Functional MRI of the human central auditory system

Citation for published version (APA):

Langers, D. R. M. (2005). *Functional MRI of the human central auditory system*. [Doctoral Thesis, Maastricht University]. Dave Langers. <https://doi.org/10.26481/dis.20060421dl>

Document status and date:

Published: 01/01/2005

DOI:

[10.26481/dis.20060421dl](https://doi.org/10.26481/dis.20060421dl)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

FUNCTIONAL MRI OF THE HUMAN CENTRAL AUDITORY SYSTEM

Copyright © 2001-2005 Dave Langers, all rights reserved

Permission is granted to anyone to make or distribute verbatim copies of this document for non-commercial purposes, in any medium, provided that this copyright notice and permission notice are preserved, and that the distributor grants the recipient permission for further redistribution as permitted by this notice.

ISBN-10 90-9019722-2
ISBN-13 978-90-9019722-7
NUR 954

Published by

Dave Langers, <http://audio-fnri.langers.nl>

Printed by

Olsztyńskie Zakłady Graficzne, <http://www.ozgraf.com.pl>
Bridge, <http://www.bridge-int.nl>

First edition

10 9 8 7 6 5 4 3 2 1

Functionele MRI van het Humane Centrale Auditieve Systeem

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus
Prof. Mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 21 april 2006 om 14:00 uur

door

Davey René Margriet Langers
Geboren 9 juli 1973 te Stramproy

Promotiecommissie

Promotores

Prof. dr. J.M.A. van Engelshoven

Prof. dr. J.J. Manni

Prof. dr. P. van Dijk (Rijksuniversiteit Groningen)

Copromotor

Dr. ir. W.H. Backes

Beoordelingscommissie

Prof. dr. H. Kingma, *voorzitter*

Prof. dr. M.A. van Buchem (Universiteit Leiden)

Prof. dr. R. Goebel

Prof. dr. K. Nicolay (Technische Universiteit Eindhoven)

‘The brain is the citadel of sense perception.’

~ *Pliny (the elder)* ~

‘Know how to listen, and you will profit even from those who talk badly.’

~ *Plutarch* ~

Contents

Titelpagina *	iii
Promotiecommissie *	iv
Contents	vii
Summary *	ix
1 Introduction.....	1
§1.1 About this Thesis	2
§1.2 The Auditory System	4
§1.3 Functional Magnetic Resonance Imaging	16
2 Hemodynamic Response Interactions	29
§2.1 Introduction	31
§2.2 Materials and Methods	33
§2.3 Results	39
§2.4 Discussion	45
§2.5 Conclusion	52
§2.6 Appendices	52
3 Intensity, Loudness, and the Linearity of Activation	57
§3.1 Introduction	59
§3.2 Materials and Methods	60
§3.3 Results	65
§3.4 Discussion	70
§3.5 Conclusion	75
4 Sensitivity to Spectro-Temporal Modulations.....	77
§4.1 Introduction	79
§4.2 Materials and Methods	81
§4.3 Results	87
§4.4 Discussion	92
§4.5 Conclusion	96
5 Tonotopy and Lateralization in the Cortex.....	97
§5.1 Introduction	99
§5.2 Materials and Methods	100
§5.3 Results	104
§5.4 Discussion	111
§5.5 Conclusion	119
6 Lateralization, Connectivity and Plasticity	121
§6.1 Introduction	123

§6.2 Materials and Methods	126
§6.3 Results	131
§6.4 Discussion	139
§6.5 Conclusion	142
7 General Discussion and Conclusions	143
§7.1 Functional Magnetic Resonance Imaging	145
§7.2 The Auditory System	150
8 Annex.....	157
§8.1 References	158
§8.2 Curriculum Vitae *	172
§8.3 Financial Support	174
§8.4 Postface *	175

- * : The sections marked with an asterisk are (also) provided in Dutch.
De paragrafen met een asterisk zijn (eveneens) in het Nederlands.

Summary

■ Compared to the visual and somatosensory systems, relatively little is known about the functioning of the central auditory system (CAS) in humans. This is in part related to the fact that although functional magnetic resonance imaging (fMRI) is a non-invasive technique with a relatively high spatial resolution, at the same time it is accompanied by intense acoustic scanner noise that tends to interfere with the experiment.

In this thesis a number of studies are presented that investigate the functioning of the human CAS by means of fMRI. They focus on the evaluation of potential improvements on the employed experimental setup, the determination of the functional organization of the CAS under normal circumstances, and the investigation of possible deviations that may occur in subjects with various kinds of hearing disorders.

Chapter 1 gives an outline of this thesis and contains a general introduction into the fields of the auditory system and fMRI.

Chapter 2 describes the influence of timing parameters in a so-called 'sparse' acquisition paradigm upon the measurable hemodynamic response in the brain. Sparse paradigms can be characterized by long silent periods of scanner inactivity between the successively performed

■ Vergeleken met zintuigsystemen zoals het visuele (gezichtsvermogen) en het somatosensorische (tast) is er over het centrale functioneren van het auditieve systeem (gehoor) in de mens nog vrij weinig bekend. Dit heeft er onder andere mee te maken dat functionele beeldvorming met magnetische kernresonantie (functional magnetic resonance imaging, fMRI) weliswaar een niet-invasieve meettechniek is met een relatief hoog ruimtelijk oplosend vermogen, maar tevens gepaard gaat met de productie van zeer luide achtergrondgeluiden door de scanner die interfereren met het experiment.

In dit proefschrift worden een aantal studies beschreven die nader ingaan op het functioneren van het centrale auditieve systeem (CAS) middels fMRI. Het accent ligt hierbij op het onderzoeken van mogelijke verbeteringen in de gebruikte experimentele opzet, het bestuderen van de functionele organisatie van het CAS onder normale omstandigheden, en het onderzoeken van de verschillen die mogelijkwerwijs optreden in individuen met gehoorafwijkingen.

Hoofdstuk 1 beschrijft de opzet van dit proefschrift en bevat een algemene inleiding op het gebied van het auditieve systeem en fMRI.

Hoofdstuk 2 beschrijft de invloed van de tijdsinstellingen in een ge-

short acquisitions. In this chapter, it is shown that the influence of a single previous scan may reduce the detectable signals by up to a factor two, which demonstrates the presence of non-linear interaction effects. A simple model is described that is able to describe the main findings in a qualitative sense, although more complicated response behavior is found for sound stimuli that resemble the scanner noise. From these results, a number of recommendations can be derived regarding optimal time settings of the employed stimulus and acquisition paradigm, including the use of scan-to-scan intervals that at least equal the typical hemodynamic response duration (~10 s) and scan durations that are small compared to the hemodynamic time-to-peak.

In Chapter 3 the linearity of the hemodynamic response magnitude as a function of the stimulus intensity is discussed. Measurements are made in an extensive range of intensities. Cortical responses prove to increase predominantly linearly, especially for moderately weak to loud stimuli. However, for stimuli near the threshold of hearing the results leave open the possibility that non-linearities exist in the form of a relatively strong initial rise in the response signal. Various explanations for such threshold effects are discussed. Moreover, in this chapter a comparison is made between the activation characteristics in subjects

spreid ('sparse') scanparadigma op de meetbare hersenrespons. Gespreide paradigma's worden gekenmerkt door lange stille periodes zonder scanneractiviteit tussen opeenvolgende korte acquisities. In dit hoofdstuk wordt aangetoond dat de invloed van een enkele voorgaande scan de meetbare respons reeds tot een factor twee kan verkleinen, hetgeen op de aanwezigheid van niet-lineaire interactie-effecten duidt. Een eenvoudig model blijkt de gevonden resultaten kwalitatief te kunnen beschrijven, hoewel voor geluidsstimuli die lijken op het scanner geluid ingewikkelder gedrag wordt gevonden. Uit de resultaten worden een aantal aanbevelingen afgeleid omtrent optimale instellingen van het te gebruiken stimulus- en scanparadigma, omvattende het gebruik van intervallen tussen scans ter lengte van ten minste de duur van de hemodynamische respons, en een scanduur die kort is vergeleken met de hemodynamische piektijd.

In Hoofdstuk 3 wordt ingegaan op de lineariteit van de hemodynamische respons als functie van de intensiteit van de waargenomen stimuli. Er worden metingen gedaan in een uitgebreid bereik van luidheden. De responsen in de hersenschors blijken voornamelijk lineair toe te nemen, zeker voor matig zachte tot luide geluiden. Voor geluiden nabij de gehoordrempel laten de resultaten echter de mogelijkheid

with normal hearing on the one hand, and subjects with elevated hearing thresholds for high frequencies on the other hand. In the latter group a distorted perception of sound may occur, termed loudness recruitment. In such cases, an increase in stimulus intensity will lead to an abnormally large rise in the corresponding perceptual loudness. It is shown that in these subjects a similar disproportionately large increase in cortical activation can be found, as compared to normally hearing subjects. This shows that the activation signal that is measured by means of fMRI is less closely related to the attributes of the sound stimulus itself than to the subjective percept it evokes.

Chapter 4 discusses the response characteristics for a type of stimulus that consists of spectrally and temporally modulated noise ('dynamic rippled noise'). This type of stimulus has some acoustic features in common with speech and is therefore expected to evoke strong brain responses. Brain activation is determined as a function of the modulation frequency and density, and the results prove to be closely related to the sensitivity to this type of stimulus as determined in psychophysical experiments. Activation levels are separable with respect to the two modulation parameters, indicating that both characteristics affect the activation independently of each other. Also, the sensitivity for spec-

open dat er een niet-lineariteit optreedt in de vorm van een relatief sterke aanvankelijke stijging in het responssignaal. Mogelijke verklaringen voor een dergelijk drempel-effect worden besproken. Daarnaast wordt in dit hoofdstuk een vergelijking gemaakt tussen het activatiegedrag in proefpersonen met normaal gehoor enerzijds en met verhoogde gehoordrempels voor hoge tonen anderzijds. In deze laatste groep treedt een verstoorde geluidswaarneming op, genaamd luidheidsrekrutering (loudness recruitment). Hierbij leidt een zekere verhoging in geluidsintensiteit tot een abnormaal grote toename in perceptuele luidheid. Aange-toond wordt dat in deze individuen eveneens een grotere toename in hersenactiviteit gevonden kan worden dan in normaalhorenden. Dit toont aan dat het middels fMRI gemeten activatiesignaal minder nauw gerelateerd is aan kenmerken van de geluidsstimulus zelf dan aan het subjectieve percept dat deze oproept.

Hoofdstuk 4 bespreekt het responsgedrag op een type stimulus dat bestaat uit spectraal en temporeel gemoduleerde ruis ('dynamische ripples'). Dit type stimulus heeft zekere akoestische kenmerken gemeen met spraak, en wordt daarom geacht tot sterke responsen te leiden. De hersenactivatie wordt bepaald als functie van de modulatiefrequentie en -dichtheid, en de uitkomsten blijken nauw overeen te

tro-temporal modulations is mapped for the entire active auditory cortex, resulting in significant spatial gradients. This suggests the presence of a topographic and hierarchical organization of the auditory cortex.

In Chapter 5, the functional organization of the auditory cortex with regard to the lateralization and the frequency of the presented stimuli is studied. An optimized scan paradigm is used involving repeated contiguous acquisitions, which provides a more sensitive detection of cerebral activation. The optimal frequency and lateralization are mapped. The primary cortical areas are found to respond more strongly to the contralateral than to the ipsilateral ear. In contrast, a similar systematic trend is not found for the secondary regions. In addition, in the primary auditory areas a tonotopic gradient is shown. The sensitivity for the lateralization and frequency of the stimuli do not interact and are independent of each other. These results form evidence for fundamental differences in the functional organization between the primary and the secondary areas in the auditory cortex.

Chapter 6 studies the preferences for stimulus lateralization in more detail and encompasses various brain centers that are involved in auditory processing. Also, the functional connectivity between these centers is considered, and plastic changes are

komen met de in psychofysische experimenten bepaalde gevoeligheid voor dit type modulaties. Het activatieniveau blijkt separabel te zijn met betrekking tot de beide modulatieparameters, wat erop duidt dat beide karakteristieken de activatie onafhankelijk van elkaar beïnvloeden. Tevens wordt de gevoeligheid voor spectro-temporele modulaties afgebeeld op de actieve hersenschors, waarbij ruimtelijke gradiënten gevonden worden. Dit suggereert de aanwezigheid van een topografische en hiërarchische organisatie van de auditieve hersenschors.

In Hoofdstuk 5 wordt de functionele organisatie van de auditieve hersenschors met betrekking tot de lateralisatie (d.w.z. links- of rechtszijdigheid) en de frequentie (toonhoogte) van de aangeboden stimuli nagegaan. Hierbij wordt een geoptimaliseerd scanparadigma gebruikt met herhaalde opeenvolgende scans dat een gevoeliger detectie van activatie bewerkstelligt. De optimale toonhoogte en lateralisatie wordt afgebeeld op de hersenschors. De primaire auditieve gebieden blijken gemiddeld sterker te reageren op het contralaterale dan op het ipsilaterale oor. Daarentegen wordt voor secundaire gebieden een dergelijke systematische tendens niet gevonden. Tevens wordt een tonotopische gradiënt gevonden in de primaire auditieve gebieden. De gevoeligheden voor stimuluslateralisatie en

assessed in unilaterally deaf individuals. Activation is demonstrated in all studied areas, although it is found most consistently across the subjects in the inferior colliculi, the medial geniculate body, and the auditory cortex in the temporal lobe. The response in the aforementioned areas is found predominantly on the side contralateral to the stimulated ear, suggesting that the contralateral activation in the cerebral cortex directly reflects the lateralization of the activation in the upper midbrain and thalamus. Connectivity analyses confirm that in humans the auditory signals that arise in both ears cross already at low levels in the brainstem, and are transmitted further upwards on the other side of the brain. In unilaterally deaf subjects, deviations in the subcortical activation are hardly found in comparison with the normal hearing subjects, although the activation in the cerebral cortex itself is distributed more symmetrically. Stimulation of the deaf ear results in small amounts of activation that can be explained by cross hearing (i.e., sound perception through the unaffected ear). The connectivity analysis in the unilaterally deaf group confirmed that the stream of information from the deaf ear has essentially vanished, while that from the normal ear remains intact. Only the deprived cortical hemisphere derives some input from the contralateral side. These results suggest that plastic processes of reorganization are largely re-

-frequentie blijken geen interactie te hebben en zijn dus onafhankelijk van elkaar. De genoemde resultaten duiden op fundamentele verschillen in de functionele organisatie tussen primaire en secundaire gebieden in de auditieve hersenschors.

In Hoofdstuk 6 wordt de gevoeligheid voor stimuluslateralisatie uitgebreider onderzocht in de verschillende gebieden die bij auditieve verwerking betrokken zijn. Tevens wordt gekeken naar de functionele verbindingsterkte tussen deze gebieden en worden plastische veranderingen in eenzijdig dove personen nagegaan. Activatie kan worden aangetoond in alle onderzochte gebieden, hoewel het meest consistent over de gehele onderzochte groep in de inferieure colliculi, het mediale geniculate lichaam, en de auditieve hersenschors in de temporaalkwab. De respons blijkt in de genoemde gebieden met name contralateraal aan het gestimuleerde oor op te treden, hetgeen erop duidt dat de contralaterale activatie in de grote hersenen een directe weerspiegeling is van de lateralisatie van de activatie in de hersenstam en thalamus. Connectiviteitsanalyses bevestigen dat in de mens de auditieve signalen vanuit de beide oren reeds op een laag niveau in de hersenstam oversteken en langs de andere zijde verder worden doorgegeven. In eenzijdig doven blijken er nauwelijks verschillen in de subcorticale activatie op te treden in reactie op stimulatie

stricted to the cerebral cortex, and that in comparison they play only a minor role in the thalamus and brainstem.

Finally, Chapter 7 integrates the results of the various studies and discusses them in broader terms.

van het normale oor, hoewel de activatie in de hersenschors zelf symmetrischer verdeeld is. Stimulatie van het dove oor leidt tot een uiterst geringe activatie en kan verklaard worden door overhoren (het waarnemen van dit geluid middels het onaangedane oor). De connectiviteitsanalyse in deze groep bevestigt dat de informatiestroom gepaard aan het dove oor nagenoeg verdwenen is, terwijl die van het normale oor intact is. Alleen de gedepriveerde corticale hemisfeer verkrijgt enige invoer van de andere zijde. Dit alles suggereert dat plastische reorganisatieprocessen beperkt zijn tot hoofdzakelijk de hersenschors en in vergelijking slechts een kleine rol spelen in de thalamus en hersenstam.

Hoofdstuk 7 tenslotte integreert de resultaten uit de diverse studies en bespreekt deze in een breder kader.

INTRODUCTION

Chapter

1

§1.1 *About this Thesis*

1.1.1 *Aims and scope*

This thesis discusses the application of functional magnetic resonance imaging (fMRI) to the central auditory system (CAS) in the human brain.

Roughly 10 % of the western population suffers from significant hearing loss, deafness or other types of hearing impairment (Karlsmose et al., 1999; Johansson et al., 2003). Although the origins of such disorders are usually found in the peripheral hearing organ, some of the complaints are related to changes in the functioning of the CAS. The CAS is known to potentially display changes in its function as a result of prolonged changes in neural input from the ear, for instance due to various types of hearing disorders. Such plastic changes presumably play a role in the development of tinnitus (Rauschecker, 1999) and are of importance for understanding the reaction of the brain to the use of hearing aids like cochlear implants (Kral et al., 2001). However, current knowledge about the CAS is only schematic and mostly based on animal research.

fMRI offers the possibility to study the active CAS non-invasively and at a detailed level. The long-term goal of the research project that led to this thesis is to gain insight into the functional changes that may occur in the CAS as a result of peripheral hearing loss. This objective demands for methodological studies and investigations of the normal brain to obtain references regarding central auditory function, as well as for studies on groups of patients with various hearing disorders.

Therefore, the aims of the research project were threefold:

- i. to develop and expand on experimental techniques for performing fMRI on the CAS, and more specifically to quantify and overcome the disturbing influence of acoustic scanner noise;
- ii. to explore and investigate in more depth various organizational characteristics of the healthy human CAS that are presumed to be of fundamental importance for the understanding of its function;
- iii. to study relevant categories of patients with impaired hearing, and to compare the functioning of the brain in such patients with that in normal hearing subjects to gain more insight in the changes that may have occurred.

For these purposes, a number of studies were designed and performed. This thesis presents the main results that were obtained.

1.1.2 Outline

This first chapter is meant to be an introduction into the field, providing a brief overview of some topics in auditory fMRI that will be of general interest to the reader. It is not meant to be complete, but it should succinctly introduce the main topics and provide some starting points for further study. In §1.2 the CAS will be presented, while in §1.3 some issues related to fMRI will be discussed. People working on related subjects will probably be acquainted with the contents of some of the paragraphs, but may find other sections useful for the understanding of the studies that are to follow. Researchers that are actively involved in auditory fMRI should encounter few unfamiliar topics.

In the Chapters 2-6 that follow, the results of five studies will be described. The various chapters will contain relevant background information and references for further reading, and can be read independently from each other. Chapter 2 investigates the influence of acoustic scanner noise on the detected hemodynamic response signals. The timing of the experimental paradigm is varied in a systematic way to estimate optimal parameter values and to record the importance and magnitude of non-linear interaction effects between the scanner noise and the stimuli of interest. The linearity of the hemodynamic response to auditory stimuli is further discussed in Chapter 3, which concerns the relation between stimulus intensity and perceptual loudness and their relation to the activation as detected by fMRI. This chapter aims to study the fMRI signal as a measure for the subjective percept in subjects with and without hearing loss. In Chapter 4 the perception of spectro-temporal modulations is studied using a type of stimulus that is widely used in this thesis. The brain's sensitivity to these modulations is spatially mapped, leading to more insight into the functional organization of the auditory cortex. Along the same lines, in Chapter 5 mappings related to two other stimulus parameters are investigated, resulting in the assessment of cortical tonotopic gradients and preferences with regard to lateralization in normal hearing subjects. Chapter 6 extends results related to lateralization of the previous chapter towards the subcortical processing centers in the thalamus, midbrain and lower brainstem. It also aims to study functional reorganization in the CAS as a whole by comparing the results in normal hearing subjects with those in unilaterally deaf patients.

Chapter 7 concludes with a general discussion, reviewing the results of the various studies that have been described, and putting them into a single comprehensive framework. Main conclusions will be drawn and prospects will be discussed.

§1.2 The Auditory System

1.2.1 Peripheral hearing

The peripheral hearing organ, i.e. the ear, can be subdivided into three partitions: the external, middle and inner ears (Fig. 1.1). The external and middle ears, which together form the conductive component of the auditory apparatus, transmit sound waves from the external environment to the inner ear. The function of the inner ear is to transform the mechanical vibrations of sound into series of neuronal discharges, which can be conducted towards the brain for further processing (Martin, 1996; Yost, 2000; Moore, 2003).

The outer ear has two major functions. First, the structure of the auricle and the ear canal focuses the sound waves towards the tympanic membrane. In the process the sound intensity is amplified, especially in a frequency region near 3 kHz where humans are most sensitive. Second, the spectrum of the impinging sound is modified due to the intricate morphological structure of the auricle. Because the frequency transfer function depends on the spatial direction of the sound source, this provides important clues especially about the elevation of the source (Van Wanrooij et al., 2005).

The main function of the middle ear is to address the impedance mismatch that exists between the atmospheric air and the aqueous fluid in the inner ear. When a wave is transferred from a low-impedance medium (air) to one of high impedance (fluid), a considerable amount of its energy is reflected (in contrast, sound can quite efficiently be transmitted to the inner ear by means of bone conduction, as bone and soft tissue have impedances close to that of water). The middle ear overcomes this problem by boosting the pressure measured at the tympanic membrane in two ways. A first and major enhancement is achieved by focusing the force impinging on the relatively large diameter tympanic membrane onto the much smaller diameter oval window. A second and related process relies on the mechanical advantage gained by the lever action of the three connected ossicles. The mechanical characteristics of these middle ear bones can be influenced by the action of two small inner ear muscles. For instance, in response to strong sounds the conductive properties can be reduced in what is called the acoustic reflex.

