Horvath, J., Jacobsen, T., Schröger, E., 2011. Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology* 48, 1276-1283.

Bäss, P., Jacobsen, T., Schröger, E., 2008. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology* 70, 137-143.

Bäss, P., Widmann, A., Roye, A., Schröger, E., Jacobsen, T., 2009. Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *European Journal of Neuroscience* 29, 1514-1521.

Carey, D., Miquel, M.E., Evans, B.G., Adank, P., McGettigan, C., 2017. Vocal tract images reveal neural representations of sensorimotor transformation during speech imitation. *Cerebral Cortex* 27, 3064-3079.

Chelazzi, L., Perlato, A., Santandrea, E., Della Libera, C., 2013. Rewards teach visual selective attention. *Vision Research* 85, 58-72.

Chen, J.L., Penhune, V.B., Zatorre, R.J., 2008a. Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex* 18, 2844-2854.

Chen, J.L., Penhune, V.B., Zatorre, R.J., 2008b. Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience* 20, 226-239.

Chen, J.L., Penhune, V.B., Zatorre, R.J., 2009. The role of auditory and premotor cortex in sensorimotor transformations. *Annals of New York Academy of Science* 1169, 15-34.

Chen, J.L., Zatorre, R.J., Penhune, V.B., 2006. Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage* 32, 1771-1781.

Christoffels, I.K., Formisano, E., Schiller, N.O., 2007. Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Human brain mapping* 28, 868-879.

Curio, G., Neuloh, G., Numminen, J., Jousmäki, V., Hari, R., 2000. Speaking modifies voice-evoked activity in the human auditory cortex. *Human brain mapping* 9, 183-191.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis - I. Segmentation and surface reconstruction. *Neuroimage* 9, 179-194.

De Martino, F., Moerel, M., Ugurbil, K., Goebel, R., Yacoub, E., Formisano, E., 2015. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. *Proceedings of the National Academy of Sciences* 112, 16036-16041.

Della Libera, C., Chelazzi, L., 2006. Visual selective attention and the effects of monetary rewards. *Psychological Science* 17, 222-227.

Downer, J.D., Niwa, M., Sutter, M.L., 2015. Task engagement selectively modulates neural correlations in primary auditory cortex. *Journal of Neuroscience* 35, 7565-7574.

Ehrsson, H.H., Fagergren, A., Jonsson, T., Westling, G., Johansson, R.S., Forssberg, H., 2000. Cortical activity in precision- versus power-grip tasks: An fMRI study. *Journal of Neurophysiology* 83, 528-536.

Eliades, S.J., Wang, X.Q., 2003. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology* 89, 2194-2207.

Eliades, S.J., Wang, X.Q., 2017. Contributions of sensory tuning to auditory-vocal interactions in marmoset auditory cortex. *Hearing research* 348, 98-111.

Engelmann, J.B., Damaraju, E., Padmala, S., Pessoa, L., 2009. Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Frontiers in human neuroscience* 3, 4.

Engelmann, J.B., Pessoa, L., 2007. Motivation sharpens exogenous spatial attention. *Emotion* 7, 668-674.

Flinker, A., Chang, E.F., Kirsch, H.E., Barbaro, N.M., Crone, N.E., Knight, R.T., 2010. Single-trial speech suppression of auditory cortex activity in humans. *The Journal of neuroscience* 30, 16643-16650.

Formisano, E., Moerel, M., Bonte, M., 2015. Functional MRI of the auditory cortex. *Fmri: from Nuclear Spins to Brain Functions* 30, 473-507.

Francis, N.A., Winkowski, D.E., Sheikhattar, A., Armengol, K., Babadi, B., Kanold, P.O., 2018. Small networks encode decision-making in primary Auditory Cortex. *Neuron* 97, 885-870.

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007a. Auditory attention—focusing the searchlight on sound. *Current Opinion in Neurobiology* 17, 1-19.

Fritz, J., Elhilali, M., Shamma, S.A., 2005a. Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing research* 206, 159-176.

Fritz, J.B., Elhilali, M., Shamma, S.A., 2007b. Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of Neurophysiology* 98, 2337-2346.

Fritz, J., Mishkin, M., Saunders, R.C., 2005b. In search of an auditory engram. *Proceedings of the National Academy of Sciences* 102, 9359-9364.

Fritz, J., Shamma, S., Elhilali, M., Klein, D., 2003. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience* 6, 1216-1223.

Gaab, N., Gaser, C., Schlaug, G., 2006. Improvement-related functional plasticity following pitch memory training. *Neuroimage* 31, 255-263.

Gaab, N., Gaser, C., Zöhle, T., Jäncke, L., Schlaug, G., 2003. Functional anatomy of pitch memory--an fMRI study with sparse temporal sampling. *Neuroimage* 19, 1417-1426.

Grady, C.L., Van Meter, J.W., Maisog, J.M., Pietrini, P., Krasuski, J., Rauschecker, J.P., 1997. Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport* 8, 2511-2516.

Greenlee, J.D.W., Jackson, A.W., Chen, F., Larson, C.R., Oya, H., Kawasaki, H., Chen, H., Howard, M.A., III, 2011. Human auditory cortical activation during self-vocalization. *PloS one* 6, e14744.

Grézes, J., Tucker, M., Armony, J., Ellis, R., Passingham, R.E., 2003. Objects automatically potentiate action: an fMRI study of implicit processing. *European Journal of Neuroscience* 17, 2735-2740.

Hackett, T.A., Preuss, T.M., Kaas, J.H., 2001. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *Journal of Comparative Neurology* 441, 197-222.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Summerfield, A.Q., Palmer, A.R., Elliott, M.R., Bowtell, R.W., 2000. Modulation and task effects in auditory processing measured using fMRI. *Human brain mapping* 10, 107-119.

Harinen, K., Rinne, T., 2013. Activations of human auditory cortex to phonemic and nonphonemic vowels during discrimination and memory tasks. *Neuroimage* 77, 279-287.

Harinen, K., Rinne, T., 2014. Acoustical and categorical tasks differently modulate activations of human auditory cortex to vowels. *Brain and Language* 138, 71-79.

Hickok, G., 2009. The functional neuroanatomy of language. *Physics of Life Reviews* 6, 121-143.

Hickok, G., 2010. The Functional Anatomy of Speech Processing: From auditory cortex to speech recognition and speech production. In: Ulmer S., Jansen O. (eds) *fMRI*. Springer, Berlin, Heidelberg.

Hickok, G., 2012. Computational neuroanatomy of speech production. *Nature Reviews Neuroscience* 13, 135-145.

Hickok, G., 2016. A cortical circuit for voluntary laryngeal control: Implications for the evolution language. *Psychonomic Bulletin and Review* 24, 55-63.

Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience* 15, 673-682.

Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407-422.

Hickok, G., Okada, K., Serences, J.T., 2009. Area spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology* 101, 2725-2732.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393-402.

Hopf, J.M., Schönfeld, M.A., Buschschulte, A., Rautzenberg, A., Krebs, R.M., Böhler, C.N., 2015. The modulatory impact of reward and attention on global feature selection in human visual cortex. *Visual Cognition* 23, 229-248.

Horváth, J., Mäss, B., Bäss, P., Toth, A., 2012. Action-sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience* 24, 1919-1931.

Hoshi, E., Shima, K., Tanji, J., 2000. Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *Journal of Neurophysiology* 83, 2355-2373.

Houde, J.F., Nagarajan, S.S., Sekihara, K., Merzenich, M.M., 2002. Modulation of the auditory cortex during speech: An MEG study. *Journal of cognitive neuroscience* 14, 1125-1138.

Häkkinen, S., Ovaska, N., Rinne, T., 2015. Processing of pitch and location in human auditory cortex during visual and auditory tasks. *Frontiers in Psychology* 6, 6.