Finally, the inner ear further amplifies the sound waves and simultaneously acts as a mechanical frequency analyzer. This is achieved by the cochlea, a coiled tubular structure consisting of three adjacent fluid compartments that are separated by flexible membranes (Fig. 1.2). One of these membranes, the basilar membrane, plays a crucial role in sound detection. The mechanical

characteristics of this membrane are such that it is wide and flexible at the apical end, and narrow and stiff at the basal end, resulting in a gradually varying resonance frequency along its length. The sound vibrations that enter via the oval window propagate along the length of the cochlea and will grow in amplitude until a maximum is reached at the site with a resonance frequency that corresponds with the sound frequency. Because the resonance frequency steadily decreases from the base to the apex of the cochlea, different sound frequencies will lead to resonances at different sites. More complex sounds result in a superposition of vibrations that resonate at different locations along the basilar membrane. Thus, the frequency content of the sound is spatially mapped onto the length of the cochlea, resulting in a frequency decomposition of the signal. As a rule of thumb, stimulus frequencies in the range of 20 Hz to 20 kHz can be detected by humans.

The basilar membrane contains hair cells with tiny processes called stereocilia that are embedded in the tectorial membrane which partly covers

Figure 1.2 A cross-section through a single winding of the cochlea (above), with a magnification of the microscopic structure of the organ of Corti (below). The arrival of pressure waves induces mechanical vibrations in the cochlea at a site of resonance. The amplified movement of the tectorial membrane causes deflections of the stereocilia of the outer (OHCs) and inner (IHCs) hair cells. Subsequently, the IHCs evoke neuronal discharges in some of the afferent fibers of the cochlear nerve.

the basilar membrane. These structures are together known as the organ of Corti. Two different types of hair cells exist which differ in their organization and function. A single row of inner hair cells (IHCs) along the length of the cochlea detects the resonant motion. When the basilar and the tectorial membranes are displaced, the hair cell's stereocilia are bended, inducing a voltage change across the hair cell membrane and triggering afferent discharges in a multitude of neurons that are attached to the hair cell. In addition, three rows of outer hair cells (OHCs) are innervated by neuronal efferents. These cells actively influence the motion of the basilar membrane in such a way that sharp frequency selectivity is accomplished.

1.2.2 Quality of hearing

The sensitivity of hearing can be described by hearing thresholds (Martin, 1996; Yost, 2000; Moore, 2003). In clinical practice, hearing thresholds are usually determined for pure tones in a typical frequency range of 125 to 8000 Hz using a logarithmic scale that is expressed in decibel (dB). Various scales that depend on different references are in use for the quantification of sound intensity. The sound intensity level (IL) and sound pressure level (SPL) refer to an absolute physical reference (equal to 10^{-12} W/m² and 20 μ Pa, respectively). Because the human threshold of hearing depends on the sound frequency, a hearing level (HL) scale has been introduced using a standardized frequency dependent reference level that corresponds with the average threshold in humans (Fig. 1.3). A frequency dependent reference

threshold can also be determined for individual subjects, leading to the sensation level (SL) intensity.

In contrast with the previously described sound intensity measures, the perceptual intensity or loudness of a stimulus is a subjective experience. A loudness scale can be defined by relating the subjective loudness of a given tone to that of a perceptually equally loud 1000-Hz reference tone. Fig. 1.3 also depicts the equal loudness contours related to the resulting phon scale. For instance, an 80-dB intensity tone at 3000 Hz has a loudness of 90 phons because it sounds equally loud as a 90-dB reference tone at 1000 Hz.

Hearing losses are most often related to temporary or permanent defects in the functioning of the peripheral hearing organ, but occasionally find their origin in central disorders. Two major types of hearing loss can be distinguished: conductive and sensorineural hearing loss. The term mixed hearing loss is used for a combination of the two.

Conductive hearing loss results from defects in the outer and/or middle ear,

Figure 1.3 The average human threshold of hearing (solid line) depends upon the tone frequency. Equal loudness contours (dashed lines) show the relationship between perceptual loudness level (number labels, in phons) and physical sound intensity level (IL) as a function of sound frequency (f).

and can be characterized by a reduced signal transmission from the surrounding environment towards the sensory cells. As a result, the hearing threshold as well as the equal loudness contours will shift upwards in Fig. 1.3 for such patients. Causes of conductive hearing loss include e.g. buildup of ear wax, perforation of the tympanic membrane, or sclerosis of the ossicular chain. In contrast with signal transmission via air conduction, bone conduction thresholds are not affected by conductive hearing loss since the middle ear ossicular chain is by-passed.

Sensorineural hearing loss relates to defects in the inner ear or the CAS, and accounts for the majority of cases of hearing loss. It is often the result of loss, damage or dysfunction of the IHCs and/or OHCs, and therefore affects both signal transmission via air and via bone conduction. Possible causes include noise trauma, various diseases, and the use of ototoxic drugs. Sensorineural hearing loss results in a deterioration of the threshold of hearing, while at the same time high-level equal loudness contours remain mostly unaffected. This causes the intermediate equal loudness contours to become compressed. This phenomenon is termed loudness recruitment and is characterized by a disproportionate increase in (perceptual) stimulus loudness with increasing (physical) stimulus intensity above threshold. Such a distorted perception may also affect the comprehension of speech in noise (Oxenham et al., 2003; Moore et al., 2004).

Central hearing loss is a type of sensorineural hearing loss that relates to disorders of the central nervous system, like acoustic nerve schwannomas (a type of tumor), focal injuries or lesions. Central hearing loss may manifest itself in an abundance of forms that will not be discussed further.

A different but related common auditory disorder is tinnitus, the presence of abnormal noises (ringing) in the ear that can not be related to any external sound source. In contrast with auditory hallucinations, tinnitus always consists of fairly simple sounds which can come and go or be continuous, and may be perceived in both ears or just in one ear. Sometimes, the tinnitus percept can be masked by other sounds. Tinnitus is often accompanied by hearing loss and can presumably have various causes related to the peripheral and/or the central auditory system. It is thought that the anomalous percept can be the result of faulty functional reorganization processes in the brain in reaction to the hearing loss. Diminished input for the affected frequency regions may cause a reduced inhibition in central auditory structures, leading to hyperexcitability of the central auditory system (Rauschecker, 1999; Eggermont et al., 2004; Weisz et al., 2005). However, little is yet known about the precise pathophysiology of tinnitus.

1.2.3 The auditory nerve

The neurons that contact the cochlear hair cells together constitute the cochlear nerve. It contains two types of neurons. Most abundant are the neurons that contact the IHCs. Their dendrites do not branch and thus make contact with only one IHC. However, each IHC is innervated by dozens of neurons. Less numerous are the neurons that connect with the OHCs. Their dendrites are ramified and contact multiple OHCs. In total, there are about 30,000 neurons in the human cochlear nerve.

The cochlear nerve joins the vestibular nerve to form the VIIIth cranial nerve. All information that is necessary for the brain to perform auditory functions like the recognition of speech and other sounds, or the localization and segregation of multiple sound sources, must somehow be encoded in the signals from the bilateral cochlear nerves. This is accomplished in various ways (Ehret et al., 1997; Yost, 2000; Taberner et al., 2005).

First, the frequency decomposition that is achieved in the cochlea is also reflected in the auditory nerve. The frequency of the tone to which a nerve fiber is most sensitive is called its characteristic frequency (CF). Adjacent frequencies can lead to excitation at higher intensity thresholds according to a tuning curve, which is usually asymmetrical and steeper for frequencies above the CF than for those below. High-frequency stimuli that excite the basal part of the basilar membrane induce discharges in the peripheral layers of the auditory nerve, whereas low-frequency sounds that stimulate the apical end lead to activation of the neurons in the nerve's core. This gradual spatial variation in CF is termed tonotopy, and is maintained in some form in the auditory processing chain at least up to and possibly including the secondary auditory cortex. As such, it is one of the major organizational principles in the auditory system.

Second, the intensity of the sound is encoded in the firing rate of the neurons. Intense stimuli generally evoke more potentials than weak stimuli. Individual neurons can be characterized by a firing threshold and a saturation level. Apart from its spontaneous firing frequency, which is inversely related to the threshold, the neuron will not fire in response to stimuli below threshold. Stimuli that are above threshold will result in an elevated firing frequency. When the stimulus level exceeds the threshold by approximately 30 dB, the neuron's maximum firing frequency has been reached and more potentials cannot be evoked by increasing the stimulus intensity further. However, because individual neurons differ in their response characteristics (more specifically in their thresholds), and because IHCs are innervated by multiple neurons, an enormous dynamic range of

some 110 dB can be obtained in the responses of the neuron population as a whole.

Third, the timing of the evoked potentials is locked to the phase of the sound stimulus. This is related to the fact that depolarization of the hair cells occurs for one specific bending direction of the stereocilia. For low sound frequencies (typically below 1 kHz), the discharges of individual neurons are able to track the undulations in the sound wave. For intermediate frequencies (~1-4 kHz), the refractory period prevents a neuron from firing with each wave front, but the responses of all neurons as a whole are still partially synchronized with the sound wave. For the highest frequencies (>4 kHz), the neurons cannot follow the phase of the sound wave (although of course the tonotopic organization in the cochlea still allows these wave frequencies to be distinguished).

Fourth, because sound is detected binaurally, differences in the responses from the left and the right ear give important clues for sound localization. For low sound frequencies, differences in the timing of discharges related to the phase difference of the impinging sound at both ears can be interpreted to yield information of the sound source in the horizontal plane. For high frequencies this source of information breaks down, but it is replaced by the emergence of intensity differences due to the shadowing effect of the head. Also, spectral information related to the head related transfer functions (HRTFs) from the left and the right ear will complement each other.

1.2.4 Subcortical pathways

The cochlear nerve terminates in the cochlear nucleus (CN) that is located near the junction between the pons and the medulla in the brainstem (Fig. 1.4). The CN acts as a switchboard, distributing auditory information to several different areas in the CAS. It also does some processing of both time and frequency information. The CN is typically divided into three regions based on the morphology of the cells they contain and the structures with which they connect. These divisions are the anterior ventral cochlear nucleus (AVCN), the posterior ventral cochlear nucleus (PVCN) and the dorsal cochlear nucleus (DCN). While the auditory nerve connects exclusively to the ipsilateral CN, the CN itself divides its output into several bilaterally ascending tracts (Ehret et al., 1997; Purves et al., 1997; Yost, 2000).

The entryway to the 'binaural pathway' is the AVCN, which provides direct input to the ipsilateral superior olivary (SO) nuclei and indirect input to the contralateral counterpart via the medial nucleus of the trapezoid body. The medial and lateral superior olives (MSO and LSO) are the two nuclei most

intimately involved in binaural processing. The MSO effectively cross-correlates the signals arriving at the two ears, thus computing interaural time-delays due to sound location along the azimuth. LSO cells, in contrast, are sensitive to interaural intensity differences also related to sound source location. The binaural pathway is therefore essential in spatial localization. The SO projects up to the inferior colliculus (IC) via a fiber tract called the lateral lemniscus (LL). Projections to the IC provide a cross-over point for information involved in source localization, allowing the segregation of information according to hemifield rather than ear source. The SO is also the point of origin of the efferent nerve fibers that innervate the outer hair cells in the cochlea.

The 'monaural contralateral pathway' begins at the DCN. As suggested by its name, this pathway carries information to the contralateral side only. Unlike the time-sensitive binaural pathway, this stream mostly conveys information about spectral content to analyze the quality of sound. The DCN is able to pick apart tiny frequency differences. This pathway projects

Figure 1.4 Schematic outline of the ascending connections of the central auditory system, showing the afferent pathways emanating from one ear. Fibers project from the inner hair cells in the cochlea towards the cochlear nucleus (CN) that acts as a switchboard and transmits signals further via at least two pathways. The binaural pathway contains connections to the bilateral superior olivary nuclei (SO), where binaural comparisons of time-delay and sound intensity are made on behalf of spatial localization. Signals are transmitted via the lateral lemniscus (LL) to the inferior colliculi (IC). In parallel, the monaural contralateral pathway directly relays spectral information from the CN via the contralateral LL to the IC. The IC is the major auditory processing center in the brainstem and has afferents that project to the bilateral medial geniculate nuclei (MG) in the thalamus. From here, signals are further transmitted to the auditory cortex (AC) in the upper surface of the temporal lobes. The auditory cortices in the cerebral hemispheres exchange information via connections through the corpus callosum.

directly to the IC, also via the LL, and is unique in that it carries a substantial amount of descending projections.

The LL is a tract of axons traveling from just above the SO all the way up to the IC. Only some of the axons synapse in the LL, but it is involved in the acoustic startle reflex. The LL carries fibers whose cell bodies reside in several different structures, the most important of which is nucleus of the LL (NLL).

All ascending auditory information passes through the IC, which receives input from a host of nuclei including the CN, the SO, and the nuclei of the LL. It contains three morphologically distinct nuclei. The largest and most well studied of these is the central nucleus (ICC). The remaining portions of the IC have been subdivided into the external nucleus and the pericentral nucleus, or further into the dorsal cortex, the dorsomedial nucleus, and the lateral nucleus.

The IC appears to be an integrative relay station. It is responsive to interaural delay and amplitude differences and may provide a spatiotopic map of the auditory environment that is maintained using amongst others visual information from the nearby superior colliculus. Changes in the spectrum of sounds such as amplitude and frequency modulations appear to have a modulotopic representation in the IC. Sound features such as stimulus duration and periodicities in the stimulus envelope are also thought to be encoded in the IC, and it might be involved in the suppression of echoes in a reverberant environment. The sensitivity to spectral changes conceivably provides the primitive building blocks for neurons responsive to specific phonemes and intonations necessary to recognize speech. In addition, the IC is involved in integration and routing of multimodal sensory perception. It is involved in the startle reflex and sends projections which are involved in ocular reflexes. It also receives input from the somatosensory system and modifies activity in regions of the brain responsible for attention and learning.

From the IC, information proceeds to the thalamus, which is the last relay site on the way to the cortex for almost all sensory information, including the auditory, visual, and somatosensory modalities. Axons of the IC synapse on the medial geniculate (MG) nuclei on the same side and on the opposite side. The MG has been divided into ventral, dorsal and medial subnuclei based on functional as well as morphological distinctions. The ascending inputs to the MG arise from the subdivisions of the IC by way of three largely segregated pathways. The projection from ICC to the ventral subnucleus of the MG preserves an orderly spatial arrangement of CFs, and shows multiple mirrored tonotopic organizations. The vast majority of cells

are binaurally sensitive. A second projection from the lateral nucleus of the IC to the medial MG forms a polysensory pathway. Finally, the projection from the dorsal cortex of the IC to dorsal MG is part of a 'diffuse' pathway, so named because of a lack of obvious tonotopic organization.

There is massive feedback from primary cortex to the MG and also from other cortical centers through the thalamus. So, the MG would appear to be the place where top-down processing comes into play, and it may be involved in focusing attention on one object or another in the image of the auditory environment, or in sharpening certain features that arise in the auditory image. The MG projects to the auditory cortex (AC), located on the upper surface of the temporal lobe.

1.2.5 The auditory cortex

The cerebral cortex contains the highest unimodal auditory processing stages, distributed over the upper part of the temporal lobe (Ehret et al., 1997; Purves et al., 1997; Konig et al., 2005). The superior temporal surface (Fig. 1.5) contains an eye-catching transverse gyrus extending in the posteromedial to anterolateral direction, called Heschl's gyrus (HG). In some individuals this gyrus may be doubled or forked, resulting in considerable morphological variability across individuals. Anterior to the HG lies the first transverse temporal gyrus that separates HG from the planum polare (PP) that extends towards the tip of the temporal lobe. Posteriorly, HG borders on Heschl's sulcus (HS) which converts to the planum temporale (PT) in the most posterior part of the Sylvian fissure. On the lateral side, these regions are lined by the superior temporal gyrus (STG) and, more inferiorly, the superior temporal sulcus (STS).

The auditory cortex can be subdivided in numerous sub-areas on the basis of neurophysiological criteria. Following the parcellation scheme by Brodmann (Brodmann, 1909), the auditory cortex roughly comprises areas 41 (BA-41, primary auditory cortex), 42 (BA-42, secondary auditory cortex) and 22 (BA-22, auditory association cortex). In most animal species that have been investigated, a primary core region is distinguished on the basis of its koniocortical architecture (Kaas et al., 1999). It receives input from the ipsilateral MG in the thalamus, but also projects to the contralateral primary auditory cortex via the corpus callosum. Although the architectonic location of the primary auditory cortex does not always register with the morphological features of the cortex due to inter-subject and inter-hemispheric variations, it is approximately situated in the medial two-thirds of HG (Rademacher et al., 2001). It can be subdivided in two or three fields

along the axis of the HG, although their precise number and designation is still being discussed (Morosan et al., 2001). In non-human mammals distinct

Figure 1.5. (a) A lateral view of the left cerebral cortex, showing a parcellation of the cortical surface according to Brodmann (Brodmann, 1909). The auditory cortex is situated in the upper part of the temporal lobe. Brodmann area 41 (BA-41) roughly corresponds with Heschl's gyrus and comprises the primary auditory cortex (core region). Brodmann area 42 (BA-42) corresponds with the planum temporale and largely contains the secondary auditory cortex (belt region). Brodmann area 22 (BA-22) covers part of the superior temporal gyrus and comprises auditory association and possibly language areas (parabelt region). (b) A T_1 -weighted oblique transverse MRI slice through the upper surface of the temporal lobe, showing (1) the planum polare, (2) Heschl's gyrus, (3) the planum temporale, (4) the superior temporal gyrus, (5) the circular sulcus, (6) the first transverse sulcus, and (7) Heschl's sulcus. (c) The same anatomical slice, indicating the approximate locations of the (i) core, (ii) belt, and (iii) parabelt regions. Subdivisions of these regions have not yet been reliably established in humans. (A=Anterior; I=Inferior; L=Lateral; M=Medial; P=Posterior; S=Superior)

mirror-symmetric tonotopic progressions have been found within these fields (Ehret, 1997), but in humans this organization has yet to be confirmed definitively (Formisano et al., 2003; Talavage et al., 2004b). In addition, the functional differences between these fields regarding the processing of sound are still obscure. Knowledge is even more limited for the surrounding secondary regions, often referred to as the belt and parabelt areas. Various subdivisions into fields have been proposed, some of which may contain similar but less distinct tonotopic organizations as the primary fields. Because this is currently still a topic of ongoing research these will not be discussed in further detail here.

Processing in the auditory cortex is likely to be performed in a serial fashion, with functional fields forming a hierarchical organization. This is corroborated by findings that show relatively sharp frequency tuning and short-latency responses in primary regions, while secondary areas required complex spectro-temporal stimulus dynamics to elicit responses (Wessinger et al., 2001). This suggests that spectral and temporal information is integrated over increasingly broad ranges in frequency and time, going from primary to secondary cortical processing centers. At the same time, multiple parallel processing streams presumably exist. Thus the various fields may accomplish complex sound analysis tasks like timbre discrimination, virtual pitch perception, foreground-background decomposition, spatial localization, or noise filtering (Recanzone et al., 2000; Lee et al., 2004; Inui et al., 2005).

Differences in the hemispheric specialization with regard to the physical characteristics of sounds have been observed. It is known that both cerebral hemispheres are most responsive to signals from the contralateral ear and to stimuli originating from the contralateral hemifield of the auditory environment (Woldorff et al., 1999; Alho et al., 2003; Krumbholz et al., 2005b). However, it has also been suggested that the left hemisphere specializes in rapid temporal processing whereas the right hemisphere is better equipped for fine spectral analysis (Zatorre et al., 2002; Sininger et al., 2004; Behne et al., 2005). This may be directly related to the language-related functions that have traditionally been ascribed to the left hemisphere (more specifically Wernicke's area near the temporo-parietal junction), as temporal variations in the sound envelope are critical for the comprehension of speech (Shtyrov et al., 2000; Crinion et al., 2003; Josse et al., 2003; Zeng et al., 2005). However, it has been noted that task and attention related top-down influences constitute a possible confound in the characterization of hemispheric specialization, complicating the interpretation of these results (Tervaniemi et al., 2003; Brechmann et al., 2005).

Although still speculative, it has also been suggested that the cortical processing results in a re-encoding of the incoming auditory signal into multiple streams associated with so-called auditory objects (Griffiths et al., 2004; Nelken, 2004; Zatorre et al., 2004). The existence of partially separate but interconnected processing streams for object identification ('what' the objects is) and object localization ('where' the object is) has been proposed; these would also respectively involve frontal and parietal cerebral regions that are known to bear some relation to high-level auditory processing (Kaas et al., 2000; Rauschecker et al., 2000; Arnott et al., 2004; Brunetti et al., 2005).

Although much work has been done to investigate the response properties of various processing centers in the CAS, even the basic nature of the auditory code (i.e., which physical attributes of the incoming auditory information are ultimately represented by the neural responses) is still largely unknown (Eggermont, 2001). Especially for the cerebral auditory cortex in humans, organizational principles have not yet been established.

§1.3 *Functional Magnetic Resonance Imaging*

1.3.1 *Physics*

The signal that is exploited in all magnetic resonance (MR) applications originates from the magnetic moment that is associated with the spin of an atomic nucleus. More specifically, MR imaging (MRI) makes use of the protons that form the nuclei of hydrogen atoms, which are abundant in body tissue in the form of water molecules. Because protons behave like tiny magnets, they can be aligned by the application of an external magnetic field B_0 (Fig. 1.6). In spite that only a slight majority of the protons will simultaneously align, the small magnetic moments of the whole proton population in a sample will merge to form a net steady state magnetization M_0 that is parallel to the main magnetic field. Although the properties of individual hydrogen nuclei can only be described by quantum-mechanical methods (Gasirowicz, 1974; Haacke et al., 1999), their overall behavior can be summarized by analyzing the net magnetization in a classical fashion (Jezzard et al., 2001; Buxton, 2002).

By means of radio frequency (RF) excitation pulses, the net magnetization can be flipped. As a result, the component of the magnetization parallel to the applied magnetic field (the longitudinal magnetization $M_{||}$) will decrease, and a component perpendicular to the magnetic field (the transverse magnetization $M_{\perp}$) will develop. Both components behave differently.

Firstly, the longitudinal component will slowly grow back to its steady-state magnitude by a relaxation process involving spin-lattice interactions. Relaxation causes the discrepancy between the momentary magnetization and its equilibrium value to decrease exponentially, with a time constant T_1 . For brain tissue, this time constant is in the order of 1 s.

Secondly, the transverse component will show a more complicated behavior, consisting of two simultaneous effects. It will relax back to equilibrium (i.e.,

zero magnitude) due to spin-spin interactions with a time constant T_2 that is different from that of the longitudinal relaxation. For brain tissue this is in the order of 0.1 s. Simultaneously, the transverse component's direction will rotate around the axis of the applied magnetic field, a process called precession. The precession frequency depends upon the local magnetic field strength as well as on the gyromagnetic ratio of the particular nucleus; for hydrogen nuclei in a typical clinically used field strength of 1.5 T, this frequency amounts to 64 MHz. This is equal to the resonance frequency of the RF-pulse that is needed to flip the magnetization. Because a rotating magnetic dipole emits electromagnetic radiation, the presence of transverse magnetization can be detected through its precession.

1.3.2 Magnetic resonance imaging

In MRI the emission of radiation from the rotating transverse magnetization is used as a mechanism to extract information about the location of the source nuclei. The basic scheme consists of the following steps (Haacke et al., 1999; Jezzard et al., 2001; Buxton, 2002). First, the magnetization is given a chance to relax back from its initial state to equilibrium. Because usually $T_1 > T_2$, this will take a time equal to a few times T_1 , in the order of seconds. Next, an RF-pulse is applied that flips the magnetization over a typical angle of 90° into the transverse plane. Then the transverse magnetization will rotate, emitting a detectable electromagnetic signal. Simultaneously, the signal-emitting transverse component will quickly shrink while the longitudinal component will slowly grow back to its original magnitude. Eventually, the magnetization will achieve its steady-state magnitude again, and a new RF-pulse can be applied a total time TR after the previous one, the repetition time, starting the process all over again.

Because the amount of emitted radiation is proportional to the density of hydrogen nuclei, and this depends upon the tissue type, anatomical images can be constructed by detecting the power of the radiation that is emitted from a particular site. However, this magnitude cannot be determined directly, because the signals from other sites that contain tissue will interfere. This problem can be solved by applying a gradient field on top of the homogeneous main field. This leads to a small spatial variation in the magnetic field strength, and this in turn will cause the flipped magnetization to precess at different frequencies at different sites.

The differences between the precession frequencies at various locations will cause the magnetic moments of different protons to dephase. This will lead to a faster decrease of the measured signal because the radiation from

different sites will no longer add up constructively. The time constant that is associated with this accelerated relaxation is called T_2^* . The signal loss can be partially recovered by the application of a second RF-pulse that flips all magnetization over 180° ; hereby, all spins with a phase lag will be turned into spins with a phase lead, and the magnetic moments will start to refocus. This will lead to a burst of electromagnetic radiation with a frequency that corresponds with the local magnetic field strength, in what is called a spin echo. Alternatively, additional gradient fields can be used to dephase all spins and refocus them again into a so-called gradient echo. The time between the initial 90° -flip and the occurrence of the echo is called the echo time TE.

Because different frequency components in a composite signal can be unraveled very efficiently by Fourier-analysis, the signals that emerge from sites with different field strengths can be separated. By using multiple excitations with different strengths and directions of the gradient field, the signals from all individual locations can be measured, and reconstructed in the form of three-dimensional image volumes with image intensities that correspond with the magnitudes of the locally emitted echo signals.

In practice, numerous variations of this basic scheme are used to obtain different kinds of images. By shortening TR, the magnetization can be flipped when not yet fully recovered. Because different types of tissue have different T_1 , this can be exploited to obtain images with different contrasts between tissues. Similarly, by varying TE, images with different T_2 - or T_2^* -contrast can be constructed. Also, the flip angle can be adjusted, as well as new RF-pulses and/or gradient switches be introduced to obtain totally new image contrasts. Some common examples are inversion recovery (IR) sequences, turbo spin echo (TSE) and echo planar imaging (EPI), diffusion weighted or tensor imaging (DWI/DTI), and flow imaging. In the studies in this thesis, EPI was used exclusively, which essentially uses the signal from multiple gradient echoes in an efficient way to construct T_2^* -sensitive images.

1.3.3 Dynamic contrast mechanisms

Functional MRI is an indirect method to measure brain activation (Jezzard et al., 2001; Buxton, 2002; Buxton et al., 2004). It does not rely directly on the electric or magnetic processes that play a role in the signal conduction by neurons, like for instance electro- and magneto-encephalography (EEG/MEG) or evoked potential (EP) measurements do. Similar to single photon emission computer tomography (SPECT) and positron emission

tomography (PET), fMRI relies on a period of increased metabolism following the actual neuronal activity. These metabolic effects in turn lead to gradual vascular responses that fMRI is able to detect (Nair, 2005). Although the temporal resolution of fMRI is very poor as compared to techniques that rely on neuronal discharges directly, its spatial resolution is superior to all neuroimaging techniques mentioned. As an aside, in recent years some of the mentioned modalities have been combined successfully to make use of their respective strengths (e.g., the temporal resolution of EEG in combination with the spatial resolution of fMRI) (Mulert et al., 2005).

Synaptic activity in neurons, whether excitatory or inhibitory, and to a lesser extent neuronal spike activity are accompanied by increased energy consumption and an increase in the cerebral metabolic rate of oxygen (CMRO₂) (Logothetis et al., 2004b; Logothetis et al., 2004a). Because the metabolic reserves within neurons and the surrounding glia cells are limited, the oxygen supply from the blood bed will have to rise. This is achieved by vascular dilatation, an enlargement of the blood vessel diameter, leading to increased cerebral blood volume (CBV) and cerebral blood flow (CBF). Because erythrocytes will then have less chance to exchange oxygen with the surrounding tissues, the oxygen extraction fraction will decrease and the oxygenation level at the venous side of the neural activity will rise. That is, somewhat counter-intuitively, the ratio of deoxygenated hemoglobin (deoxy-Hb) to oxygenated hemoglobin (oxy-Hb) will drop.

Because hemoglobin contains a ferrous core that changes with respect to its magnetic properties when it binds to oxygen, a change in oxygenation level will lead to a change in the magnetic susceptibility of the blood tissue. More specifically, oxy-Hb is diamagnetic, similar to the surrounding brain tissue, and does not lead to significant distortions of the magnetic field. However, deoxy-Hb has paramagnetic properties that strongly differ from its surroundings, and this difference causes the magnetic field lines to deform locally. The irregularities in the magnetic field strength that arise from these deformations lead to a quick dephasing of the nuclear magnetic moments, reducing the net transverse magnetization. Thus, deoxygenated blood has a relatively short T_2^* . In part, this dephasing can be recovered by a 180° refocusing excitation pulse (such that $T_2 > T_2^*$), but due to random molecular motion, some residual phase differences will remain.

In blood oxygenation level dependent (BOLD) fMRI, this mechanism is used to obtain image contrast from activated brain areas (Fig. 1.7). Regions of the brain that are active will respond with an increase in CBF and CBV, leading to an increase in the local blood oxygenation level. This will in turn result in a reduction in the local field inhomogeneities, and a rise in T_2^* . In a T_2^* -sensitive EPI sequence, this will show up as a local increase in image intensity called the hemodynamic response signal. By performing acquisitions both during alleged rest (i.e., baseline) and during active task performance, the resulting differences in image intensity can be detected. Essentially, when the corresponding images are subtracted the active brain locations will light up.

Other contrast mechanisms do exist. For instance, activation is also followed by changes in T_2 and T_1 that can be detected. However, the activation signal magnitude in T_2^* images is superior (Parkes et al., 2005), and because the signal differences are small by nature this mechanism is traditionally used most often. However, completely different mechanisms can be suitable for specific applications. Changes in CBV are detectable using IR sequences (a technique called vascular space occupancy, VASO), while changes in CBF can be detected by arterial spin labeling (ASL) techniques (Lu et al., 2003; Stefanovic et al., 2004; Yang et al., 2004). Possibly, in the future the magnetic effects that neuronal discharges have on the MR-signal can even be measured directly. The various techniques can differ in spatial or temporal resolution, their sensitivity upon vessel size, the importance of susceptibility artifacts, their dependence on magnetic field strength, etc. Because the development of pulse sequences or contrast mechanisms is not subject of this thesis, in these studies T_2^* -based BOLD fMRI was employed exclusively.

1.3.4 Experimental designs

Regarding the timing of experimental fMRI paradigms, a distinction is usually made between block-related and event-related studies (Fig. 1.8) (Jezzard et al., 2001; Buxton, 2002). In block-related experimental designs, stimuli are presented during long on- versus off-intervals while MR acquisitions are performed continuously. This allows hemodynamic BOLD responses to the stimulus to reach a steady state which can reliably be measured relative to the baseline signal. In this way, active brain regions can be localized more or less independently of the BOLD response dynamics.

Figure 1.8 The typical timing (t) of the acquisitions, stimuli and responses that occur in experiments differs between block-related (above), event-related (middle) and sparse designs (below). Block-related designs allow for an effective assessment of the response magnitudes, whereas in event-related designs the response shape can be sampled efficiently. In fMRI studies on the CAS that use sparse designs, the acquisitions are spaced by periods of scanner silence to reduce the influence of the responses to acoustic scanner noise (dashed curves).

Advantages of such block- or state-related techniques are that the requirements on the acquisition rate of the scanner are not so stringent, and that the contrast is relatively high. Thus, a measure for the magnitude of the local hemodynamic response can be obtained quite efficiently. However, in spite that block-related experiments allow the signal rise from baseline and signal fall to baseline to be detected, these techniques are poorly efficient to study signal dynamics and do not allow the shape of the hemodynamic response to be determined directly.

As rapid fMRI techniques like EPI and spiral imaging have become commonly available, research is directed more and more towards measuring the shape of the hemodynamic response function to gain more insight in the underlying neural, metabolic and vascular mechanisms. Therefore, event-related paradigms have been developed that essentially study the shape of the BOLD response to short stimuli. If MR-acquisitions can be performed sufficiently fast, consecutive scans can be executed following a short stimulus, thus sampling the hemodynamic response to a single stimulus at multiple moments in time.

Finally, various mixed designs have been tried in search of an optimal compromise between signal detection and signal shape measurement. These include designs where block-related periods are alternated with event-related periods, designs with randomly varying block lengths or randomly timed events, and designs based on m-sequences (Dale, 1999; Buracas et al., 2002; Liu, 2004).

1.3.5 Acoustic scanner noise

Apart from the technological challenges associated with fMRI in general, fMRI of the CAS in particular is hampered by the influence of acoustic scanner noise. The EPI acquisitions that are commonly used in fMRI paradigms are accompanied by sound waves that can reach intensities over 100 dB SPL for standard 1.5-T scanner systems (Cho et al., 1997). These are generated by mechanical vibrations that occur as a result of the interaction between the fast switching magnetic gradients and the strong homogeneous magnetic field in the scanner. The scanner noise may interfere with the stimuli that are presented to the subject in the context of the protocol of the experiment, thus affecting the auditory perception directly. Also, scanner noise is likely to elicit hemodynamic responses in the same brain areas that are involved in the processing of the experimental stimuli, and may therefore affect the signal that is measured. A number of approaches can be

followed to avoid or minimize these undesirable effects (Amaro et al., 2002; Di Salle et al., 2003a; Moelker et al., 2003).

First, silent acquisition types can be used to reduce the scanner noise intensity. This can for instance be achieved by decreasing the slew rate, i.e. the speed with which the gradient coils are switched. Sometimes this can be achieved without changing the imaging method, but the potential gain will be constrained by considerations related to the TE of the acquisition. To further decrease the maximum slew rates, the sampling of Fourier space (or *k*-space) can be optimized by employing different *k*-space trajectories like in spiral imaging or sensitivity encoding (SENSE) (Hennel, 2001; de Zwart et al., 2002). Finally, various other imaging parameters like the TE or the EPI factor can be varied. However, this will also influence the image contrast and the strength of image artifacts, and can therefore detrimentally affect the signal to noise ratio.

Second, various noise attenuation methods can be used. A distinction can be made between passive and active noise attenuation techniques. Passive noise attenuation comprises the use of sound insulation. For instance, head phones and/or ear plugs may be used that can reduce the intensity of ambient noise by approximately 30 dB if combined (Moelker et al., 2005). Also, more radical approaches may be tried involving the use of helmets or a screening of the entire scanner bore (Ravicz et al., 2001). In addition, active noise attenuation could potentially reduce the perceived scanner noise by another 30 dB. Sound with an opposite phase is generated to largely cancel the scanner noise through destructive interference, either real-time or by using pre-recorded scanner sounds (Chambers et al., 2001). However, such techniques are technologically very challenging and are not yet available for use in a strong field scanner environment. Also, in practice active noise attenuation is only efficient for low sound frequencies and is further limited by the occurrence of bone conduction to the ear.

Third, the experimental paradigm may be modified to reduce the effect of scanner noise on the measured signals. So-called sparse imaging paradigms have been proposed (Fig. 1.8) in which functional images are not acquired continuously, but separated by relatively long periods of scanner inactivity (~10 s) (Hall et al., 1999). Auditory stimuli can be presented during these gaps of scanner silence, avoiding the direct interference of scanner noise with the perception of the presented stimulus (Novitski et al., 2003). At the same time, the delay in the hemodynamic response can be exploited by delivering the stimuli shortly before the acquisitions (~4 s), such that the timing of an actual data acquisition coincides with the maximum in the hemodynamic response to the previously presented stimulus. For optimal

results, slice acquisitions should be clustered and together should not take longer than approximately 2 s per image volume (Talavage et al., 1999).

1.3.6 Data processing

Typically, hemodynamic response amplitudes have a magnitude of only a few percent of the baseline signal level. At the same time the signal noise, consisting of measurement noise and physiological fluctuations, has a similar magnitude. Therefore, activation signals can only be discriminated from the noise floor by performing numerous acquisitions, and by applying statistical techniques to determine at which sites the measured signal courses contain significant amounts of hemodynamic response contributions.

Before the actual signal detection can take place, a number of preprocessing steps is necessary (Jezzard et al., 2001). First and foremost, some spatial realignment has to be performed to correct for subject movement during the course of the functional acquisitions. This is usually done by automated registration using rigid body rotations. Optionally, an additional (possibly piecewise) affine registration can be performed to bring the functional volumes of multiple subjects into alignment with each other, or with a canonical reference brain volume. This corrects for subject to subject differences in the overall anatomy of the brain to some part, and allows the functional activation data of multiple subjects to be pooled. Standardized coordinate systems have been developed for this purpose, e.g. stereotaxic Talairach coordinates. Alternative visualization and registration techniques have also been developed based on virtual inflation and flattening of the cortical surface. Next, the volumetric or surface data can be smoothed by means of a convolution with a Gaussian kernel. Although this decreases the spatial resolution, it improves the signal to noise ratio. Furthermore, it has the advantage of fixating the spatial correlation structure in the data, which may be required to make accurate statistical inferences. Finally, it can be necessary to correct for drifts in the acquired scanner signal over the duration of the functional sessions.

Following these preprocessing steps, the acquired data are analyzed for each voxel individually to establish the presence of any hemodynamic responses to the experimental stimulus conditions. For this purpose, a general linear model (GLM) is set up that incorporates the expected time courses of the responses to each of the employed stimulus types. Using linear regression analysis, the amplitudes of these responses are fitted (Friston et al., 1998a;

Norman et al., 2000; Buxton, 2002). Alternatively, a correlation analysis can be used for simple experimental setups with few stimulus types.

Statistics are performed on the resulting regression or correlation coefficients to determine the statistical significance of the responses to the respective stimuli. Analysis of variances (ANOVA) can be performed, resulting in *F*-test statistics. In a sense, this test assesses whether inclusion of a certain stimulus condition into the model is necessary in order to describe the measured data reliably. The importance of multiple conditions collectively can also be tested. An advantage of *F*-tests is that few prior assumptions have to be made on the sign and relative strength of the responses. Alternatively, Student *T*-test statistics can be used to test whether the magnitude of a particular regression coefficient is large enough not to be explained only by chance. These are suitable when the sign of the coefficients is important, e.g. when only positive responses should be considered. Again, more complex contrasts can be constructed by combining multiple coefficients into linear sums with predetermined weightings.

The resulting significance levels from individual voxels are combined into whole-brain statistical parametric mappings (SPMs), which can be thresholded at a certain *p*-value to obtain three-dimensional information on which brain regions are active. Thresholds can for instance be chosen to restrict the statistically expected family wise error (FWE) rate or the false discovery rate (FDR) below an acceptable level (typically 5 %) (Nichols et al., 2003). Further restrictions can be imposed by limiting the activation to a region of interest (ROI), or by requiring a minimal cluster size of adjacent active voxels.

Several types of measures for the amount of activation in the brain can be constructed. Apart from differences in their designations and from the intricate details of the calculations, two measures are common. On the one hand, an activation volume (AV) can be calculated by counting the number of voxels that is significantly active. This measure reflects the spatial extent of the activated brain areas. On the other hand, an activation level (AL) can be computed by averaging the percentage signal change over a predetermined set of active voxels. This measure is an indication of the amount of neural activity that is present within the selected voxels.

The results from multiple individuals can be combined into analyses at the group level. This can be done using fixed effects analyses, which are very sensitive to activation but which assume the effect of interest to be present in all subjects in equal fashion. Because of this assumption it is not possible to make inferences regarding the significance of the detected effects in the population as a whole, although this can be partially overcome by the use of

conjunction analyses (Friston et al., 1999). Alternatively, random effects analyses can be employed which allow the strength of the effect to differ between subjects, and which allow population inferences to be made. However, as a result these analyses are much less sensitive. Another option is to perform statistics on the summary activation characteristics of individual subjects (e.g., the AV and AL). Such methods combine good sensitivity with the power to determine the significance of effects on the population level, but are often unable to account for differences in the spatial characteristics of activation patterns.

The regression analysis that is used in the evaluation of the GLM is very effective in the detection of response signals if the hemodynamic response behavior can be predicted with some accuracy. This is the case if the stimulus conditions are unambiguously defined. However, model-free approaches exist that can detect significant activation patterns irrespective of the delivered stimuli.

A basic technique that is able to extract overall response behavior from the data itself is principal component analysis (PCA) (Viviani et al., 2005). In essence, this mathematical technique describes the signal variations by means of a series of successively less important terms. Using factor analysis, the most important terms can be extracted. Although this technique allows for the detection of activation patterns with a minimal amount of initial assumptions regarding the nature of these activation patterns, it is unable to separate multiple simultaneous signal contributions. Independent component analysis (ICA) can overcome this disadvantage when the various signal contributions can be assumed to show non-Gaussian variations (Hyvarinen et al., 2000). At the moment, both PCA and ICA strongly rely upon retrospective interpretation to assign functional meaning to the signals that are found, making it hard to draw objective conclusions. While PCA is very useful for the compression of data with minimal loss of potential effects of interest, ICA is currently gaining popularity as a means for detecting brain activation patterns and promises to become a serious alternative for the more traditional GLM approaches.

Another class of methods does not measure the strength of localized responses to certain stimuli, but quantifies the extent to which brain areas are connected (Jezzard et al., 2001; Friston, 2002). Although fiber bundles can actually be traced by MRI using DTI, fMRI connectivity measures should not be interpreted in the anatomical context of intrinsic structural connectivity; only functional meaning should be assigned (Ramnani et al., 2004). A basic functional connectivity measure can be calculated by correlating the signal courses in different brain areas with each other. Brain regions involved in

closely related processing types will exhibit similar responses, and will therefore result in high mutual correlation coefficients. Such a measure can be calculated without prior knowledge of the delivered stimulus conditions. Indeed, no stimuli need to be presented at all, and this has led to the identification of resting state networks (van de Ven et al., 2004). More advanced connectivity measures take into account the influence that various regions exert on each other simultaneously. Using a connectivity model, the importance of multiple inputs can be quantified by means of path analysis (Norman et al., 2000). Various variations and extensions of such methods have been proposed, including canonical correlation analysis (CCA) (Friman et al., 2001), partial least squares (PLS) (McIntosh et al., 2004), structural equation modeling (SEM) (Horwitz et al., 1999), and dynamic causal modeling (DCM) (Friston et al., 2003), to name a few. However, these promising methods are still under development and need to be assessed and validated further.

HEMODYNAMIC RESPONSE INTERACTIONS

Published in modified f
Magn. Res. Med. 2005;5
doi:10.1002/mrm.20315

Abstract

In fMRI experiments on the central auditory system, activation caused by acoustic scanner noise is a dominating factor that partially masks the hemodynamic response signals to sound stimuli of interest. In this study, the non-linear interaction between auditory responses to single scans and those to tone stimuli was investigated. By using irregular acquisition repetition times and quasi-random stimulus timings, the brain responses to pure tone stimuli were analyzed, as well as their interaction with scanner noise. The tone frequencies were chosen to match either the fundamental frequency of the scanner noise (730 Hz) or a region with little spectral power (4.70 kHz). The hemodynamic responses could be characterized by amplitudes of 1.3 % and a time-to-peak of 4.0-4.5 s in the absence of scanner noise. Interaction effects due to a single previous scan typically decreased the response magnitudes to 0.9 %. The functional shape of the interaction was analyzed and could be described by a highly separable, dominantly symmetric interaction function that fairly agreed with a low-order Volterra expansion of a simple non-linear model. Interactions were stronger and more complex in shape when the spectral content of the tone stimulus and the scanner noise were more similar.

§2.1 Introduction

In functional MRI (fMRI), non-linearities in the hemodynamic response behavior can severely complicate the interpretation of experimental results. In particular, considerations regarding non-linear interactions between stimuli are of importance in fMRI studies involving the central auditory system, where the response to sound stimuli can be saturated by the presence of acoustic scanner noise (Eden et al., 1999; Edmister et al., 1999; Elliott et al., 1999; Hall et al., 1999; Amaro et al., 2002; Moelker et al., 2003). When using continuous block-related acquisition schemes, the response to the background scanner noise will lead to an elevated baseline activation level and possibly a reduction in the dynamic range of the responses that experimental stimuli will additionally induce.

To overcome the signal interactions due to acoustic scanner noise, sparse scanning paradigms have been developed. These are inherently similar to event-related paradigms, with the additional feature that acquisitions are spaced by relatively large intervals of scanner silence. Auditory stimuli can be presented during these gaps, thus avoiding any direct interference from scanner noise. Also, hemodynamic responses to previous scanner noises will have declined to low levels at the onset of an acquisition. Short acquisitions are used such that acoustic influences of the concurrent scan do not have time to accumulate. A major disadvantage of sparse paradigms is their inefficiency regarding data acquisition, since few scans can be executed in a given amount of time. A compromise might be achieved by shortening the silent scan interval (Shah et al., 2000). However, not only can auditory stimuli show interactions when presented simultaneously, but also successively if intervals are shorter than the typical duration of the hemodynamic response (~10 s) (Friston et al., 1998b). To determine optimal settings, more insight into the non-linear mechanisms behind the hemodynamic response and knowledge about the interaction between acoustic scanner noise and other auditory stimuli is desirable.