Häkkinen, S., Rinne, T., 2018. Intrinsic, stimulus-driven and task-dependent connectivity in human auditory cortex. *Brain Struct Funct*, 223, 2113-2127.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. *Fsl. Neuroimage* 62, 782-790.

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

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences* 97, 11793.

Kauramäki, J., Jääskeläinen, I.P., Hari, R., Möttönen, R., Rauschecker, J.P., Sams, M., 2010. Lipreading and covert speech production similarly modulate human auditory-cortex responses to pure tones. *Journal of Neuroscience* 30, 1314-1321.

Kilgard, M.P., 2012. Harnessing plasticity to understand learning and treat disease. *Trends in Neurosciences* 35, 715-722.

Krebs, R.M., Böehler, C.N., Egner, T., Woldorff, M.G., 2011. The neural underpinnings of how reward associations can both guide and misguide attention. *Journal of Neuroscience* 31, 9752-9759.

Kölsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., Gruber, O., 2009. Functional architecture of verbal and tonal working memory: an fMRI study. *Human brain mapping* 30, 859-873.

Lakatos, P., Barczak, A., Neymotin, S.A., McGinnis, T., Ross, D., Javitt, D.C., O'Connell, M.N., 2016. Global dynamics of selective attention and its lapses in primary auditory cortex. *Nature Neuroscience* 19, 1707-1717.

Leung, A.W.S., Alain, C., 2010. Working memory load modulates the auditory 'What' and 'Where' neural networks. *Neuroimage* 55, 1260-1269.

Liu, T., Pestilli, F., Carrasco, M., 2005. Transient attention enhances perceptual performance and fMRI response in human visual cortex. *Neuron* 45, 469-477.

Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. *Nature* 453, 869-878.

Loose, R., Kaufmann, C., Auer, D.P., Lange, K.W., 2003. Human prefrontal and sensory cortical activity during divided attention tasks. *Human brain mapping* 18, 249-259.

Makris, S., Grant, S., Hadar, A.A., Yarrow, K., 2013. Binocular vision enhances a rapidly evolving affordance priming effect: Behavioural and TMS evidence. *Brain and cognition* 83, 279-287.

Martikainen, M.H., Kaneko, K., Hari, R., 2005. Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex* 15, 299-302.

Maunsell, J.H.R., 2004. Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences* 8, 261-265.

McLaren, D.G., Kosmatka, K.J., Oakes, T.R., Krönke, C.D., Kohama, S.G., Matochik, J.A., Ingram, D.K., Johnson, S.C., 2009. A population-average MRI-based atlas collection of the rhesus macaque. *Neuroimage* 45, 52-59.

Minamimoto, T., La Camera, G., Richmond, B.J., 2009. Measuring and modeling the interaction among reward size, delay to reward, and satiation level on motivation in monkeys. *Journal of Neurophysiology* 101, 437-447.

Minamimoto, T., Saunders, R.C., Richmond, B.J., 2010. Monkeys quickly learn and generalize visual categories without lateral prefrontal cortex. *Neuron* 66, 501-507.

Moerel, M., De Martino, F., Formisano, E., 2014. An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience* 8, 225.

Nelson, A., Schneider, D.M., Takatoh, J., Sakurai, K., Wang, F., Mooney, R., 2013. A circuit for motor cortical modulation of auditory cortical activity. *Journal of Neuroscience* 33, 14342-14356.

Niwa, M., O'Connor, K.N., Engall, E., Johnson, J.S., Sutter, M.L., 2015. Hierarchical effects of task engagement on amplitude modulation encoding in auditory cortex. *Journal of Neurophysiology* 113, 307-327.

Northam, G.B., Adler, S., Eschmann, K.C.J., Chong, W.K., Cowan, F.M., Baldeweg, T., 2018. Developmental conduction aphasia after neonatal stroke. *Annals of Neurology* 83, 664-675.