Previous studies have shown non-linearities in the hemodynamic response to auditory stimuli. For instance, in some experiments the stimulus rate was varied by changing the interval between consecutive short stimulus presentations (Binder et al., 1994a; Rees et al., 1997; Buchel et al., 1998; Friston et al., 1998b; Inan et al., 2004; Harms et al., 2005). A non-linear dependence of the activation level on the presentation rate of the stimuli was found. The response was found to be saturable, signifying that the response level increased less strongly for high than for low presentation rates. In some

studies the response characteristic did not only asymptotically flatten, but even decreased again for sufficiently high presentation rates. For single stimuli such observations can be understood as resulting from a short refractory period (~6 s) following each response.

Other experiments studied the dependence of the activation level on the stimulus duration (Robson et al., 1998; Glover, 1999; Jäncke et al., 1999; Bohning et al., 2003; Birn et al., 2005). If responses were linear, the shape of the responses to long stimuli could be predicted by convolving the response to a shorter stimulus with a suitable repetition function. Whereas such convolution models lead to reasonable results when applied to long stimuli only, large deviations appear when short stimuli are included (typically smaller than 6 s). Generally, responses to short stimuli are more strongly peaked than expected. This suggests some form of adaptation when stimuli persist for sufficiently long.

Although a number of experiments have considered the influence of the intensity or loudness of the stimulus, their results regarding the linearity of the response are inconclusive (Millen et al., 1995; Jäncke et al., 1998; Mohr et al., 1999). Whereas some authors observe a monotonic increase of activation with sound intensity (Hall et al., 2001; Hart et al., 2003a), the works of others demonstrate much weaker correlations and even report evidence for a topographic principle of intensity encoding called amplitopicity (Bilecen et al., 2002; Brechmann et al., 2002).

Finally, multiple studies investigated the influence of scanner noise and a variety of ways to minimize it by varying the timing of the presentation of stimuli and the acquisition of images (Bandettini et al., 1998; Eden et al., 1999; Edmister et al., 1999; Elliott et al., 1999; Hall et al., 1999; Shah et al., 1999; Talavage et al., 1999; Shah et al., 2000; Yang et al., 2000; Le et al., 2001; Schmithorst et al., 2004; Talavage et al., 2004a). Significant differences in response amplitude have been found as the timing was varied, demonstrating some influence of scanner noise upon the detectable signals. However, most of these experiments are not very suitable to draw direct conclusions about non-linearities in the blood oxygenation level dependent (BOLD) response characteristics or about the exact shape of interactions between the hemodynamic response to scanner noise and that to other stimuli.

This study was designed to investigate the shape of the interaction between acoustic scanner noise and pure tone stimuli and to quantify the level of brain response saturation for different tones. The hypothesis is that interaction effects will decrease the signal levels as compared to the predictions of a linear model. Furthermore, it is expected that interaction

effects are strongest when the spectral content of the sound stimuli and the scanner noise are more similar.

§2.2 *Materials and Methods*

2.2.1 *Subjects and preparation*

Twelve healthy subjects with normal hearing were recruited (9 male, 3 female; aged 21-59 years, mean 33 years) and gave written informed consent to participate in this study, in approved accordance with the requirements of the medical ethical committee of the institute. Of the subjects, 7 were right-handed, 3 left-handed and 2 ambidextrous (Oldfield, 1971). During the experiment, the scanner room was dimly lit and subjects were asked to close their eyes. They wore ear plugs and a noise protection headset with built-in piezo-electric speakers. The subjects were asked to press one of two fibre-optic button-devices that were held in the left and right hands whenever they perceived an auditory stimulus of a low or a high frequency respectively. The purpose of this simple task was to let subjects focus on the stimuli; the results were used only to monitor that the subjects were able to discern the stimuli, understood the task, and stayed alert.

2.2.2 *MR acquisitions*

Subjects were placed in supine position in the bore of a clinical 1.5-T MRI system (Philips Intera, Philips Medical Systems), which was equipped with a standard quadrupolar head receiver coil. A sagittal 3D T_1 -weighted fast-field echo scan was acquired for anatomical orientation (repetition time TR 9.85 ms; echo time TE 3.5 ms; flip angle 15°; matrix 256×256×50; voxel dimensions 0.94×0.94×3.0 mm). On these high-resolution images a central imaging plane was positioned approximately parallel to the Sylvian fissure of both hemispheres, containing the superior temporal gyrus of both temporal lobes. An additional 3D T_1 -weighted fast-field echo scan (TR 10.37 ms; TE 3.5 ms; flip angle 15°; matrix 256×256×20; voxel dimensions 1.25×1.25×1.75 mm) was conducted parallel to this imaging plane to serve as anatomical reference. The functional scans were acquired covering this same brain volume, and consisted of 1.3-s single-shot T_2^* -sensitive echo planar imaging (EPI) sequences with ten 3.5-mm slices (TE 50 ms; flip angle 90°; matrix 96×96; field of view 240 mm) which were performed at irregular repetition times (TR = 2, 4, 6, 8 or 10 s) by varying the silent gap between consecutive scans

(Fig. 2.1). The 23 mT/m maximum gradient field strength was switched in 0.22 ms (105 T/m/s slew rate); the fundamental switch frequency for the gradient fields amounted to 730 Hz approximately. For each subject the functional session consisted of six runs of 47 acquisitions each, of which the two initial scans were discarded.

2.2.3 Stimulus paradigm

Two types of auditory stimuli were used in this experiment. First, pure tones

Figure 2.1 Schematic representation of a fragment of the stimulus paradigm in time (t). The timing of the acquisitions and tone presentations was compiled such that for each acquisition at most one preceding tone presentation (of either high or low frequency) and at most one previous scan occurred at intervals Δt and TR smaller than 10 s. Scans and tones that preceded an acquisition by 10 s or more were assumed to have no influence, as well as stimuli that coincided with the acquisition. In the above figure, part of such a paradigm is pictured along with the expected hemodynamic responses. Because of the irregular acquisition scheme, differences in T_1 -relaxation at the onset of a scan lead to strong contrast differences between scans with different TR.

were presented by means of head phones connected to a signal generation set-up (System II 16-bit DA converter, Tucker Davis Technologies). These tones had a duration of 1.3 s, a frequency of 730 Hz or 4.70 kHz, and an intensity of 70 dB sound pressure level (SPL) after correction for ear plug damping. Second, the scanner noises that were generated during all acquisitions were simultaneously used as stimulus sources (1.3 s duration, 65 dB SPL corrected for damping). The fundamental frequency of the scanner spectrum (Fig. 2.2) was 730 Hz (i.e., equal to the frequency of the low tone stimuli); in contrast, only a small amount of spectral power was present at 4.70 kHz (i.e., the frequency of the high tones). Stimuli that preceded an acquisition by less than an interval of 10 s were considered relevant exclusively; all stimuli that started before or after this interval were assumed not to contribute to the BOLD response, because the signal either would have faded or could not have reached a significant level yet. Within this 10-s time window, the onset of the tones was timed to occur at an interval $\Delta t = 2, 4, 6$ or 8 s before the start of the subsequent acquisition (Fig. 2.1). Because of the irregular acquisition scheme, scanner noise stimuli could also occur within this window at an interval $TR = 2, 4, 6$, or 8 s before an

Figure 2.2 The acoustic spectrum of the EPI acquisitions consists of a fundamental frequency (f) of 730 Hz plus higher harmonics. The intensity (I) of the scanner noise is approximately 65 dB SPL (corrected for damping by the headset and ear plugs). The low-frequency tone stimulus (730 Hz) was chosen to coincide with the fundamental frequency; the high-frequency tone stimulus (4.70 kHz) was positioned between two relatively weak higher harmonics. The intensities of both tones were similar (70 dB SPL, corrected for damping).

acquisition. The paradigm was set up in such a way that at most one tone and at most one previous scan occurred in the 10-s window before each acquisition.

Using this experimental setup, 9 different conditions existed with regard to the timing of the tone (either a high- or a low-frequency tone 2, 4, 6 or 8 s before an acquisition; or no tone at all in the 10-s window, modelled as $\Delta t = 0$ s). Simultaneously, 5 different conditions existed with regard to the presence of scanner noise (again 2, 4, 6 or 8 s before an acquisition; or no scan at all in the preceding 10-s window when $TR = 10$ s). All 45 possible combinations of previous scan and tone presentations were included exactly once in each functional run. The order and timing of the acquisitions and stimuli was chosen differently in all six runs.

2.2.4 Model

To characterize the hemodynamic response quantitatively, four contributions to the image intensity I of a voxel were distinguished: (1) a baseline signal I_0 , (2) a hemodynamic response to the preceding tone in the absence of scanner noise H_{tone} , (3) a hemodynamic response to the previous scan in the absence of tones H_{scan} , and (4) an interaction between the responses to tones and scanner noise H_{interact} . The model equation combining these responses is

$$I(\Delta t, TR) = I_0 \cdot (1 + H_{\text{tone}}(\Delta t) + H_{\text{scan}}(TR) + H_{\text{interact}}(\Delta t, TR)) . \quad [2.1]$$

The baseline level I_0 equals the signal when no stimuli are of any influence; under the assumption that only stimuli in the preceding 10-s time window are relevant, this corresponds with the experimental data for which $\Delta t = 0$ s and $TR = 10$ s, implying $I_0 = I(0,10)$. To yield all other terms dimensionless, I_0 was treated as a multiplicative constant. H_{tone} can be determined from the data for which no scan took place in the interval preceding the acquisition (i.e., $TR = 10$ s). Analogously, H_{scan} can be derived from the data for which no tone was presented precedingly (i.e., $\Delta t = 0$ s). All remaining signal is accounted for by the interaction term H_{interact} . When this last term is omitted, a conventional linear model with two stimulus types is obtained.

As the TR is varied, resulting differences in T_1 -contrast in the images will constitute an additional source of signal variation. Therefore, the baseline image intensity I_0 will depend on TR . However, since Eq. 2.1 already contains the term H_{scan} that also depends upon TR only, it is impossible to make a distinction between the effects of these two unknown functions. Because hemodynamic response amplitudes are two orders of magnitude

smaller than the overall signal level, T_1 -relaxation effects will likely obscure the scanner noise responses.

Using the approximation $\ln(1+x) \approx x$ for small x , the following approximation can be derived from Eq. 2.1.

$$\ln(I(\Delta t, TR)) = \ln(I_0(TR)) + H_{\text{tone}}(\Delta t) + H_{\text{interact}}(\Delta t, TR). \quad [2.2]$$

The function H_{scan} was omitted as a separate term because of the T_1 -relaxation effect previously mentioned, and is now assumed to be incorporated into the value of $I_0(TR)$. Equation 2.2 can be used directly to segregate the data into signal components without assuming any specific functional forms for the functions I_0 , H_{tone} and H_{interact} . The function $I_0(TR)$ can be determined from the subset of intensity data $I(\Delta t, TR)$ for which $\Delta t = 0$ s. Once this function is known, it is possible to directly determine the function H_{tone} from the subset of data for which $TR = 10$ s. Finally, H_{interact} is obtained by subtracting these two functions from the original data set. It follows that by definition $H_{\text{tone}}(0) = H_{\text{interact}}(0, TR) = H_{\text{interact}}(\Delta t, 10) = 0$.

When the functional forms of H_{tone} and H_{interact} are prescribed a priori, Eq. 2.2 can be expressed as a linear regression model according to

$$\ln(I(\Delta t, TR)) = \sum_{k=1}^5 \ln(I_0^k) \cdot \delta^k(TR) + \beta \cdot h_{\text{tone}}(\Delta t) + \gamma \cdot h_{\text{interact}}(\Delta t, TR). \quad [2.3]$$

The function $I_0(TR)$ has been modelled by a sum of five separate parameters I_0^k for each of the five values for TR ; the functions $\delta^k(TR)$ are set to 1 for data for which the TR equals the k^{th} possible value ($TR = 2, 4, 6, 8$ or 10 s, respectively), and 0 elsewhere. The hemodynamic response term and the interaction term have been written as the product of amplitude coefficients (β and γ) and unit-amplitude hemodynamic response and interaction functions h . These last two terms should be duplicated to account for the presence of a low or a high frequency tone separately. Thus, the complete regression model allows nine regression parameters to be calculated: five baseline levels I_0^k , two tone response amplitudes β , and two interaction amplitudes γ .

2.2.5 Data analysis

The functional images were motion-corrected and spatially smoothed with a 5-mm Gaussian kernel using the SPM99 software package (Wellcome Department of Cognitive Neurology). In the MatLab environment (The MathWorks Inc.), temporal drifts of the baseline signal were corrected for by a first order polynomial for each of the six runs separately, preserving signal

differences due to differences in T_1 -contrast. Because the timing of the presented series of stimuli was compiled differently in each run, with all conditions occurring only once in a run, and because the average scan interval was moderately large (6 s), the signal noise in consecutive acquisitions was assumed to be uncorrelated to simplify the analysis. The natural logarithm of all data was taken and the low and high tone data were collected into a 9×5 -matrix as a function of the presence of tone stimuli (9 conditions) and scanner noise (5 conditions). Hereafter, the analysis described next was carried out in two similar steps to reduce any bias from the initial choice of model parameters.

First, a simplified regression model was constructed on the basis of Eq. 2.3, initially omitting the interaction term (effectively forcing $\gamma = 0$). The hemodynamic response was modelled by a unit-amplitude gammavariate function $h_{v,\tau}$ according to

$$h_{v,\tau}(t) = \left(\frac{t}{\tau} \exp \left(1 - \frac{t}{\tau} \right) \right)^v. \quad [2.4]$$

The parameter v is a measure for the relative narrowness of the response function, and was initially set to $v = 3$; the parameter τ equals the time-to-peak and was initially chosen equal to $\tau = 5$ s on the basis of estimates from literature (Cohen, 1997; Aguirre et al., 1998; Friston et al., 1998a; Glover, 1999; Le et al., 2001). After evaluation of the model, voxels were considered active if β for the low or high tone response was significantly larger than zero using Student T -test statistics; p -values were thresholded such that extensive activation was detected, primarily confined to the auditory cortex ($p < 0.001$, uncorrected). Next, Eq. 2.2 was used to segregate the original data from all voxels into a TR-dependent baseline contribution and a Δt -dependent tone response contribution; the interaction term was again ignored initially. The data from the voxels that were previously determined to be active were averaged, yielding average functions $I_0(\text{TR})$ and $H_{\text{tone}}(\Delta t)$. To determine optimal values for the parameters v and τ , the resulting function $H_{\text{tone}}(\Delta t)$ was fitted non-linearly to a gammavariate function (Eq. 2.4) using a Levenberg-Marquardt algorithm. The purpose of this initial linear analysis was to determine values for the parameters v and τ on the basis of the data themselves and thus minimize any bias with respect to the response shape in a second iteration of these analyses.

In the second step, similar analyses were performed including the interaction term. The full regression model (Eq. 2.3) was applied using the parameter values v and τ that resulted from the initial linear analysis. As

explained in appendix 2.6.1 of this chapter, the interaction term was modelled as a direct product of two gammavariate functions

$$h_{\text{interact}}(\Delta t, \text{TR}) = h_{v,\tau}(\Delta t) \cdot h_{v,\tau}(\text{TR}). \quad [2.5]$$

The set of active voxels for which $\beta > 0$ for the low- or high-frequency tone was determined anew ($p < 0.001$, uncorrected) using interaction term regressors that were orthogonalized with respect to the linear response terms. Next, the data matrices and the values of the response and interaction amplitudes were averaged over all active voxels for each subject individually.

The data matrices were segregated into multiple contributions according to Eq. 2.2. The response contribution H_{tone} was non-linearly fitted to a gammavariate function (Eq. 2.4). The baseline data $I_0(\text{TR})$ were non-linearly fitted to a function

$$I_0(\text{TR}) = I_{\infty} \cdot \left(1 - \exp\left(-\frac{\text{TR}}{T_1}\right) \right), \quad [2.6]$$

describing the effect of T_1 -relaxation. The response interaction term H_{interact} was analyzed for its functional dependence on the arguments Δt and TR: its magnitude was characterized by averaging all non-zero entries in the data matrix; its shape was quantified with respect to separability, symmetry and correlation with Eq. 2.5 as described in appendix 2.6.2.

§2.3 Results

2.3.1 Linear model

Using the initial regression model without interaction terms, the location of all significantly active voxels was determined. In a number $N = 1131 \pm 241$ of voxels the activation threshold was reached (mean $\pm$ standard error). For this set of voxels, average response amplitudes to the 730-Hz and 4.70-kHz tones were $\beta = 1.20 \pm 0.13 \%$ and $\beta = 1.37 \pm 0.13 \%$, respectively. Tone response shape parameters were fitted by means of the segregation model. The average time-to-peak was $\tau = 4.60 \pm 0.08$ s for the low frequency tone and $\tau = 4.65 \pm 0.08$ s for the high frequency tone. The values for the relative response narrowness diverged more widely, $v = 4.59 \pm 0.51$ for the low tone and $v = 5.36 \pm 0.35$ for the high tone. Neither τ nor v differed significantly for the two types of tone stimuli in a paired Wilcoxon signed ranks test ($p > 0.1$). However, the average values did deviate significantly from the values that

were initially chosen for the model functions in the regression model ($\tau = 5$ s; $\nu = 3$). Representative estimates for the response shape parameters $\tau = 4.6$ s and $\nu = 5$ were deduced from these results, and these values were used in the models including interaction that will be discussed in the subsequent paragraphs.

2.3.2 Regression amplitudes

By including interaction terms into the regression model in the full non-linear analysis, extensive activation could be detected in all subjects. No further restrictions were placed on the cluster size or location of the activation, but the activation was consistently found to be predominantly concentrated in the temporal lobes and organized into multiple stripe-like patterns running in the (postero)medial to (antero)lateral direction. There

Figure 2.3 Low- and high-frequency tone activation in two subjects. The activation level of all significantly active voxels was color-coded and projected perpendicularly on an anatomical slice through the upper surface of the temporal lobe. No restrictions were imposed on the cluster size or location of the activation. The upper pictures show the activation level in response to a low frequency tone stimulus; the lower ones to a high frequency stimulus. Two subjects are shown with small (left) and large (right) amounts of activation respectively. The activation is concentrated in the temporal lobes and is organized in stripe-like patterns running diagonally, approximately parallel to Heschl's gyrus. Generally the response to high frequency tones was slightly larger than that to low frequency tones, but the location of activation foci was similar.

was a substantial range of activations over various subjects, both regarding the extent (i.e., the total volume of significantly active voxels) and the magnitude (i.e., the average response amplitude) of the activation. In Fig. 2.3, low and high tone activation patterns are projected on an anatomical background for two subjects with relatively small and large amounts of activation respectively.

The total number of active voxels $N = 1909 \pm 388$ showed an increase as compared to the first iteration. Also, the number of voxels with significant activation in response to the high-frequency tone ($N = 1581 \pm 341$) was larger than that to the low-frequency tone ($N = 1068 \pm 255$), but the difference was

only marginally significant ($p < 0.1$). The average response amplitudes over all significantly active voxels resulting from the regression model were equal to $\beta = 1.25 \pm 0.16 \%$ and $\beta = 1.37 \pm 0.19 \%$ for the low and high tone respectively. According to Student T -tests, the interaction coefficient $\gamma = -0.41 \pm 0.14 \%$ differed significantly from zero for the low frequency tone ($p < 0.05$), but not in the case of $\gamma = -0.31 \pm 0.17 \%$ for the high frequency tone ($p > 0.1$). The sign of the interaction term was negative, indicating that the interference from scanner noise tends to decrease the amplitude of the response to tone stimuli. Therefore, when the influence of the preceding scan is at its maximum, the fitted response amplitudes are reduced to 0.8 % (for the 730-Hz tone) or 1.0 % (for the 4.70-kHz tone).

The interaction amplitudes γ from the regression model for the various subjects correlate well with the response amplitudes β (Fig. 2.4). The corresponding Pearson correlation coefficients equalled $R = -0.80$ and $R = -0.88$ for the low- and high-frequency tone respectively, indicating that high activation levels in subjects were accompanied by strong interactions as a rule. The average ratio γ/β equalled -0.25 for the low tone and -0.16 for the high tone. For the low tone this ratio is marginally significantly different from zero ($p < 0.1$) while for the high tone the difference is insignificant ($p > 0.1$).

2.3.3 T_1 -relaxation effects

Fig. 2.5a shows the value of I_0 as a function of TR as obtained with the segregation model. Also shown is an exponential fit (Eq. 2.6) corresponding to parameter values $T_1 = 1.02 \pm 0.04$ s and $I_\infty = 4.31 \pm 0.02$. The vertical spread in the data points is mainly caused by variations in I_∞ over all subjects, and much less so because of variations in T_1 . Furthermore, the TR-dependent I_0 data were fitted for all active voxels individually and the resulting values for I_∞ and T_1 were used to normalize the data. Typical T_1 -values were in the range 0.77-1.42 s (10-90 % percentiles). In Fig. 2.5b, the individual voxels' I_0/I_∞ are plotted as a function of TR/T_1 . If Eq. 2.6 were exactly valid, all data points would lie on a single exponential curve. However, five partially overlapping lobes of data points can be distinguished, each corresponding with a set of data belonging to one of the five different values for TR. The standard deviation of the normalized data from the exponential function amounts to 1.2 %. A substantial part of this variance is related to a systematic tendency for the data points to lie above the theoretical curve for $TR = 2$ s as well as for $TR = 8$ and 10 s, whereas I_0 is systematically lower for $TR = 4$ and 6 s.

Figure 2.5 TR-dependence of the baseline signal level. In panel a, the TR-dependent baseline level I_0 is shown, averaged over all active voxels. A fitted exponential function is shown by means of the dashed curve, corresponding with $T_1 = 1.02$ s. In panel b, a similar fit was performed on the data of all active voxels individually, and the resulting parameters I_{∞} and T_1 for each voxel were used to normalize the data along both axes. All data points cluster along the theoretical curve $I_0/I_{\infty} = 1 - \exp(-TR/T_1)$, but separate lobes of data points can be identified that correspond with different values of TR. Systematic deviations of these lobes from the exponential curve exist that can be attributed to the influence of partial-volume effects. To illustrate this, the inset shows a simulation of single voxel data according to a bi-exponential relationship (solid curve). Similar to the actual data, for intermediate TR (4/6 s) the data points lie below the best exponential fit (dashed curve), whereas for both lower and higher TR-values the data points lie above this curve. Each voxel contributes one data point to each of the five lobes in the main figure.