Pa, J., Hickok, G., 2008. A parietal-temporal sensory-motor integration area for the human vocal tract: Evidence from an fMRI study of skilled musicians. *Neuropsychologia* 46, 362-368.

Parker Jones, O.P., Prejawa, S., Hope, T.M.H., Oberhuber, M., Seghier, M.L., Leff, A.P., Green, D.W., Price, C.J., 2014. Sensory-to-motor integration



during auditory repetition: a combined fMRI and lesion study. *Frontiers in human neuroscience* 8, 24.

Patel, G.H., Yang, D., Jamerson, E.C., Snyder, L.H., Corbetta, M., Ferrera, V.P., 2015. Functional evolution of new and expanded attention networks in humans. *Proceedings of the National Academy of Sciences* 112, 5377-5377.

Peck, C.J., Salzman, C.D., 2014. The amygdala and basal forebrain as a pathway for motivationally guided attention. *Journal of Neuroscience* 34, 13757-13767.

Peschke, C., Ziegler, W., Eisenberger, J., Baumgörtner, A., 2012. Phonological manipulation between speech perception and production activates a parieto-frontal circuit. *Neuroimage* 59, 788-799.

Peschke, C., Ziegler, W., Kappes, J., Baumgörtner, A., 2009. Auditory-motor integration during fast repetition: The neuronal correlates of shadowing. *Neuroimage* 47, 392-402.

Pessoa, L., 2015. Multiple influences of reward on perception and attention. *Visual Cognition* 23, 272-290.

Petkov, C.I., Kang, X., Alho, K., Bertrand, O., Yund, E.W., Woods, D.L., 2004. Attentional modulation of human auditory cortex. *Nature Neuroscience* 7, 658-663.

Petkov, C.I., Kikuchi, Y., Milne, A.E., Mishkin, M., Rauschecker, J.P., Logothetis, N.K., 2015. Different forms of effective connectivity in primate frontotemporal pathways. *Nature Communications* 6.

Purcell, D.W., Munhall, K.G., 2006. Compensation following real-time manipulation of formants in isolated vowels. *Journal of the Acoustical Society of America* 119, 2288-2297.

Rauschecker, J.P., 2010. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing research* 1-2, 16-25.

Rauschecker, J.P., Romanski, L.M., 2011. Auditory cortical organization: evidence for functional streams. In: Winer J., Schreiner C. (eds) *The auditory cortex*. Springer, Boston, MA.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience* 12, 718-724.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences* 97, 11800-11806.

Rauschecker, J.P., Tian, B., Hauser, M., 1995. Processing of complex sounds in the macaque nonprimary auditory-cortex. *Science* 268, 111-114.

Recanzone, G.H., Cohen, Y.E., 2010. Serial and parallel processing in the primate auditory cortex revisited. *Behavioural brain research* 206, 1-7.

Reynolds, J.H., Heeger, D.J., 2009. The normalization model of attention. *Neuron* 61, 168-185.

Reznik, D., Henkin, Y., Schadel, N., Mukamel, R., 2014. Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nature Communications* 5, 4059.

Riecke, L., Peters, J.C., Valente, G., Poser, B.A., Kemper, V.G., Formisano, E., Sorger, B., 2018. Frequency-specific attentional modulation in human primary auditory cortex and midbrain. *Neuroimage* 174, 274-287.

Rimmele, J.M., Gudi-Mindermann, H., Nolte, G., Roeder, B., Engel, A.K., 2019. Working memory training integrates visual cortex into beta-band networks in congenitally blind individuals. *Neuroimage* 194, 259-271.

Rinne, T., 2010. Activations of human auditory cortex during visual and auditory selective attention tasks with varying difficulty. *Open Neuroimaging Journal* 4, 187-193.

Rinne, T., Ala-Salomaki, H., Stecker, G.C., Pätynen, J., Lokki, T., 2014. Processing of spatial sounds in human auditory cortex during visual, discrimination and 2-back tasks. *Frontiers of Neurosciences* 8, 220.