2.3.4 Response shapes

Fig. 2.6 shows the value of H_{tone} as a function of Δt , averaged over all subjects. The data were fitted to a gammavariate function (Eq. 2.4) with an amplitude A , resulting in parameters $A = 1.18 \pm 0.01 \%$, $\tau = 4.12 \pm 0.03 \text{ s}$ and $\nu = 2.85 \pm 0.08$ for the low tone, and $A = 1.45 \pm 0.11 \%$, $\tau = 4.20 \pm 0.18 \text{ s}$ and $\nu = 4.28 \pm 0.79$ for the high tone.

In Fig. 2.7 the average value of H_{interact} as a function of TR and Δt is outlined by means of contour plots for both tone frequencies. These functions equal zero for TR = 10 s and for $\Delta t = 0 \text{ s}$ by definition, and have been extrapolated to zero for TR = 0 s and for $\Delta t = 10 \text{ s}$. In general terms, a maximally negative magnitude is reached for intermediate values of both arguments ($\sim 4 \text{ s}$). If all non-zero entries in the data matrix H_{interact} are averaged to obtain a measure for its magnitude, $H_{\text{interact}} = -0.43 \pm 0.19 \%$ for the low tone and $H_{\text{interact}} = -0.38 \pm 0.18 \%$ for the high tone; for both tones this is significantly different from zero ($p < 0.05$). In similar fashion, the standard errors in these interaction data have been outlined in Fig. 2.7. Some deviations occur from

subject to subject for those values of TR and Δt that correspond with the largest interaction effects. Still, the uncertainties in the interaction data are typically about twice as small as the actual values themselves.

When the separability of the interaction term was calculated (see appendix 2.6.2), for the low-frequency tone $\alpha_{\text{sep.}} = 0.92$ and for the high-frequency tone $\alpha_{\text{sep.}} = 0.94$. The degree of symmetry of the averaged interaction functions equalled $\alpha_{\text{sym.}} = 0.64$ and $\alpha_{\text{sym.}} = 0.82$ respectively. The difference in the interaction symmetry between the two tones is also directly visible from Fig. 2.7; particularly the half of the 730-Hz tone graph below the main diagonal (for which $\text{TR} < \Delta t$) has an irregular character, with a notable deviation at $(\Delta t, \text{TR}) = (8, 4)$ s. The Pearson correlation between H_{interact} and the model function (Eq. 2.5) amounted to $R = -0.51$ for the low-frequency tone and $R = -0.78$ for the high-frequency tone. Both correlations were highly significant ($p < 0.01$).

§2.4 Discussion

2.4.1 Hemodynamic tone responses

The hemodynamic response to tone stimuli was modelled by a gammavariate function (Cohen, 1997; Backes et al., 2002). The results from the segregation model confirmed that this model was sufficiently accurate (see Fig. 2.6). Estimates could be derived for the amplitude, time-to-peak and relative narrowness of the response. The response to high-frequency tones was generally larger than that to low-frequency tones, both with regard to the amplitude of the activation as well as its spatial extent. Despite the fact that the sound delivery system was calibrated to deliver equally intense stimuli, such a difference is not an unusual finding. The recruitment of auditory cortical neurons differs between high and low frequencies, with a tendency for the high frequency tones to be more effective in eliciting extensive cortical responses (Lockwood et al., 1999; Hart et al., 2003b). The time-to-peak τ that is derived from the model was consistently close to 4.0-4.5 s. Although this deviated from our initial estimate, it is in agreement with results from a variety of studies (Aguirre et al., 1998; Friston et al., 1998b; Belin et al., 1999; Backes et al., 2002; Pfeuffer et al., 2003). Regarding the response narrowness there was a somewhat wider range of values that could be found: $\nu = 3-5$, depending upon the model and tone frequency. Differences in the overall response shape as determined by the parameters τ and ν were insignificant between the two stimulus frequencies.

Since the activation level of a voxel was determined from a regression model that requires prior assumptions about the shape of the response, some bias

Figure 2.7 The magnitude and error of the interaction between low- and high-frequency tones with scanner noise. The mean magnitude of the interaction H_{interact} between a 730-Hz or a 4.70-kHz tone stimulus and scanner noise is plotted in the upper panels as a function of the scan-to-scan and stimulus-to-scan intervals (TR and Δt , respectively) by means of contour lines. The negative sign of the interaction indicates a decrease in hemodynamic response signal due to interference between both types of stimuli. Although it is apparent that the interaction for the low-frequency tone is more complexly shaped, both interactions were highly separable and dominantly symmetrical. The lower panels show similar plots of the corresponding standard errors. Conditions with strong interaction effects are accompanied by larger standard errors.

in the assessment of these parameters is inevitable. In this study this bias was minimized by performing an initial analysis where the interaction terms were neglected to reduce the model to a traditional linear model. The fact that the resulting response shape parameters τ and ν differed significantly from the initial estimates can be brought forward as evidence that the model is sufficiently flexible to detect responses of various shapes. Changes in the parameter values still existed after the second application of the model when interaction terms were included, but were less pronounced. The remaining uncertainties in the exact response shape parameters were not considered to be serious, because the outcomes for most of the quantities of interest were not sensitive to such differences. This can for instance be illustrated by the fact that the resulting response amplitude β did not differ significantly between the linear and the non-linear model.

2.4.2 T_1 -relaxation

For MR-acquisitions the signal amplitude is proportional to the net longitudinal magnetization at the onset of a scan. When acquisitions take place at irregular time intervals, the magnitude of the magnetization will depend on the repetition time TR due to T_1 -relaxation effects. By plotting the results as averages over the different subjects (Fig. 2.5a), an exponential relationship was found as expected. The resulting value for T_1 (1.0 s) is characteristic for gray matter at 1.5 T (Haacke et al., 1999).

A more detailed picture was obtained when the data were plotted for voxels individually. To obtain meaningful results, the data were normalized to correct for the large differences in T_1 and I_∞ that occur from voxel to voxel. This was achieved by plotting $I_0(\text{TR})/I_\infty$ as a function of TR/T_1 (Fig. 2.5b), where T_1 and I_∞ were obtained for each individual voxel by fitting its data to Eq. 2.6. Now the data points appeared to be organized into a number of distinct lobes, each corresponding to a particular TR value. The data points for the smallest and the largest values of TR systematically lied in a lobe somewhat above the theoretical relationship according to Eq. 2.6, whereas for intermediate TR the signal levels were slightly smaller. These deviations cannot be explained from the hemodynamic response to the scanner noise that was incorporated into the definition of $I_0(\text{TR})$ in Eq. 2.2; such responses would lead to a slight increase in the baseline signal level for intermediate TR, and most likely only decrease the T_1 that would be fitted. However, the observed deviations are compatible with the presence of partial-volume effects: a bi-exponential function for instance would lead to better fits, given that such a function shows a stronger initial increase as compared with a

similar mono-exponential function, followed by a somewhat more flattened course at intermediate argument values accounting for the deficits near $TR = 4/6$ s (see the inset in Fig. 2.5b). Still, given the relatively small number of data points (5 per voxel) and the large number of degrees of freedom of a bi-exponential function, it was not practical to impose such a functional relationship upon $I_0(TR)$.

In fMRI, the amplitude of any hemodynamic response is generally small in relation to the baseline signal, in the order of a few percent. In this study, typical response amplitudes equalled 1.3 %. If images with different TR are to be included in a single analysis, corrections for the aforementioned T_1 -relaxation will have to be made with an accuracy that is preferentially better than about 1 % in order not to obscure the response signal. Our results show that such accuracy is hardly achievable because of the presence of partial volume effects. Analyses that combine images with different TR are feasible if the stimulus onset can be varied independently of the timing of the acquisitions (Schmithorst et al., 2004). When investigating the response to scanner noise this could for instance be achieved by including gradient switches without RF-pulses (Bandettini et al., 1998; Elliott et al., 1999) or prerecorded scanner noise stimuli in the paradigm. Although this would lead to an interesting alternative experimental setup, in this study all scanner noise was generated by actual acquisitions.

2.4.3 Interaction magnitudes

A marked increase in the activation volume (i.e., the number of active voxels) from the initial linear model to the subsequent non-linear model could be noted, whereas the mean activation level (i.e., the response amplitude β) remained approximately the same. The selection of active voxels was based on statistical Student T -values that depended not only on the amplitudes of the activation, but also on their respective residual variances. The increase in the number of active voxels indicates that the additional interaction term in the model explains a substantial amount of this residual variance, suggesting that non-linear interactions play an important role.

The importance of interaction effects was directly confirmed by the fact that the amplitude of the interaction term in the model took on values that significantly deviated from zero. However, there was a difference between the interactions of high- and low-frequency tones with scanner noise: the interaction was less strong for the high tone, both in absolute sense (i.e., with regard to γ) and in relative sense (with regard to the ratio γ/β). This was also

reflected by the fact that the interaction amplitude for the low frequency tone significantly deviated from zero, whereas for the high tone it did not. Such differences can be related to the fact that the frequency of the high tone corresponded with a region in the spectrum of the scanner noise that contained relatively little acoustic power (Fig. 2.2). Because there are indications that at least the primary processing stages of the auditory system are organized in a tonotopic fashion (Wessinger et al., 1997; Schönwiesner et al., 2002; Talavage et al., 2004b), it is reasonable to assume that the scanner noise was mainly processed in different regions of the primary auditory cortex than the 4.70-kHz stimuli. This spatial separation of the involved processing centers would reduce the importance of interaction effects. In contrast, the 730-Hz stimulus coincided with the primary peak in the scanner spectrum. Therefore, any areas that were involved in the processing of the low-frequency tones were also likely to respond to scanner noise, thus enhancing interactions (Scarff et al., 2004; Talavage et al., 2004a).

The aforementioned presence of tonotopic gradients is not strictly necessary if the non-linear interactions occur on a neural or metabolic level; neurons showing frequency-selectivity without being organized in a spatial pattern suffice to explain that stimuli of different frequencies show less interaction. However, the typical interval between tone stimuli and scanner noise amounted to several seconds. Thus, it is unlikely that the observed negative interactions are related to neural non-linearities, and consequently they must be the result of accompanying slow vascular mechanisms. Since veins drain considerable regions of brain tissue containing numerous neurons (Turner, 2002), the presence of tonotopic gradients on the scale of several millimeters is necessary to explain the difference in importance of interaction effects for different tone frequencies.

The mentioned vascular non-linearities can possibly originate from physiological limitations; the increase in the diameter of cerebral veins and venules will be limited, and therefore some saturation of the vascular dilatation can be expected. However, other mechanisms that depend on both Δt and TR should possibly be considered as well. For instance, the inflow of fresh saturated blood between scans into the acquired slice can affect signal levels. Inflow enhancement should lead to a signal increase for the lower values of TR (say, ≤ 4 s), in combination with Δt -values close to the peak in the hemodynamic response (~ 4 s). Since the observed interaction effects conversely have a negative magnitude, the decrease in response signal due to interactions in our experiment is incompatible with inflow enhancement and most likely arises from physiological limitations in vasodilatation

processes. Moreover, inflow mechanisms do not confound the comparison of interactions for different stimulus frequencies.

2.4.4 Interaction shapes

Two different types of models were applied to complement each other. On the one hand, a regression model is very suitable to determine the amplitude of responses and interactions, and their statistical significance. However, it requires prior knowledge about the shape of the response and interaction functions. On the other hand, the segregation model was set up to investigate the functional shape of the interaction term. However, this model requires prior input about which voxels to consider active. Both models agreed that the magnitude of the interaction was negative and that interaction effects had a more significant effect for the low frequency tones. However, small differences in the absolute response magnitudes did exist. Regarding the shape of the interaction, the segregation model revealed some irregularities, especially for the low frequency stimulus.

The differences in both the quantitative and qualitative outcome of both models can possibly be explained by two arguments. First, the segregation model is rather susceptible to noise and artefacts. For each combination of Δt and TR, only a small subset of all data is used to sample the interaction (whereas in the regression model all data are used simultaneously to estimate the parameter γ). In particular, any deviations in the determination of the TR-dependent baseline (i.e., the function $I_0(\text{TR})$) will lead to corresponding systematic errors in the calculation of the remaining signal due to the interaction term H_{interact} (Eq. 2.2). Second, the shape of the regression interaction term as it was adopted based on a simple model may be too simple. More complex interaction types might be necessary to describe the data accurately, possibly taking into account non-linearities arising at multiple levels of the auditory processing chain.

To quantify the agreement between the proposed regression term and the actual interaction, the Pearson correlation coefficient was determined. Also, since Eq. 2.5 is completely separable and symmetrical with respect to its arguments, quantitative measures for both features were computed for the actual data. The correlation was reasonably strong and very significant; the interaction was predominantly symmetrical in nature, and highly separable. However, the irregularities in the low-frequency tone interaction term decreased the correlation, symmetry and separability as compared to the high-frequency tone. Especially the presence of a secondary extremum in the interaction term was obvious for the 730-Hz stimuli (Fig. 2.7). The

corresponding timing parameters $\Delta t = 8$ s and $TR = 4$ s indicate that the preceding scan took place 4 s after the preceding tone presentation, such that the hemodynamic response to the preceding tone would have reached its maximum at the time of the scanner noise. It might be that under those conditions the response to the scanner noise diminished because the blood oxygenation already was at an elevated level. However, it remains unclear why a similar mechanism would not play a role in the case where for instance the preceding tone and scan were reversed ($\Delta t = 4$ s and $TR = 8$ s). Apparently, tone stimuli and scanner noise are not completely interchangeable with regard to the behaviour of stimulus interactions.

Although a conclusive explanation for the observed irregularities in the interaction term is hard to give, judging from the degree of separability, symmetry and correlation, interactions certainly were more intricate for the low frequency tone than for the high frequency tone. This may again be related to the tonotopic organization of the auditory cortex. Interactions between spectrally similar sounds may be facilitated on a neural, metabolic and vascular level. In contrast, interactions between differing sounds may depend more on slowly occurring vascular mechanisms, giving rise to a larger smoothness of the response and interaction terms.

2.4.5 Implications for sparse paradigms

Sparse scanning paradigms can typically be characterized by timing parameters in the order of $\Delta t = 5$ s and $TR \geq 10$ s. An important disadvantage of such paradigms is their data-taking inefficiency due to the large gaps between subsequent acquisitions. One goal of this study was to gain insight whether the efficiency could be improved by shortening the TR, allowing more acquisitions to be performed in a given time span. For sparse paradigms, Δt will not exceed TR, and for the regime $\Delta t < TR$ the interaction terms behave quite similarly for both tone frequencies (Fig. 2.7). When the TR is decreased to ~6 s while setting $\Delta t = 4$ s, interaction effects can reach magnitudes in the order of 0.5 %, decreasing the response amplitudes by almost a factor two. Such signal decreases will likely outweigh the beneficial effect of an increased acquisition density on the signal error.

From the characteristics of the interaction term one could erroneously conclude that the interaction effects are small not only for high TR, but also for very small TR. This would lead to the conspicuous conclusion that in sparse paradigms optimal results would be achieved by making the scan interval either large (e.g., 10 s) or very small (2 s). However, our experiment was designed such that the influence of only one single previous scan was

investigated. In practice a large number of preceding scans will be present that will all contribute to the background activation. For instance, in regular acquisition schemes, if $TR = 2$ s, previous scans will also have occurred at intervals of 4, 6, 8 and 10 s before an acquisition. These will all interfere with the stimulus response, accumulating in stronger combined interaction effects.

In practice, a trade-off has to be made between a decrease in interaction effects for large TR , and a decrease in the error level when more acquisitions can be averaged at small TR . Although care should be taken in generalizing our results to other stimulus types and intensities, a noticeable decrease in signal strength should be expected when decreasing the scan interval below 10 s.

§2.5 Conclusion

In summary, we have shown that on top of direct hemodynamic responses, interactions between different stimuli play an important role on similar time scales as the hemodynamic response itself (~ 10 s). As a result, saturation effects tend to decrease the response to auditory stimuli if background scanner noise is present. Non-linear interactions seem to be caused dominantly because of physiological limitations at a vascular level. In accordance with the tonotopic organization of the auditory cortex, interactions are more significant when the spectral content of the stimuli is similar. Although modelling can be used to characterize these interactions, irregularities are prominent in the interaction term especially for interactions between similar sounds, suggesting that more complex mechanisms also contribute to the apparent non-linearities.

§2.6 Appendices

2.6.1 The interaction shape according to a non-linear model

In order for interactions to occur between stimuli, non-linearities must be present. In Fig. 2.8, a basic model is outlined consisting of an initial linear stage followed by the application of a single non-linear function f . Such a model would for instance be natural when the initial neural and metabolic responses are treated in a linear fashion, and the dominant non-linearities in

Figure 2.8 Schematic representation of a simple non-linear response model. The above model is able to account for non-linear behaviour in a voxel's BOLD response to two types of input stimuli $s_1(t)$ and $s_2(t)$. The experimental pure tone stimuli are represented by $s_1(t)$; scanner noise is represented by $s_2(t)$. The model consists of initial linear stages similar to those in traditional fMRI response models, followed by the application of a single non-linear function f that models saturation effects to yield the image intensity $I(t)$. A_1 and A_2 represent measure for the response amplitudes to the two separate stimulus types, and h_v , models the hemodynamic impulse response function.

fMRI are assumed to arise from subsequent vascular mechanisms. The function f should approximate the identity function for small enough arguments, but should asymptotically level off at higher input values to account for saturation effects. This model is comparable with that introduced by Friston et al. (Friston et al., 1998b); an important difference is that in this study two different stimulus sources were used, represented by $s_1(t)$ for tone stimuli and $s_2(t)$ for acoustic scanner noise.

A truncated Volterra series expansion for this model would yield

$$\begin{aligned}
 I(t) = & K^{00} + \\
 & + \int_0^\infty d\tau K^{10}(\tau) s_1(t-\tau) + \int_0^\infty d\tau K^{01}(\tau) s_2(t-\tau) +, \\
 & + \int_0^\infty d\tau \int_0^\infty d\tau' K^{11}(\tau, \tau') s_1(t-\tau) s_2(t-\tau')
 \end{aligned} \tag{2.7}$$

where K^{00} represents the baseline level, the Volterra kernels K^{10} and K^{01} equal the impulse response functions from conventional linear theory (i.e., the hemodynamic response functions), and the kernel K^{11} is the lowest-order interaction kernel. Note that this second-order kernel does not necessarily have to be symmetrical with respect to its arguments, in contrast with the case where only non-linearities in response to a single stimulus type are considered. For this model it can straightforwardly be shown that

$$K^{11}(\tau, \tau') = f''(0) \cdot K^{10}(\tau) \cdot K^{01}(\tau') = f''(0) \cdot (A_1 \cdot h_{v,\tau}(\tau)) \cdot (A_2 \cdot h_{v,\tau'}(\tau')), \tag{2.8}$$

where both kernels K^{10} and K^{01} have been modelled by gammavariate functions $h_{v,t}$ (Eq. 2.4). It follows that according to this model the interaction term is completely separable and symmetric with regard to its arguments and proportional to the product of two hemodynamic response functions of the two kernel arguments. The corresponding constant of proportionality is directly related to the non-linearity in the function f , as expressed by its second derivative $f''(0)$. Since in this study all stimuli were short (1.3 s) compared to the typical duration of the hemodynamic response (~ 5 s), the stimulus functions s_i can be approximated by Dirac delta functions; combined with Eqs. 2.7 and 2.8 this leads to the direct conclusion that the interaction term in the regression model (Eq. 2.5) should have the same functional form as the interaction kernel K^{11} (Eq. 2.8).

2.6.2 Quantitative characterization of the interaction shape

The shape of the interaction term H_{interact} was characterized by means of its (1) separability, (2) symmetry and (3) correlation with the model equation (Eq. 2.5).

First, separability expresses whether a function can be written as a product of functions depending on only one of the arguments; in this case $H_{\text{interact}}(\Delta t, \text{TR}) = H_{\Delta t}(\Delta t) \cdot H_{\text{TR}}(\text{TR})$. The separability was quantified by the method of singular value decomposition (SVD) which factorizes the interaction data matrix as a product $\mathbf{M} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^t$. In this expression, $\mathbf{\Lambda}$ is a positive definite diagonal matrix containing the singular values $\lambda_1 \dots \lambda_n$ in decreasing order, and $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices. The degree of separability $\alpha_{\text{sep.}}$ was calculated using $\mathbf{M}'$'s best separable approximation $\mathbf{M}' = \lambda_1 \cdot \mathbf{U}_{11} \cdot \mathbf{V}_{11}^t$, according to

$$\alpha_{\text{sep.}} = \frac{\|\mathbf{M}'\|_{\text{HS}}^2}{\|\mathbf{M}\|_{\text{HS}}^2} = \frac{\lambda_1^2}{\sum_{m=1}^n \lambda_m^2}; \quad 0 < \alpha_{\text{sep.}} \leq 1, \quad [2.9]$$

where $\|\cdot\|_{\text{HS}}$ denotes the Hilbert-Schmidt matrix norm; $\alpha_{\text{sep.}}$ equals the proportion of $\mathbf{M}'$'s total power which is accounted for by its best separable approximation (Depireux et al., 2001; Linden et al., 2003). A result equal to 1 would indicate that the data are completely separable.

Second, the degree of symmetry of the interaction term was expressed by means of the power of the symmetric and the anti-symmetric components of the data matrix $\mathbf{M}$. A measure $\alpha_{\text{sym.}}$ was defined according to

$$\alpha_{\text{sym.}} = \frac{\|\mathbf{M} + \mathbf{M}^\dagger\|_{\text{HS}}^2 - \|\mathbf{M} - \mathbf{M}^\dagger\|_{\text{HS}}^2}{4 \cdot \|\mathbf{M}\|_{\text{HS}}^2}; \quad -1 \leq \alpha_{\text{sym.}} \leq 1, \quad [2.10]$$

where $\mathbf{M}^\dagger$ is the transposed matrix. The value of $\alpha_{\text{sym.}}$ ranges from -1 for perfectly anti-symmetric matrices to $+1$ for perfectly symmetric ones.

Third, Pearson's correlation coefficient R was computed for the values of H_{interact} and those of the model function according to Eq. 2.5 to quantify their mutual agreement (see appendix 2.6.1 for more information about the model). Note that the model function is both separable and symmetric with regard to its arguments.