Rinne, T., Koistinen, S., Salonen, O., Alho, K., 2009. Task-dependent activations of human auditory cortex during pitch discrimination and pitch memory tasks. *The Journal of Neuroscience* 29, 13338-13343.

Rinne, T., Koistinen, S., Talja, S., Wikman, P., Salonen, O., 2012. Task-dependent activations of human auditory cortex during spatial discrimination and spatial memory tasks. *Neuroimage* 5, 4126-4131.

Rinne, T., Muers, R.S., Salo, E., Slater, H., 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* 36, 1-14.

Rinne, T., Pekkola, J., Degerman, A., Autti, T., Jääskeläinen, I.P., Sams, M., Alho, K., 2005. Modulation of auditory cortex activation by sound presentation rate and attention. *Human brain mapping* 26, 94-99.

Rivier, F., 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.

Rogalsky, C., Poppa, T., Chen, K.H., Anderson, S.W., Damasio, H., Love, T., Hickok, G., 2015. Speech repetition as a window on the neurobiology of auditory-motor integration for speech: A voxel-based lesion symptom mapping study. *Neuropsychologia* 71, 18-27.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience* 2, 1131-1136.

Rushworth, M.F.S., Nixon, P.D., Passingham, R.E., 1997. Parietal cortex and movement .1. Movement selection and reaching. *Experimental brain research* 117, 292-310.

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., Alho, K., 2009. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain research* 1286, 155-164.

Sanmiguel, I., Todd, J., Schröger, E., 2013. Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology* 50, 334-343.

Santangelo, V., Fagioli, S., Macaluso, E., 2010. The costs of monitoring simultaneously two sensory modalities decrease when dividing attention in space. *Neuroimage* 49, 2717-2727.

Schafer, E.W.P., Marcus, M.M., 1973. Self-stimulation alters human sensory brain responses. *Science* 181, 175-177.

Scheich, H., Brechmann, A., Brosch, M., Budinger, E., Ohl, F.W., 2007. The cognitive auditory cortex: task-specificity of stimulus representations. *Hearing research* 229, 213-224.

Schneider, D.M., Mooney, R., 2018. How movement modulates hearing. *Annual Review of Neuroscience* 41, 553-572.

Schneider, D.M., Nelson, A., Mooney, R., 2014. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513, 189–194.

Schneider, D.M., Sundararajan, J., Mooney, R., 2018. A cortical filter that learns to suppress the acoustic consequences of movement. *Nature* 561, 391–395.

Schröger, E., Marzecova, A., SanMiguel, I., 2015. Attention and prediction in human audition: a lesson from cognitive psychophysiology. *European Journal of Neuroscience* 41, 641–664.

Schulze, K., Vargha-Khadem, F., Mishkin, M., 2012. Test of a motor theory of long-term auditory memory. *Proceedings of the National Academy of Sciences* 109, 7121–7125.

Scott, B.H., Mishkin, M., Yin, P.B., 2012. Monkeys have a limited form of short-term memory in audition. *Proceedings of the National Academy of Sciences* 109, 12237–12241.

Seitz, A.R., Dinse, H.R., 2007. A common framework for perceptual learning. *Current Opinion in Neurobiology* 17, 148–153.

Seitz, A.R., Watanabe, T., 2005. A unified model for perceptual learning. *Trends in Cognitive Sciences* 9, 329–334.

Shuster, L.I., Lemieux, S.K., 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language* 93, 20–31.

Simmonds, A.J., Wise, R.J.S., Dhanjal, N.S., Leech, R., 2011. A comparison of sensory-motor activity during speech in first and second languages. *Journal of Neurophysiology* 106, 470–478.

Simmonds, A.J., Leech, R., Collins, C., Redjep, O., Wise, R.J.S., 2014a. Sensory-motor integration during speech production localizes to both left and right plana temporale. *Journal of Neuroscience* 34, 12963–12972.