**INTENSITY,
LOUDNESS, AND
THE LINEARITY OF
ACTIVATION**

Chapter

3

Abstract

In people with elevated hearing thresholds due to sensorineural hearing loss, the perception of sound levels is distorted. The neural correlate of these distortions is not well understood. The purpose of this fMRI study was to relate cortical responses to both the (physical) intensity and (perceptual) loudness of sound. Subjects with normal hearing and subjects with high-frequency hearing loss were included. Stimuli consisted of frequency modulated tones in either a low (0.5-1 kHz) or a high (4-8 kHz) frequency range, with intensities in an equivalent loudness range of 0-70 dB. The activation in the subjects with hearing loss increased significantly more strongly with stimulus intensity than in the normal hearing subjects. This reflects the loudness recruitment phenomenon, characterized by a disproportionate increase in loudness percept with stimulus intensity. In contrast, the rate of activation increase did not differ between both subject groups as a function of loudness. Responses were found to increase predominantly linearly with sound level. Response saturation at the highest intensities was not apparent. However, activation exhibited a steep rise between 0 and 10 dB for the low-frequency stimuli, contributing to a significant non-linearity in the impaired subjects. Our results demonstrate that fMRI activation levels are less related to the physical attributes of a stimulus than to the percept it evokes.

§3.1 Introduction

In spite of extensive research, still relatively little is known about how sound is processed in the brain and how various sound attributes are neurally represented. Questions remain even regarding very basic attributes like sound level.

In people with sensorineural hearing loss, hearing thresholds are elevated over some range of frequencies while uncomfortable loudness levels have not changed notably. This is likely related to defects in the inner ear feedback mechanisms, which are required to achieve good sensitivity to soft sounds (Moore, 1996). In such cases, the relation between loudness perception and sound intensity is modified: a large rise in perceptual loudness can be generated by a relatively small increase in physical sound intensity. This phenomenon is known as loudness recruitment (Martin, 1996; Oxenham et al., 2003), and in the context of functional brain imaging should not be confused with neurophysiological recruitment of neurons or capillaries in the cortex. Loudness recruitment influences the perception of sound, and may for instance also affect the comprehension of speech. As the input to the brain has often permanently changed in people with sensorineural hearing loss, qualitative modifications in the activation characteristics as a function of sound level may occur. Hence, people with cochlear hearing loss are of interest in gaining more knowledge on the cortical effects of sound level as a stimulus parameter and potentially for studying plasticity of the auditory cortex using fMRI (McDermott et al., 1998; Rauschecker, 1999).

Previous functional magnetic resonance imaging (fMRI) studies have systematically investigated sound level dependent activation in normal hearing subjects exclusively. It has been shown that the hemodynamic response signal in the auditory cortex increases with increasing sound level (Jäncke et al., 1998). However, due to the hindrance of acoustic scanner noise as well as the limited signal detection power of fMRI, the determination of sound level dependent activation has often been limited to high sound intensities over 60 dB sound pressure level (SPL) (Mohr et al., 1999; Hall et al., 2001; Bilecen et al., 2002; Mulert et al., 2005) or moderately high sound intensities over 40 dB SPL (Brechmann et al., 2002; Hart et al., 2002; Hart et al., 2003b). Using sparse imaging techniques (Hall et al., 1999) auditory stimuli can be delivered at low intensities with limited interference from acoustic scanner noise (Yetkin et al., 2004b). Thus, it should be possible to

characterize cortical activation levels over a much larger range of sound intensities than has been done before.

The linearity of the hemodynamic response strength as a function of the sound level is an important consideration, for instance for making comparisons between the results of studies that employ stimuli at different levels. Some previous studies have claimed that the relationship between the stimulus intensity and the measured response is essentially linear (Hall et al., 2001), while others report significant non-linearities (Hart et al., 2002) and possibly even non-monotonic relationships with a spatially dependent amplitopic organization (Bilecen et al., 2002). Results also vary according to the way that activation is quantified, e.g. for activation extent vs. level (Hart et al., 2002). However, none of these studies comprise stimulus levels near the auditory threshold. Because non-linear feedback mechanisms may play an important role at low sound intensities especially, it seems necessary to include low stimulus intensities into the study design in order to accurately assess the linearity of the response.

In this study, responses will be characterized in a broad intensity range of 0-70 dB above the hearing threshold for subjects with normal hearing, spanning most of the range that is of daily relevance in life. Also, activation will be studied in subjects with elevated hearing thresholds for an equivalent loudness range. The magnitude of cortical activation will be described both as a function of physical sound intensity and perceptual loudness to provide further insight in the relation between fMRI activation, the strength of the stimulus and its percept. Moreover, the linearity of the increase in response level with sound intensity and loudness will be assessed.

§3.2 *Materials and Methods*

3.2.1 *Subjects*

Ten subjects with normal hearing were recruited (8 male, 2 female; age 22-61 years, mean 37 years). In addition, ten subjects (9 male, 1 female; age 57-70 years, mean 64 years) with bilaterally symmetrical moderately severe high-frequency sensorineural hearing loss were selected from a patient database of the audiology department at the Maastricht University Hospital (Maastricht, the Netherlands). Patient selection criteria comprised left and right hearing thresholds better than 20 dB hearing level (HL) for tone frequencies equal to 250, 500, and 1000 Hz and thresholds of 40-70 dB HL for

frequencies equal to 4 and 8 kHz, with the average difference between the left and right ear thresholds not exceeding 10 dB, and air-bone gaps (if available) not exceeding 10 dB. The selected subjects were invited to participate in the study on the basis of written informed consent, in approved accordance with the requirements of the medical ethical committee of the institute. To assess current hearing thresholds, all subjects were re-examined prior to the experiment by an audiologist using standard pure tone audiometry in a sound booth (Katz, 2001).

3.2.2 Auditory stimuli

The auditory stimuli consisted of frequency modulated (FM) pure tones with sinusoidal modulations on a logarithmic frequency scale. Modulations had a rate of 5.0 Hz and spanned a spectral range of 0.5-1.0 kHz or 4.0-8.0 kHz. Stimulus waveform data were generated digitally in advance and were stored in the form of standard 16-bit uncompressed sound files with a sampling frequency of 44.1 kHz. A spectrally dynamic stimulus type was chosen because such stimuli are expected to elicit strong responses in both primary and secondary areas of the auditory cortex (Thivard et al., 2000). At the same time, FM-tones can be constructed to have spectral power in a limited range of frequencies only, such that stimuli could be generated that coincided with either a region of normal or impaired hearing in the patient group.

In a silent room, left and right ear hearing thresholds for the FM-tone stimuli were determined, and a loudness matching task was performed to obtain loudness-matched low- and high-frequency stimulus pairs as follows. 1.0-s FM-tone fragments with low and high frequency ranges were alternately presented in a continuous fashion. The low-frequency stimuli served as reference stimuli and were binaurally presented at a constant level. This level was chosen equal to a multiple of 10 dB above the individual hearing threshold, with a range of 10 to 70 dB. Subjects were instructed to adjust the presentation level of the high-frequency FM-tones by means of a keyboard such that both alternately presented stimuli had equal perceptual loudness levels. This matching task was performed twice for the set of all stimuli, once in order of increasing intensity (i.e., from 10 to 70 dB reference intensity in 10-dB steps) and once in reverse order. The resulting intensities for the high-frequency FM-tones were averaged between both series. The outcome of the entire procedure was reproducible within subjects with an approximate accuracy of 5 dB.

The subject-dependent perceptual loudness of the FM-tones was quantified by an equivalent loudness (EL) scale analogous to the phon scale for pure tone stimuli (Moore, 2003). For the low-frequency reference FM-tones the EL was by definition equal to the intensity above the individually determined threshold. For the high-frequency FM-tones the EL equaled the intensity of the low-frequency reference FM-tone with equal perceptual loudness. For example: if a low-frequency FM-tone at 60 dB above the corresponding individual threshold was perceptually matched in loudness with a high-frequency FM-tone at 40 dB above threshold, then these tones had an intensity of 60 and 40 dB sensation level (SL) respectively, while both tones were said to have a loudness of 60 dB EL. The threshold stimuli were assigned an SL and an EL of 0 dB. Using these definitions, the previously described procedure served to determine corresponding sets of low- and high-frequency FM-tone stimuli with loudness levels of 0-70 dB EL in steps of 10 dB EL.

Loudness recruitment in the high frequency range was quantified in all subjects individually by relating the required increase in stimulus intensity (Δ SL) to the corresponding increase in loudness (Δ EL) using linear regression.

3.2.3 Functional imaging sessions

Subjects were placed in supine position in the bore of a clinical 1.5-T MRI system (Philips Intera, Philips Medical Systems), which was equipped with a standard quadrupolar head receiver coil. A sagittal 3D T_1 -weighted fast-field echo scan was acquired for anatomical orientation (repetition time TR 10.20 ms; echo time TE 3.5 ms; flip angle 15°; matrix 256×256×75; voxel dimensions 1.0×1.0×2.0 mm). On these high-resolution images a central imaging plane was positioned approximately parallel to the Sylvian fissure of both hemispheres, containing the superior temporal gyri of both temporal lobes. An additional 3D T_1 -weighted fast-field echo scan (TR 10.34 ms; TE 3.5 ms; flip angle 15°; matrix 256×256×12; voxel dimensions 1.0×1.0×2.0 mm) was acquired parallel to this imaging plane and served as anatomical reference. The functional scans were acquired covering the same brain volume, and consisted of a dynamic series of 2.5-s single-shot T_2^* -sensitive echo planar imaging (EPI) sequences with twelve 2.0-mm thick adjacent slices (TR 10 s; TE 50 ms; flip angle 90°; matrix 192×192; field of view 192×192 mm).

A sparse clustered volume scanning paradigm was employed to reduce the influence of acoustic scanner noise (Hall et al., 1999; Talavage et al., 1999). Functional acquisitions were performed every 10 s, leaving 7.5-s gaps of

scanner silence between consecutive scans during which auditory stimuli were presented (see Fig. 3.1). The stimuli were separated from the subsequent scan by a silent period of 3.0 s and consisted of pairs of similar 2.0-s FM-tones. Both tones had equal frequency characteristics but marginally different intensities: the loudness of one of the tones was a multiple of 10 dB EL, while the intensity of the other randomly differed by 3 dB. This 3-dB difference was just large enough to be discernible, but was small in comparison with the 10-dB EL differences between the employed stimuli across the entire experiment. Subjects were instructed to indicate by means of two fiber-optic button devices in both hands which of the two FM-tones appeared loudest. This active listening task was performed to enhance the recorded brain responses (Hall et al., 2000), and the results were used only to monitor the subjects' performances and to confirm that they understood the task and stayed alert during the fMRI session.

For each subject the functional imaging session consisted of six runs of 32 volume acquisitions each, excluding initial start-up scans. Stimuli were delivered by means of modified electrostatic head phones (Sennheiser HE 60, Sennheiser) built into a noise protection headset that were connected to a PC setup equipped with a sound card with 16-bit DAC output capability (Sound Blaster Live! Platinum 5.1, Creative). In the odd runs high-frequency FM-tones were presented, while in the even runs low-frequency FM-tones were employed. In each run, eight different stimulus intensities corresponding with perceptual loudness levels of 0-70 dB EL in 10-dB steps were presented in pseudo-random order. This allowed each of the stimulus intensities to be presented four times per run. For 4 out of 20 subjects deviations occurred between the stimulus presentation levels during the loudness matching task and during the functional imaging sessions due to minor differences in hardware settings. For these subjects the actual presentation levels during functional imaging were taken into account in the fMRI data analysis that followed.

3.2.4 Data analysis

In the image analysis we made use of the MatLab programming environment (The MathWorks Inc.) and pre-processing routines from the SPM2 software package (Wellcome Department of Cognitive Neurology). The functional image volumes were corrected for motion effects using 3D rigid body transformations. To improve signal-to-noise ratio characteristics, spatial smoothing was performed using an isotropic 5-mm Gaussian kernel. For each of the functional runs, linear drifts of the scanner baseline signal were corrected for by means of linear regression. Images were normalized

by dividing the signal for each individual voxel by its mean across all 192 functional acquisitions to express the signal variations as a percentage relative to the mean voxel signal.

For each voxel, the signal across all acquisitions was correlated with the corresponding stimulus loudness levels. For every subject the 100 voxels with the most significant positive correlation coefficients were selected to form a reference set of voxels that responded most strongly to the presented stimuli. Signals were averaged over these voxels and over all acquisitions corresponding with a certain stimulus condition to obtain average signal levels (S_i) corresponding with all 16 different stimulus conditions (i.e., 8 different loudness levels for each of the two FM-tone frequency ranges). The activation level is considered to be a more reliable and robust measure of activation than activation extent (Mohr et al., 1999; Hall et al., 2001), especially in experiments with many stimulus conditions.

The signals corresponding with the low- and high-frequency FM-tones were separately analyzed as a function of the eight different stimulus intensities and the equivalent loudness levels (collectively referred to as L_i) using linear regression according to a quadratic model

$$S_i = a + b \cdot L_i + c \cdot [L_i^2]_{\text{orth}} + e_i, \quad [3.1]$$

by minimizing the sum of squares of the error terms e_i . The fMRI baseline signal level was not calculated as the average signal at 0 dB, but was instead set equal to the fitted value of the offset a in Eq. 3.1. This baseline level was subsequently subtracted from the measured signal values to obtain the activation levels (A_i) corresponding with each of the stimulus conditions: $A_i = S_i - a$. The described method to determine the baseline signal provides greater accuracy in the calculated activation levels by reducing inter-subject variability due to systematic errors in the subtraction. However, as a result the activation level measured at 0 dB may deviate from 0 % due to non-linear signal variations.

The quadratic regressor (L_i^2) was orthogonalized with respect to the lower order terms to avoid interference with the estimation of the coefficients a and b . Therefore, the regression coefficient b quantifies the average linear increase in activation with increasing stimulus levels over the entire range of stimulus conditions, rather than the local slope at 0 dB. The coefficient c measures the overall strength of quadratic non-linearities.

Eq. 3.1 was applied to estimate the coefficients a , b and c in all subjects individually. Single sample Student T -test statistics were used to determine the significance of the linear and quadratic coefficients b and c across the two subject groups. Independent samples T -tests were performed to calculate the

Figure 3.2 Subject group hearing thresholds. Hearing thresholds for subjects with normal hearing (NH) and impaired hearing (IH) were measured in a frequency (f) range of 0.25-8.0 kHz using pure tone audiometry. Thresholds were normal (i.e., better than 20 dB hearing level, HL) for low-frequency stimuli (0.25-1.0 kHz) in both subjects groups. For high frequencies (4.0-8.0 kHz), thresholds in the impaired subjects were highly elevated as compared to the normal hearing subjects. The error bars indicate the group standard deviation around the mean.

significance of differences between the groups of normal hearing and impaired subjects.

§3.3 Results

3.3.1 Audiometry

Pure tone audiometry was performed prior to the functional sessions. For the normal hearing subjects, air conduction thresholds were determined at 7 ± 5 dB hearing level (HL), averaged over both ears and over frequencies of 0.25-8.0 kHz (mean $\pm$ standard deviation across subjects). Because in these subjects hearing was normal, bone conduction thresholds were not determined. In the subjects with hearing loss, average air conduction thresholds equaled 13 ± 8 dB HL for frequencies of 0.25-1.0 kHz, and 66 ± 11 dB HL for frequencies of 4.0-8.0 kHz (see Fig. 3.2). The absolute difference in threshold between the left and right ears averaged over all frequencies equaled 8 ± 4 dB. The bone conduction thresholds equaled 11 ± 7 dB HL at 0.25-1.0 kHz, and 59 ± 7 dB HL at 4.0 kHz (for 8.0 kHz bone conduction thresholds were not determined). These results confirmed that all impaired subjects suffered from moderately severe symmetrical bilateral

Figure 3.3 Stimulus intensity vs. loudness in subjects with normal and impaired hearing. The intensity above threshold (sensation level, SL) of high-frequency FM-tones is plotted as a function of their equivalent loudness (EL) for all subjects. The equivalent loudness equals the intensity of a low-frequency reference FM-tone with equal perceptual strength. In the impaired subjects, the increase in stimulus intensity that is required to evoke a certain rise in perceptual loudness is smaller than in normal hearing subjects. This phenomenon is commonly referred to as loudness recruitment, and is characteristic for sensorineural hearing loss. The gray band indicates the 95-% confidence interval of the quadratic polynomial fit through the data points of all subjects collectively.

sensorineural hearing loss for high frequencies (≥ 4 kHz), while hearing thresholds were normal for low frequencies (≤ 1 kHz).

3.3.2 Loudness vs. Intensity

The relationship between the loudness and the intensity of the high-frequency stimuli according to the results of the loudness matching task is displayed in Fig. 3.3. Results are shown for the individual subjects. In this and the following figures, the 95-% confidence interval for the quadratic polynomial fit to all data points collectively is shaded to indicate a group trend. The intensity and loudness of the low-frequency stimuli are equal by definition and will therefore not be discussed.

For all individual subjects, the intensity increased significantly with the loudness and vice versa ($p < 0.001$). For the group of normal hearing subjects, the high-frequency intensity increased with loudness by 0.75 ± 0.06 dB/dB ($\Delta SL/\Delta EL$, mean $\pm$ standard error). For the impaired subjects this

increase was significantly smaller ($p < 0.01$), and equaled 0.55 ± 0.03 dB/dB. For both groups the coefficient of the quadratic term was insignificant across subjects ($p > 0.25$).

3.3.3 Low-frequency activation

Fig. 3.4 displays the average activation level of the 100 most active voxels for each subject as a function of the low-frequency stimulus intensity, which equals the loudness for this stimulus frequency. In all individual subjects the activation increased significantly with the stimulus intensity ($p < 0.05$). The average rates of increase in activation (b) are listed in Table 3.1 for both subject groups. The increase rates did not differ significantly between the two groups.

Table 3.1 Rates of increase in activation with sound level.

Stimulus frequency	Sound level measure	<i>b</i> [10^{-3} %/dB]		<i>p</i> -value
		NH	IH	
0.5-1.0 kHz	SL, EL	24.3 ± 3.3	28.8 ± 2.0	0.25
4.0-8.0 kHz	SL	20.5 ± 2.3	37.2 ± 4.8	0.005
	EL	14.8 ± 1.7	20.4 ± 2.7	0.10

The rate of increase in activation (*b*, mean ± standard error) as a function of the physical intensity (sensation level, SL) or perceptual loudness (equivalent loudness, EL) of the two stimulus types is listed in subjects with normal hearing (NH) or impaired hearing (IH). For the low frequency stimulus, EL equals SL by definition. The null hypothesis that both subject groups show equal increase was tested using an independent samples *T*-test; increase rates differed significantly only ($p < 0.05$) for high-frequency stimuli as a function of stimulus intensity.

For the threshold stimulus condition (0 dB) the response level across subjects was observed to systematically lie below the 0-% baseline level for both subject groups. This does not represent inhibition or deactivation at 0 dB stimulation: because Eq. 1 was used to estimate the baseline level on the basis of the entire data set, this indicates only that the signal at 0 dB tends to lie below the linear trend lines of individual subjects. At low intensities (0-10 dB), a steeper increase of the activation with stimulus level was observed than at moderate and high intensities (20-70 dB). In relation to this finding, the coefficient of the quadratic term was significant for the impaired subject group ($p < 0.01$) and equaled $c = (-0.56 \pm 0.11) \cdot 10^{-3} \text{ \%/dB}^2$. However, it was insignificant for the normal hearing subject group ($p > 0.1$).

3.3.4 High-frequency activation

The activation levels in response to the high-frequency stimuli are shown in Fig. 3.5. In all individual subjects the activation increased significantly ($p < 0.05$) with stimulus intensity (Fig. 3.5a). Also, in the impaired subjects the activation increased significantly more strongly as a function of stimulus intensity than in the normal hearing subjects (see Table 3.1). Similarly, Fig. 3.5b displays the activation levels as a function of the equivalent loudness. In all individual subjects the activation increased significantly with loudness ($p < 0.05$). However, in contrast with the findings as a function of stimulus intensity, the difference in *b* between both groups if calculated as a function of stimulus loudness was insignificant (Table 3.1).

In contrast with the low frequency FM-tones, activation data clustered around the baseline level ($A = 0 \%$) for the 0-dB condition, and no excessive

increase in activation was observable between 0 and 10 dB as compared to that at moderately low to high stimulus levels of 20-70 dB (compare Fig. 3.5 with Fig. 3.4). The coefficient of the quadratic term was not significant for

either subject group and for either stimulus intensity or loudness ($p > 0.1$).

§3.4 Discussion

Using fMRI, we determined the magnitude of the cortical response to FM-tones in subjects with normal hearing as well as subjects with high-frequency sensorineural hearing loss. The activation was studied both in relation to the sound intensity and loudness. In addition, the large range of sound levels (0-70 dB) allowed us to assess the linearity of the response more accurately than has been done before.

3.4.1 Intensity, loudness, and activation

In fMRI studies on the central auditory system, often no explicit distinction is made between the intensity and loudness of stimuli. Intensity refers to the objective physical characteristics of sound, while loudness refers to the subjective strength of the percept it generates. When normally hearing subjects are involved exclusively this is hardly an issue, because then the relationship between intensity and loudness is reasonably well defined and documented (Moore, 2003). In such cases, sound intensity is usually preferred over loudness as the experimental parameter because it can be controlled much more easily and objectively. However, the distinction between sound intensity and loudness should ideally be accounted for if types of stimuli are used that are perceived with different loudness even though their intensities are the same. This is for instance the case for stimuli with different frequencies, different bandwidths (Hall et al., 2001), monaural and binaural stimuli (Jäncke et al., 2002a), or stimuli that interact differently with the acoustic scanner noise (Scarff et al., 2004). Differences between intensity and loudness can become considerable if subjects are studied with disturbed sound perception, for example due to hearing loss or tinnitus. In this experiment, subjects with normal hearing as well as with high-frequency sensorineural hearing loss were included to differentiate between intensity and loudness as the parameter for sound level.

For the low-frequency stimuli, both groups of subjects displayed normal hearing thresholds and the hemodynamic response did not show a significant difference in the rate of increase with sound level. However, for the high-frequency stimuli, hearing thresholds in the impaired subjects were worse than in those with normal hearing. In addition, loudness recruitment was present: the perceived loudness of high-frequency stimuli increased

more strongly with intensity in impaired subjects than in normally hearing subjects. We also found that the magnitude of cortical activation increased more strongly with intensity in impaired subjects than in normally hearing subjects. This suggests that the magnitude of cortical activation is more closely related to stimulus loudness than to stimulus intensity. Indeed, in spite of the severely disturbed sound perception in the impaired subjects, the increase in cortical activation was not significantly different between both subject groups if expressed as a function of perceptual loudness.

While loudness recruitment is a symptom that is commonly associated with inner ear impairment, the corresponding neural mechanisms are poorly understood. It has been reported that the afferent signals in the auditory nerve fibers do not provide a simple representation of the excitation in the basilar membrane in people that display loudness recruitment (Heinz et al., 2004). Therefore, it remains unclear how the loudness percept is generated by the brain from available input signals, especially in pathological conditions. The present study is the first to document the cortical responses related to loudness recruitment in humans using fMRI. Not only does the stimulus percept grow disproportionately strong in people with loudness recruitment, but previous studies using magnetoencephalography (MEG) have also suggested that brain activity increases more quickly with stimulus intensity in such individuals (Morita et al., 2003). Our results agree with these findings and confirm that cortical activity is more strongly linked to the perceptual loudness of sound than to its intensity. This forms objective evidence that the fMRI signal can more or less reflect the strength of the subjective individual percept, which is of utmost importance for the interpretation of fMRI results in general.

3.4.2 Linearity at moderate and high intensities

The cortical response behavior that was found in this study was dominantly linear as a function of stimulus level. For the FM-tone stimuli with a frequency range of 4-8 kHz, non-linear (quadratic) contributions to the response characteristics were not significant for either the subjects with normal or impaired hearing, irrespective of whether the stimulus intensity or its equivalent loudness was used as the independent variable. For the FM-tone stimuli in a range of 0.5-1 kHz, quadratic non-linearities were significantly present in the impaired subjects, but not in the normal hearing subjects.

Neither for the low- or the high-frequency FM-tones did the response signal seem to reach an obvious upper limit at high intensities. Although at some

point saturation effects should become dominant, this point evidently lies more than 70 dB above threshold. This agrees with previous studies that suggest that saturation doesn't occur below levels of 80 to 90 dB SPL (Jäncke et al., 1998; Mohr et al., 1999; Hart et al., 2002). Furthermore, our general findings indicate that from approximately 20 dB above threshold the response magnitude as a function of stimulus level is essentially linear. This is in agreement with reports in literature (Hall et al., 2001) and we were able to extend such findings towards lower stimulus intensities.

Although fMRI is a method with relatively good spatial resolution as compared to other neuroimaging techniques (e.g. MEG), it is not justified to extend our conclusions concerning response linearity to the level of individual neurons. It is possible for a population of neurons to respond linearly as a whole, even if neurons show highly non-linear responses individually. Also, because fMRI is an indirect measure of activation that represents metabolic rate more than neuronal firing rate, a distinction between firing activity, synaptic excitation and synaptic inhibition cannot easily be made (Buxton et al., 2004; Logothetis et al., 2004a). Therefore, the relation between the measured fMRI response and neuronal activity is not trivial.

The fact that the outcomes of analyses are inevitably dependent upon the criteria that are used to select active regions of interest forms another methodological limitation. In this study, per subject only the 100 most active voxels were included on the basis of the correlation of their signals with stimulus loudness. This may prevent some voxels with atypical response behavior from being included (for instance voxels that respond strongly to moderate intensities but only weakly to both low and high intensities, as might occur in the presence of an ampliotopic organization). This might introduce a weak bias towards the selection of voxels with linear response behavior. However, our conclusions would still apply to at least a large proportion of highly active voxels. Also, because stimulus intensity and loudness are essentially linearly related in both subject groups, such effects would be equally important for intensity and loudness. Therefore, voxel selection criteria should introduce no bias in comparisons between the subject groups or sound level measures.

3.4.3 Non-linearity at low intensities

In spite of the predominantly linear increase of activation at moderate and high intensities, our results (and those of others) leave open the possibility that non-linearities play a role at extremely low stimulus levels (0-10 dB). It

was noted that for the low frequency FM-tones the fMRI signal in response to stimuli near threshold (i.e., 0 dB) was relatively low for both the normal hearing and the impaired subjects. Between 0 and 10 dB the signal showed a pronounced increase, followed by a less strong monotonic rise at moderate and high stimulus intensities. This was observed for the normally hearing as well as the impaired subjects. In contrast, for the high frequency FM-tones the fMRI signal levels near the auditory threshold clustered around the 0-% baseline level, and non-linearities at low stimulus levels in particular were not apparent.

A number of possible causes may give rise to non-linearities. First, non-linear feedback mechanisms play a crucial role in the mechanics in the inner ear (Oxenham et al., 2003). Efferent connections that innervate the outer hair cells in the cochlea induce hair cell motility that actively influences the transfer characteristics of the system as a whole. Such mechanisms are necessary to obtain strong frequency selectivity and good sensitivity to weak sounds. The non-linear response behavior that arises in the cochlea can perpetuate towards the central auditory system, leading to non-linear response characteristics in the auditory cortex. Second, neural processing is a highly non-linear type of activity. Individual neurons inherently show non-linear response behavior (London et al., 2005). Neural systems as a whole contain numerous excitatory and inhibitory feedback and feedforward connections that can result in highly specific responses to certain input signals. Also variations in the state of the individual neurons and of the entire system allow complex types of behavior to emerge (e.g., adaptation, attention, memory, and priming). Third, fMRI measures blood oxygenation level dependent (BOLD) hemodynamic responses that are only an indirect measure of the underlying neural activity. The metabolic and vascular dynamics upon which fMRI depends are not yet fully understood, but can surely display non-linear behavior (Buxton et al., 2004). The fact that the dynamic range of various physiological parameters (e.g., metabolic conversion rate, blood vessel dilatation, or capillary recruitment) is limited, implies that at some point saturation of the measurable response should occur. Finally, the presence of ambient noise during the functional imaging sessions might play a role. Although in this experiment a silent acquisition protocol was used (Hall et al., 1999), some background noises related to the fMRI environment were inevitable (e.g., those caused by the helium pump). Weak background noises are not likely to interfere significantly with loud stimuli, but in spite of the noise protecting headset they might influence the responses to stimuli near the auditory threshold.

We argued that non-linear behavior, if present, occurs especially in a limited range of low stimulus levels, and manifests itself as a relatively strong initial increase in activation followed by a gradual rise at higher levels. Significant quadratic contributions to the overall response characteristic seem to originate mainly from non-linearities near threshold instead of saturation effects at high intensities. This is well compatible with the provided explanations that non-linearities potentially arise from cochlear mechanics or from neural processing. However, it seems to exclude that metabolic and vascular neurophysiological mechanisms play an important role in this respect (at least across the behaviorally relevant range of 0-70 dB), since these would mainly be expected to lead to saturation effects at high intensities. Furthermore, the presence of background noises would mask the weakest experimental stimuli, and result in a smaller rate of increase in activation at low levels. This contrasts with our findings.

The observed type of non-linearity near threshold might be related to the type of stimulus that was employed in this study. FM-tones elicit responses in cortical regions that are responsive to spectral dynamics, for instance areas located in secondary auditory cortex (Thivard et al., 2000). Such areas can show tuning to complex sound parameters with responses depending less strongly upon sound level (Hart et al., 2002). The response characteristic will then show a more or less on/off type of behavior, with relatively strong responses being attained already for low-intensity stimuli as long as sufficiently complex spectral dynamics are discernible. Then, if the activation is averaged across multiple cortical regions containing both neurons that show the aforementioned on/off responses as well as neurons that show a more gradual increase in firing rate with stimulus level, an overall response behavior may be obtained that is linear at medium and high stimulus levels, but that displays a steeper increase just above threshold.

However, such arguments do not explain why similar effects are absent in the high-frequency data. Experiments comprising various stimuli that vary in spectral complexity may shed more light on this issue.

§3.5 Conclusion

In summary, we found that the perceptual loudness of stimuli, in contrast with sound intensity, related strongly to cortical activation even when comparing groups of subjects with vastly different sound perception. This

suggests that fMRI activation can be interpreted as a correlate of the subjective strength of the stimulus percept. Furthermore, it was found that cortical responses increased predominantly linearly over a large range of stimulus levels. Non-linear contributions were mostly insignificant and, if present, seemed to be caused by the presence of a relatively strong rise in activation just above threshold.

SENSITIVITY TO SPECTRO- TEMPORAL MODULATIONS

Chapter

4

Published in modified form:
Neuroimage 2003;20(1):265-275
doi:10.1016/S1053-8119(03)00258-1

Abstract

Functional MRI was used to investigate the characteristics of the human cerebral response to dynamic ripples. Dynamic ripples are sound stimuli containing regular spectro-temporal modulations, which are of major importance in speech processing; however, in contrast to speech, dynamic ripples can be characterized fully by a limited number of parameters. Extensive activation consisting of multiple separate regions was found bilaterally in the auditory cortex, particularly along the Heschl's gyri. This agrees with the presence of a structural cortical subdivision into functional fields. The level and the volume of activation were measured, and correlated highly ($R^2 = 0.97$). Both measures depended strongly on the spectral density, temporal frequency, and amplitude of the modulations, and matched the perceptual discernibility of the spectro-temporal modulations. The largest responses occurred for parameter values near the optimal human sensitivity. The drift direction of the modulations did not influence the activation. No quantitative differences were found between the two hemispheres. Average brain activation levels proved to be separable with regard to the spectral density and temporal frequency of the modulations. Topographic mappings of the modulation density and frequency onto the cortical surface were shown, approximately in posterolateral-to-anteromedial and lateral-to-medial direction respectively. Posterolateral regions were most sensitive to spectro-temporal features at a scale similar to phonemes. Anteromedial regions, however, were also relatively sensitive to smaller-scale acoustic features. This spatially dependent sensitivity suggests a functional topographic and hierarchical organization of the auditory cortex.

§4.1 Introduction

In this study we focus on sound properties related to speech, the processing of which obviously is an important function of the human brain (Belin et al., 2000; Scott et al., 2000; Wise et al., 2001; Belin et al., 2002; Jäncke et al., 2002b; Wong et al., 2002). In systematic studies, the structural complexity of vocal sounds is a significant problem when defining their exact properties (Hickok et al., 2000). Therefore, simpler sound signals have been derived that can be defined with mathematical exactitude, but that still contain some essential characteristics of vocal sounds. An important category of such signals comprises modulated noise. The most basic of these modulated noise signals contain either temporal amplitude modulations (Wakefield et al., 1990; Giraud et al., 2000; Liegeois-Chauvel et al., 2004) or static spectral modulations also known as rippled noise (Calhoun et al., 1998; Versnel et al., 1998; Supin et al., 1999). However, modulations in the spectrum of speech are dynamic. Similar dynamic behaviour can be incorporated into modulated noise by combining temporal and spectral modulations, resulting in a spectral ripple drifting in time along the frequency axis (Fig. 4.1). Such

signals have a similar spectro-temporal envelope as speech, and can model formant transitions between vowels and changes in pitch due to intonation. However, the resemblances between dynamic ripples and vocal sounds are limited to acoustic and phonetic properties. Dynamic ripples are not associated with speech on a phonological or linguistic level, involving the recognition of syllables or the comprehension of words.

Dynamic ripples can be characterized by four parameters (Fig. 4.2): the spectral modulation density Ω (in cycles per octave, c/o), the temporal modulation frequency ω (in cycles per second, c/s), the modulation amplitude A (%), and the drift direction d (downward or upward). Any sound pattern can be regarded as a superposition of such dynamic ripples, since they form a complete orthonormal basis of functions. In electrophysiological animal studies, dynamic ripples play a useful role in characterizing neuronal spectro-temporal response fields (Kowalski et al., 1996; Klein et al., 2000; Depireux et al., 2001; Miller et al., 2002). Selective neuronal responses to the speed and direction of spectral changes have been reported using a range of stimuli (deCharms et al., 1998; Tian et al., 1998; Liang et al., 2002), as well as corresponding topographic mappings on the brain's cortical surface (Mendelson et al., 1993; Shamma et al., 1993; Versnel et al., 1995). However, in humans electrophysiological experiments are infeasible due to the invasive nature of this technique. Psychoacoustical experiments have been performed to measure thresholds for the perceptual

Figure 4.2 The spectral power (I) of a dynamic ripple as a function of frequency (f) is based on pink noise, but contains a modulation that drifts in time along the logarithmic frequency axis in either direction d . The modulation density Ω equals the number of spectral peaks present in a frequency band of one octave. The modulation frequency ω is equal to the number of spectral peaks that passes per second. The modulation amplitude A is expressed relative to the noise level.

detection of spectro-temporal modulations. Near an optimum at $\Omega = 0.5$ c/o and $\omega = 4$ c/s, dynamic ripples with amplitudes A as low as 5 % can be discerned from noise without modulations (Chi et al., 1999).

We used functional MRI (fMRI), which is a volume imaging technique that is suitable for humans, with the additional advantage of a high spatial resolution. Although the poor time resolution prevents us from directly detecting any synchronization of neuronal responses with the temporal sound modulations, sustained changes in overall activity due to selective neuronal activation will be measurable (Seifritz et al., 2002). Given that the presence of tonotopy in the auditory cortex has previously been demonstrated using fMRI (Wessinger et al., 1997; Bilecen et al., 1998; Talavage et al., 2000; Le et al., 2001; Engelien et al., 2002; Formisano et al., 2003; Talavage et al., 2004b), it may also be able to reveal any topographic organization associated with dynamic ripples. In contrast to animal studies, the activation of the human auditory cortex in response to dynamic ripples has not yet been studied systematically. In this paper we will investigate the magnitude and spatial distribution of the cortical activation as a function of the previously mentioned dynamic ripple characteristics.

§4.2 Materials and Methods

4.2.1 Subjects

Nine healthy subjects with normal hearing were recruited (8 male, 1 female; aged 23-59 years, mean 37 years) and gave written informed consent to participate in this study. Of all subjects, 4 were right-handed, 3 left-handed, and 2 ambidextrous (Oldfield, 1971). Subjects received explanation and a short demonstration beforehand, but were not trained to the task.

4.2.2 Stimulus paradigm

A sound generation set-up was used to present the auditory stimuli. Each stimulus presentation consisted of a series of two 3.25-s dynamic ripples (Fig. 4.3). The first sound consisted of a reference stimulus with certain parameters (Ω, ω, d), and $A = 60$ %. This was immediately followed by a second stimulus with equal parameters (Ω, ω, d), but A equal to 30 % or 100 %. Values used for Ω were: $\frac{1}{4}$, 1 and 4 c/o; for ω : 2, 8 and 32 c/s; d was taken downward or upward. All possible combinations of Ω , ω , A and d were included in the experiment, giving a total of 36 dynamic ripple

stimulus series. In addition, 4 identical stimulus series were added that consisted of two consecutive noise stimuli without any modulations. Effectively this condition corresponds to $A = 0\%$. These noise stimulus series were used as the baseline condition in the subsequent image analysis. The experiment was performed in six blocks, each containing the total of 40 stimulus series in a different random order.

4.2.3 Task

Subjects were asked to compare both consecutively delivered auditory stimuli in each series, and to press one of two fibre-optic button-devices to indicate whether the perceived modulation amplitude of the second stimulus was larger or smaller than the first. The main purpose of this two-alternative forced-choice task was to let subjects focus actively on the auditory stimuli. The results were used only to verify that the subjects understood the task and that they were able to discern the spectro-temporal modulations. Given the fact that they judged 85 % of all stimulus series correctly, this was indeed the case.

4.2.4 Stimulus generation

Stimuli were generated using a real-time signal processing system (System II 16-bit DA converter, Tucker Davis Technologies) at a 32.8 kHz sampling

rate. The baseline stimulus was computed by inverse Fourier transformation of a 2^{15} -point complex array z_k according to

$$z_k = \frac{e^{i\varphi_k}}{\sqrt{|k|}}, \quad [4.1]$$

with quasi-random phases φ_k in the range $-\pi < \varphi_k \leq \pi$. The amplitude spectrum defined in Eq. 4.1 ensures that the logarithmic density of the spectral power $|z_k|^2$ is constant (i.e. pink noise). A spectro-temporal modulation was added by transforming the complex spectrum z_k to obtain a spectrum s_k according to

$$s_k = z_k + \frac{1}{2} A \cdot z_{k+\omega} \cdot e^{-2\pi i \cdot \frac{1}{2} \log(|k+\omega|) \cdot \Omega} + \frac{1}{2} A \cdot z_{k-\omega} \cdot e^{2\pi i \cdot \frac{1}{2} \log(|k-\omega|) \cdot \Omega}. \quad [4.2]$$

The drift direction d is controlled by choosing the sign of Ω and ω appropriately. The $\log$ -functions in Eq. 4.2 result in a constant number of modulations per octave. Because the obtained duration of the waveform after inverse Fourier transformation (1 s) is a multiple of the period of all dynamic ripples that were used, the waveforms could be repeated periodically. Signals were scaled to obtain equal intensity and loudness, and windowed with a 100-ms rise/fall time to obtain the stimuli.

The stimuli were delivered by means of headphones consisting of piezo-electric speakers built into a noise protection headset, with a lower cut-off frequency of 0.2 kHz. In addition to the headphones, subjects wore ear plugs, resulting together in 40 dB damping of the acoustic scanner noise of 105 dB peak sound pressure level (SPL). The auditory stimuli were presented binaurally at 80 dB SPL and were damped by 20 dB by the ear plugs alone. The spectral power at different frequencies in the range 0.2-16.3 kHz varied smoothly within 10 dB.

4.2.5 Imaging

Subjects were placed in a supine position in the bore of a 1.5-T MRI system (Philips Intera, Philips Medical Systems), which was equipped with a standard quadrupolar head receiver coil. A 3D T_1 -weighted fast-field echo scan with 50 contiguous sagittal slices was acquired for anatomical orientation (TR 9.85 ms; TE 3.5 ms; flip angle 15°; matrix 256×256; voxel size 0.94×0.94×3.0 mm). On these high-resolution images an imaging plane was positioned parallel to the Sylvian fissure of both hemispheres, containing the superior temporal gyrus of both temporal lobes. An additional 3D T_1 -weighted fast-field echo scan (TR 10.37 ms; TE 3.5 ms; flip angle 15°; matrix 256×256; voxel size 1.25×1.25×1.75 mm) was made of 20 slices parallel to this

imaging plane in order to serve as anatomical reference of the subsequent functional scans. The functional scans were made using an identically oriented scan volume as in the anatomical reference scan. These acquisitions consisted of a single-shot T_2^* -sensitive echo planar imaging (EPI) sequence over 10 slices (TR 9 s; TE 50 ms; flip angle 90° ; matrix 96×96 ; voxel size $2.5 \times 2.5 \times 3.5$ mm). For each subject the functional session consisted of six blocks of 40 acquisitions each; all dynamic ripple conditions were acquired 6 times in total, the baseline condition was acquired 24 times.

A considerable disadvantage of fMRI is the influence of acoustic noise produced by the scanner itself, evoking auditory responses that interfere with the responses to the stimuli being presented. This problem was reduced by using a sparse clustered volume scanning technique (Edmister et al., 1999; Hall et al., 1999; Yang et al., 2000; Le et al., 2001; Amaro et al., 2002). The functional acquisitions were performed at a rate of one per 9.0 s and took 1.25 s per image volume, leaving 7.75 s of scanner silence between acquisitions. An electronic trigger signal was used to start the presentation of the 6.5 s stimulus series precisely 1.0 s after the start of an acquisition (Fig. 4.3). The 9-s acquisition spacing allowed the hemodynamic response to the scanner noise to be largely faded at the time the next acquisition took place (Belin et al., 1999; Backes et al., 2002). Insofar as there is a remaining contribution of scanner noise it will be the same in all acquisitions, and thus it is assumed to cancel in the subsequent image subtraction analysis. The immediate influence of the scanner noise on the measurement itself is negligible because all slices were acquired within less than two seconds (Talavage et al., 1999).

4.2.6 Data analysis

In the data analysis we made use of the MatLab environment (The MathWorks Inc.) and routines from the SPM99 software package (Wellcome Department of Cognitive Neurology). The functional images were motion-corrected and spatially smoothed using a 5-mm Gaussian kernel. Images were thresholded to omit voxels outside the brain. Areas responding to the auditory stimulation were identified by analyzing the data of each of the voxels individually. We set up a general linear regression model containing 36 conditions representing all different dynamic ripple parameter combinations. To correct for drifts of the scanner signal, the baseline was modelled by a first order polynomial for each of the six acquisition blocks separately. Because the acquisitions were spaced at as much as 9 s, and because the order of the presented series of stimuli was randomized in each

block of acquisitions, we assumed that consecutive acquisitions were uncorrelated to simplify the analysis. We performed T-statistics for contrasts between subsets of dynamic ripple conditions and the baseline condition to obtain statistical parametric maps (SPMs). These subsets were made to include all conditions having the value of one dynamic ripple parameter in common. For instance, all dynamic ripples with $\Omega = 1/4$ c/o were contrasted together against the baseline condition to obtain one SPM. With the mentioned values for the parameters Ω , ω , A and d , this leads to a total of 10 contrasts. Furthermore, one SPM was constructed using the contrast of all dynamic ripple conditions against the baseline condition. Within the SPM formalism, the contrast vector was constructed such that the conditions with $A = 100\%$ were weighted twice as strong as conditions with $A = 30\%$.

Two different measures of the cerebral activation were employed. First, we defined the activation volume (AV) as the total volume of all voxels in the brain becoming significantly active, using a threshold of $p < 0.001$. The AV was determined as a function of each dynamic ripple parameter separately. For each of the 10 contrasts corresponding with a certain value of a dynamic ripple parameter, the significantly activated voxels in the brain were counted in the respective SPM, without constraints on location or cluster size. Second, the activation level (AL) was defined as the percentual signal change due to the hemodynamic response, averaged over all significantly activated voxels in the brain. In this calculation, the contrast of all dynamic ripple conditions against the baseline condition was used to determine one single set of voxels that significantly responded in general. For each of the 36 dynamic ripple conditions the percentual signal change caused by the hemodynamic response was averaged over this set of voxels to obtain AL. In this way we were able to determine AL as a function of all possible combinations of parameter values.

4.2.7 Separability

Both in electrophysiological (Kowalski et al., 1996) as well as in psychoacoustical (Chi et al., 1999) experiments it has been found that the response to dynamic ripples is separable with regard to Ω and ω . This means that there is no mutual dependence of these parameter characteristics, and the overall response is simply a product of the response characteristics to both of these modulation parameters:

$$R(\Omega, \omega) = R_{\Omega}(\Omega) \cdot R_{\omega}(\omega). \quad [4.3]$$

To investigate the separability of AL in our experiment, the method of singular value decomposition (SVD) was applied. In order to use this

method the data were shaped into a matrix $\mathbf{M}$, which was subsequently factored as $\mathbf{M} = \mathbf{U} \times \mathbf{\Lambda} \times \mathbf{V}^\dagger$. In this expression $\mathbf{\Lambda}$ is a diagonal matrix containing the singular values $\lambda_1 \dots \lambda_n$ in decreasing order, and $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices containing the corresponding vectors in both dimensional directions of the matrix. The degree of separability $\alpha_{\text{sep.}}$ is defined as

$$\alpha_{\text{sep.}} = \frac{\lambda_1^2}{\sum_{m=1}^n \lambda_m^2}. \quad [4.4]$$

This equals the proportion of $\mathbf{M}$'s total power which is accounted for by its best separable approximation (Depireux et al., 2001; Linden et al., 2003).