Simmonds, A.J., Wise, R.J.S., Collins, C., Redjep, O., Sharp, D.J., Iverson, P., Leech, R., 2014b. Parallel systems in the control of speech. *Human brain mapping* 35, 1930–1943.

Slater, H., Milne, A.E., Wilson, B., Muers, R.S., Balezeau, F., Hunter, D., Thiele, A., Griffiths, T.D., Petkov, C.I., 2016. Individually customisable non-invasive head immobilisation system for non-human primates with an option for voluntary engagement. *Journal of neuroscience methods* 269, 46–60.

Symes, E., Tucker, M., Ellis, R., Vainio, L., Ottoboni, G., 2008. Grasp preparation improves change detection for congruent objects. *Journal of Experimental Psychology-Human Perception and Performance* 34, 854-871.

Talja, S., Alho, K., Rinne, T., 2015. Source analysis of event-related potentials during pitch discrimination and pitch memory tasks. *Brain Topography* 28, 445-458.

Tian, B., Reser, D., Durham, A., Kustov, A., Rauschecker, J.P., 2001. Functional specialization in rhesus monkey auditory cortex. *Science* 292, 290-293.

Timm, J., SanMiguel, I., Keil, J., Schröger, E., Schönwiesner, M., 2014. Motor intention determines sensory attenuation of brain responses to self-initiated sounds. *Journal of cognitive neuroscience* 26, 1481-1489.

Timm, J., SanMiguel, I., Saupe, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. *Bmc Neuroscience* 14, 2.

Tourville, J.A., Reilly, K.J., Guenther, F.H., 2008. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39, 1429-1443.

Tucker, M., Ellis, R., 2001. The potentiation of grasp types during visual object categorization. *Visual Cognition* 8, 769-800.

Vainio, L., Tiainen, M., Tiippana, K., Vainio, M., 2014. Shared processing of planning articulatory gestures and grasping. *Experimental brain research* 232, 2359-2368.

Uluc, I., Schmidt, T.T., Wu, Y.H., Blankenburg, F., 2018. Content-specific codes of parametric auditory working memory in humans. *Neuroimage* 183, 254-262.

Warren, J.E., Wise, R.J., Warren, J.D., 2005. Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends in Neuroscience* 28, 636-643.

Weinberger, N.M., 2011. Reconceptualizing the primary auditory cortex: Learning, memory and specific plasticity. In: Winer J., Schreiner C. (eds) *The auditory cortex*. Springer, Boston, MA.

Weis, T., Brechmann, A., Puschmann, S., Thiel, C.M., 2013. Feedback that confirms reward expectation triggers auditory cortex activity. *Journal of Neurophysiology* 110, 1860-1868.

Wessinger, C.M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., Rauschecker, J.P., 2001. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience* 13, 1-7.

Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. *Neuroimage* 92, 381-397.

Wise, R.J.S., Scott, S.K., Blank, S.C., Mummery, C.J., Murphy, K., Warburton, E.A., 2001. Separate neural subsystems within 'Wernicke's area'. *Brain* 124, 83-95.

Woods, D.L., Alain, C., 2009. Functional imaging of human auditory cortex. *Current opinion in otolaryngology and head and neck surgery* 17, 407-411.

Woods, D.L., Herron, T.J., Cate, A.D., Yund, E.W., Stecker, G.C., Rinne, T., Kang, X., 2010. Functional properties of human auditory cortical fields. *Frontiers in Systems Neuroscience* 4, 155.

Woods, D.L., Stecker, G.C., Rinne, T., Herron, T.J., Cate, A.D., Yund, E.W., Liao, I., Kang, X., 2009. Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS one* 4, e5183.

Zatorre, R.J., Chen, J.L., Penhune, V.B., 2007. When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience* 8, 547-558.

## ORIGINAL PUBLICATIONS