4.2.8 Topographic mappings

The presence of any topographic organization in the auditory cortex with respect to the parameters Ω , ω , A and d was studied by determining which parameter values were able to elicit the largest response in individual voxels. Because the brain's overall sensitivity to the presence of dynamic modulations strongly depends on the values of these parameters, a normalization step was carried out first. For each dynamic ripple parameter, the median response level of all significantly activated voxels was calculated as a function of the parameter value. Hereafter, the measured response levels of individual voxels were divided by this median brain response level to obtain a normalized response level. In each voxel, the parameter value resulting in the largest normalized response was determined for each separate parameter. In order to determine the presence of a topographic gradient, a linear function

$$f(x, y, z) = a \cdot x + b \cdot y + c \quad [4.5]$$

of the lateral-to-medial and posterior-to-anterior coordinates (x and y) was fitted to the normalized response patterns. As can be seen in Eq. 4.5, the inferior-to-superior coordinate (z) was not included in the fit, because the superior temporal surface hardly extends in that direction. For each activated voxel, f refers to the parameter value which gave the largest normalized response. For example, when investigating the topography of ω , discrete values equal to 1, 2 or 3 were fitted for f , indicating that the corresponding voxel gave a largest normalized response for $\omega = 2, 8$ or 32 c/s respectively. Fits were carried out for the largest cluster of significantly activated voxels in the left and in the right auditory cortices separately. The gradient vectors (a, b) of these fitted functions were used to indicate the

direction of the topographic mapping of the parameter (e.g. ω) onto the cortex, and to determine its significance ($p < 0.05$).

§4.3 Results

4.3.1 Activation volume and activation level

The cerebral activation caused by the spectro-temporal modulations in dynamic ripples depended on the parameters of the stimulus. Fig. 4.4 shows the results with respect to the activation volume (AV), averaged over all subjects. We found that AV is maximal for low modulation densities ($\Omega = 1/4$ c/o) and small modulation frequencies ($\omega = 2$ c/s), and decreases monotonically for both Ω and ω . With regard to the modulation amplitude we found that AV increased with A . In contrast, we did not find any significant changes in AV with the drift direction d . In Table 4.1 we computed the ratio of AV for the extreme values of the relevant parameters, with p -values indicating the level of significance of the differences.

Similar results were found with respect to the activation level (AL). Again, AL decreased with increasing values of Ω and ω , and increased with A . The relative differences were smaller as compared to AV, but still significant. Again there was no dependency on d .

To confirm the similarity between AV and AL, the correlation coefficient between the subject-averages of AV and AL for the various parameter

Table 4.1 A comparison of the activation for different values of various dynamic ripple parameters, as well as for the left and right hemispheres.

parameter	comparison	AV		AL	
		ratio	p -value	ratio	p -value
	Ω : $1/4$ c/o vs. 4 c/o	1.93	0.05	1.26	0.06
	ω : 2 c/s vs. 32 c/s	20.0	0.03	2.05	0.03
	A : 100% vs. 30%	3.56	0.03	1.65	0.03
	d : down vs. up	1.12	0.5	1.02	0.8
	hemisphere: right vs. left	1.07	0.9	1.02	0.9

The p -values were determined by the Wilcoxon matched-pairs signed-ranks test after which Hochberg-correction for multiple comparisons was applied (Hochberg, 1988), and indicate whether the values for the activation volume AV as well as the activation level AL for the given parameter values differ significantly.

values was computed. The result amounted to $R^2 = 0.97$, indicating a strong linear correlation between AV and AL ($p < 0.001$).

We applied the SVD-method to analyze the separability of the AL data with regard to Ω and ω . To obtain a matrix of values depending on Ω and ω only, the ALs were averaged over A and d . For the individual subjects, the degree of separability α_{sep} was in the range 0.974–0.996 with a mean of 0.989. When

Figure 4.4 The dependency of the activation volume AV (upper panel) and mean activation level AL (lower panel) on the dynamic ripple parameters Ω , ω , A and d . Results were averaged over all subjects, with error bars visualizing the standard error of the mean. The changes in activation as a function of the parameters are all significant, except for those in d .

the data were averaged over all subjects before the separability analysis, α_{sep} equalled 0.998, indicating almost perfect separability.

Additionally, we compared AV as well as AL in the left and right hemisphere separately. We found no significant quantitative hemispherical difference for any dynamic ripple stimulus or any set of stimuli (Table 4.1).

4.3.2 Activation pattern

To investigate the spatial distribution of cerebral activation, the signal change from baseline in response to the set of all dynamic ripple conditions was calculated for each activated voxel. A few representative images are given in Fig. 4.5a. The anatomical T_1 -weighted background of these images was extracted from the slice in the scan volume that contained the largest mean activation, and therefore shows the anatomical structures that correspond with the main regions of activation. The maximal activation level over all ten slices was colour-coded and projected perpendicularly onto this anatomical plane. In practice, the activation was largely confined to about three slices in the temporal lobe.

For each subject substantial regions of activation could be discerned

bilaterally, with an organization into multiple separate regions. These regions mainly extend in the posteromedial-to-anterolateral direction, approximately parallel to the Heschl's gyri. Generally at least two main regions could be identified in each hemisphere, stretching out along the borders of Heschl's gyrus. Usually more regions were observed, sometimes located anteriorly in the planum polare (in three subjects) but more often posteriorly in the planum temporale (in five subjects). However, there is considerable variability in the pattern of activation between subjects, and within subjects there is quite some variability between both hemispheres. Fig. 4.5b shows the signal change averaged over subjects in Talairach coordinate space. Two separate regions of activation extending along Heschl's gyrus are clearly visible in both hemispheres, while in the auditory cortex as a whole activation is present on a lower average level.

Similar regression analyses using SPM's were carried out to identify regions in the auditory cortex where the presence of dynamic modulations decreased the baseline activation caused by noise without modulations. No regions with significant signal reductions were found.

4.3.3 Topographic organization

In Fig. 4.6 we illustrated the parameter values to which each voxel responded best after normalization to correct for differences in sensitivity of the whole brain. In this figure the optimal parameter values for the voxels that showed significant activation were colour coded, as indicated in the legends, and projected perpendicularly onto the anatomical slice with the highest number of activated voxels. The upper left panel of the figure illustrates that the posterior (or posterolateral) part of the auditory cortex is more responsive to dynamic ripples with low spectral densities, whereas the anterior (or anteromedial) part responds relatively well to high values of Ω . This is visible in the form of a yellow-to-red colour gradient. This feature was noted by three independent observers in five out of nine subjects; in the other four subjects this pattern was visible retrospectively.

To quantify this observation, the gradient of the 2D linear fit to the spatial distribution of optimal Ω -values was calculated. The magnitude of this gradient was significantly different from zero for 6 subjects (for both hemispheres). Its average direction is indicated in the figure by the white arrows, and deviated from the posterior-to-anterior axis by approximately $30 \pm 50^\circ$ (left hemisphere) and $30 \pm 20^\circ$ (right hemisphere) towards the lateral-to-medial direction (mean $\pm$ standard deviation). Furthermore, it was observed in six out of nine subjects that the activated voxels in a lateral

position generally responded best to low modulation frequencies, while voxels located more medially responded relatively well to higher ω . Quantitatively, the magnitude of the gradient was significant in 5 and 6 subjects for the left and right hemispheres respectively. The average direction deviated approximately $65 \pm 60^\circ$ (left hemisphere) and $105 \pm 30^\circ$ (right hemisphere) from the anterior-to-posterior axis, as indicated in Fig. 4.6. The subjects that showed less clear parameter mappings for both Ω and

Figure 4.6 The spatial distribution of the optimal modulation density Ω , modulation frequency ω , modulation amplitude A , and drift direction d . The colours represent the parameter values that lead to an optimal response in a voxel, after normalization to correct for differences in sensitivity of the whole brain. When multiple voxels project to the same pixel, the corresponding colours are mixed. Ω and ω show a topographic mapping on the auditory cortex in posterolateral-to-anteromedial and lateral-to-medial direction respectively, which is visible as a red-to-yellow colour gradient. Results are shown for the significantly activated voxels in one subject that showed both mappings. The white arrows indicate the direction of the gradient averaged over multiple subjects. The images concerning A and d show a smoothed random distribution of optimal sensitivity, and no representative gradient is present. All images are shown in radiological coordinates.

ω were precisely those that showed the smallest amount of activation. Both topographic mappings stretch out over the entire region of activation in the auditory cortex, which amounts to approximately 4 cm in the posterolateral-to-anteromedial direction for the dynamic ripple density mapping, and to 3 cm for the lateral-to-medial mapping of dynamic ripple frequency.

With regard to A and d , the observers found no systematic evidence of a topographic organization subjectively. For these two parameters the normalized optimal response appeared to show a smoothed random distribution. Still, up to 3 subjects quantitatively showed a gradient that differed from zero. However, the directions of these gradients varied greatly between subjects (with spreads in excess of 90° in both directions), and no representative average could be calculated.

§4.4 Discussion

4.4.1 Activation dependencies

We considered two quantitative measures: the volume (AV) and the level (AL) of the activation. In principle it is possible for these two measures to show different dependencies on the dynamic ripple parameter values. After all, large volumes that show marginally significant activation might well be imagined, as well as confined regions with high activation levels. Nevertheless, our data show an almost perfect correspondence between both measures ($R^2 = 0.97$); AV shows the same general trends as AL. This suggests that at least in this experiment both are adequate measures of cerebral activation. The origin of the stimulus-dependence of AV is not obvious, but has also been shown in other experiments, e.g. for stimulus loudness (Jäncke et al., 1998; Lockwood et al., 1999; Mohr et al., 1999; Brechmann et al., 2002; Hart et al., 2002). It may partly be related to physiological phenomena like the recruitment of a larger neuronal population, or to an increased response contribution from draining veins, and also to the spatial smoothing of the data (Hall et al., 2001). In the following discussion we will refer to both measures simultaneously.

Our results show that the perception of dynamic ripples is accompanied with significant amounts of bilateral activation in an extensive part of the superior surface of the temporal lobes. The activation differs considerably as a function of the parameters Ω and ω , although the spatial locations of maximal activation approximately stay the same. Significantly less voxels reach the activation threshold as Ω and ω take on larger values. Our findings

indicate that the human auditory processing on a cortical level mainly focuses on spectro-temporal modulations that are quite sparse in spectral sense ($\Omega = 1/4$ c/o) and relatively slow in temporal sense ($\omega = 2$ c/s). This agrees with the psycho-acoustical data that show that the sensitivity in humans to dynamic ripples with these parameters is close to the optimum (Chi et al., 1999). These values also fall within the range of modulations that are present in speech, and correspond best with formant transitions. For even lower values of Ω and ω , we would expect the brain's sensitivity to the presence of spectro-temporal modulations to diminish again.

Because we observed a similar Ω -dependence of the cortical activation regardless of ω , and a similar ω -dependence regardless of Ω , we reasoned that the activation might be separable with regard to Ω and ω . To confirm this impression we applied the SVD-method to our AL data. The high degree of separability proved that the deviations from complete separability were negligible. Despite the limited amount of data, this is an important finding because it enables one to determine the brain response to all possible parameter combinations by only measuring the dependencies on Ω and ω separately, keeping the other parameter constant. However, from our data this conclusion is only justified for average responses in the brain as a whole, or at least large regions containing a multitude of neurons. We cannot extrapolate this to single neuron response characteristics.

4.4.2 Activation pattern

The large extent of activation shows that the processing of the spectro-temporal dynamics of sound is not confined to one specialized region in the auditory cortex. Instead numerous regions of activation were shown to exist. These are spread over the entire auditory cortex, mainly along Heschl's gyrus, but also extend posterolaterally onto the planum temporale and to a lesser extent anteriorly onto the planum polare. There is considerable variability in the precise location and number of involved regions between subjects and hemispheres. Nevertheless, the average activation persistently shows two regions of activation flanking Heschl's gyrus, in both hemispheres (Fig. 4.5b). The anterior region extends along Heschl's gyrus, somewhat on its anteromedial side towards the circular sulcus. The posterior region also extends along Heschl's gyrus, but is located on its posterolateral side, bordering on Heschl's sulcus. Both regions primarily occupy the lateral part of the auditory cortex. This pattern may correspond with functional areas in the cortex, and is quite consistent with the stripe-like patterns of activation described by others and denoted as T1b and T2

(Scheich et al., 1998; Di Salle et al., 2001; Brechmann et al., 2002). These regions might correspond with the most lateral portion of the auditory core region, Te1.0 and Te1.2, and the auditory belt region, Te2 (Morosan et al., 2001).

There are strong indications from literature that differences exist between both hemispheres concerning the processing of sound, especially at higher processing levels (Zatorre et al., 1992; Binder et al., 1997; Gage et al., 2002; Palomaki et al., 2002; Zatorre et al., 2002). Nevertheless, we did not find any statistically significant differences in the amount of activation between the left and right hemisphere. This was valid for the responses to the individual ripple stimuli, as well as for the mean response to the combined set of all stimuli. The precise patterns of activation differ between both hemispheres within subjects, but this variability exists between subjects as well. Because dynamic ripples contain similar spectro-temporal modulations as speech fragments, a left hemisphere preference might have been expected. However, the lack of phonological content apparently results in similar activity of both hemispheres. Current evidence indeed suggests that there are functional hierarchies, such that the early stages of processing take place in auditory core areas bilaterally. High cognitive functions like speech recognition and comprehension depend on these areas, but also involve more extensive regions predominantly in the left hemisphere (Demonet et al., 1992; Binder et al., 2000; Hickok et al., 2000; Zatorre et al., 2002; Liebenthal et al., 2005)

4.4.3 Topographic organization

The auditory cortex contains mappings of certain sound characteristics (Schreiner, 1998). Tonotopy is best known, and refers to the fact that the frequency of the tone that evokes the largest response depends on the cortical location of the neurons. In this experiment mappings of dynamic ripple parameters onto the auditory cortex were considered. It would be little meaningful to determine the parameter values for which the absolute response of a voxel is largest, because we already showed that the mean response in the brain is highly parameter dependent. The overall differences in sensitivity would overshadow the differences that might exist between voxels. In fact, almost all voxels did indeed respond best in an absolute sense to the same set of parameter values ($\Omega = 1/4$ c/o, $\omega = 2$ c/s, $A = 100\%$). Therefore we corrected for the brain's overall responsiveness by normalizing. Then, if no topographic organization of dynamic ripple parameters were present, approximately identical distributions of activation would be found for all dynamic ripple conditions. For each voxel it would

be a matter of chance which dynamic ripple evoked the largest normalized response. Using this normalized definition of topicity we found strong indications that a topographic organization indeed exists for Ω and ω , but no systematic topicity was clear for d or A .

With respect to the temporal modulation frequency ω , we found a topographic mapping approximately in the lateral-to-medial direction. Regarding the spectral modulation density Ω there is a mapping approximately in the posterolateral-to-anteromedial direction. This topicity was found in both hemispheres, although the results for the right hemisphere were slightly more consistent between subjects. With respect to temporal modulations, similar large-scale topographic organizations were found in humans (Giraud et al., 2000; Liegeois-Chauvel et al., 2004) and non-human vertebrates (Schreiner et al., 1988; Miller et al., 2002). The mapping regarding spectral modulations can be related to findings concerning a spatially dependent selectivity to the spectral broadness of stimuli (Rauschecker et al., 1995) and the sweep rates of FM stimuli (Tian et al., 2004). Both topographic organizations span the entire auditory cortex. This suggests a hierarchical organization of the human auditory cortex as a whole with regard to modulation density and frequency. The anteromedial region responds relatively well to sounds with small-scale spectral and temporal features (i.e. high Ω and ω), representing acoustic details like specific voice characteristics or the presence of echoes. This is in line with the primary auditory cortex being located in this region (Galaburda et al., 1980; Morosan et al., 2001; Rademacher et al., 2001). The posterior and lateral regions on the other hand respond best to the slowly changing large-scale features in the spectrum (i.e. small Ω and ω), both in absolute and in relative sense, by performing some spectro-temporal integration (Thivard et al., 2000). This indicates a specialization towards the processing of formants and phonemes. This agrees with the fact that speech-related areas lie in these regions (Binder et al., 1994b; Hickok et al., 2000; Wise et al., 2001).

In electrophysiological experiments, neuronal preferences were shown for the direction (d) of spectral drifts (Shamma et al., 1993; deCharms et al., 1998; Tian et al., 1998). Several reasons can explain why our results did not demonstrate any differences in the average brain response or any topographic mapping with respect to d . First of all, the number of responding neurons may be small. Note that only a small fraction of the neurons might respond to a stimulus with a given set of parameter values (Ω , ω , d and A) when a selectivity exists for certain combinations of these parameters. Functional MRI is a technique with a limited signal-to-noise ratio, and will not be sensitive enough to distinguish a sufficiently small

fraction from the background activation of the complete cell population. Second, a population bias toward either drift direction may be absent in humans. If neurons with preferences for upward drifts have the same spatial density as neurons that preferentially respond to downward moving ripples, then the measured macroscopic response patterns may be similar. Literature suggests that neurons with directional preferences are clustered at small scales (<0.2 mm) within cortical fields (Nelken et al., 2000), while at larger scales directional preferences do not differ between multiple fields (Kowalski et al., 1995; Tian et al., 1998). This scale at which clustering is found is too small to resolve using fMRI at 1.5 T.

§4.5 Conclusion

A shift was found in the brain's sensitivity from small-scale to large-scale spectro-temporal features as signals are mediated from primary to secondary auditory cortex. Such a notion is of fundamental importance in the understanding of the hierarchical processing as it takes place in the auditory cortex.

TONOTOPY AND LATERALIZATION IN THE CORTEX

Chapter

5

Abstract

In this study, functional MRI was performed to investigate the functional organization of the auditory cortex. A modified sparse acquisition scheme was used involving pairs of contiguous scans spaced by periods of scanner inactivity during which auditory stimuli were presented. Signal detection was optimized by varying the weighting of the acquired images. Thus, an improvement in statistical significance was achieved for individual voxels. Activation could be detected more extensively, and with smaller average activation levels. The method was used to map the characteristics of the auditory cortex with respect to stimulus lateralization and frequency. In the regions of Heschl's gyrus and Heschl's sulcus that correspond with the primary auditory cortex in the temporal lobe, activation was strongest contralateral to the ear of stimulation. In the surrounding active regions, including the planum polare and the planum temporale, systematic preferences with respect to stimulus lateralization were absent. Regarding stimulus frequency, low- to high-frequency tonotopic gradients were evident along the Heschl's gyrus and Heschl's sulcus in anterolateral to posteromedial direction, especially in the right hemisphere. In the surrounding activated areas frequency preferences were erratic. Mappings with respect to lateralization and frequency showed hardly any interaction in the sense that lateralization maps were independent of stimulus frequency and tonotopic maps were independent of stimulus lateralization. Our results suggest that the primary auditory cortex is topographically organized with respect to the physical stimulus properties regarding lateralization and frequency, but these organizational principles are no longer obvious in the secondary areas.

§5.1 Introduction

The cerebral hemispheres receive sensory input from both ears (Moore, 1991; Woldorff et al., 1999; Suzuki et al., 2002). Nevertheless, the human auditory cortices are known to respond most strongly to stimulation of the contralateral ear (Scheffler et al., 1998; Woldorff et al., 1999; Bilecen et al., 2000; Jäncke et al., 2002a). Such cortical response patterns directly reflect the contralateral preferences that have also been found in the inferior colliculus in the midbrain and the medial geniculate nuclei in the thalamus (Melcher et al., 2000). This suggests that each hemisphere focuses on sound sources in the contralateral hemifield of the auditory environment (Lipschutz et al., 2002), similar to and compatible with the functioning of the visual and somatosensory systems.

Another eye-catching organizational feature of the auditory system is its tonotopic representation, a gradual spatially dependent selectivity for acoustic stimulus frequency that finds its origin in the cochlea and that persists in the brainstem nuclei as well as in subdivisions of the primary and possibly even secondary auditory cortices (Pickles, 1988). Detailed spatial dependencies of the frequency characteristics of individual neurons in the auditory cortex have been recorded in studies in primates and other mammals (Merzenich et al., 1973; Reale et al., 1980; Kelly et al., 1986; Stiebler et al., 1997; Bonham et al., 2004; Lee et al., 2004), but in humans electrophysiological recordings of tonotopy are exceptional (Howard et al., 1996). Most studies in humans that use non-invasive imaging techniques agree on the existence of an overall gradient in the auditory cortex from low- to high-frequency selectivity towards the medial and sometimes posterior direction (Romani et al., 1982; Lauter et al., 1985; Wessinger et al., 1997; Bilecen et al., 1998; Muhlneckel et al., 1998; Schmid et al., 1998; Cansino et al., 2003; Fujioka et al., 2003; Yetkin et al., 2004b). Yet, some evidence suggests that these gradients lie in the exactly opposite direction (Yang et al., 2000; Engelien et al., 2002), and numerous further studies report a diverse range of findings involving more complex arrangements (Pantev et al., 1996; Langner et al., 1997; Lockwood et al., 1999; Talavage et al., 2000; Liegeois-Chauvel et al., 2001; Schönwiesner et al., 2002; Formisano et al., 2003; Scarff et al., 2004; Talavage et al., 2004b). Among other factors, the presence of large inter-subject variabilities (Ehret, 1997; Weisz et al., 2004) and, in the case of functional magnetic resonance imaging (fMRI), the interference of acoustic scanner noise (Le et al., 2001; Scarff et al., 2004) have hindered a more detailed understanding of the tonotopic organization in humans. In conclusion, although tonotopic organization is a key property of the

auditory system, its expression in the human auditory cortex has not yet been established unambiguously.

Previous studies that have investigated response lateralization or tonotopic organization often only report summary statistics for the auditory cortex as a whole, for instance by calculating lateralization indices for left and right activation volumes to monaural stimulation, or by detecting shifts in the effective location of activation as a function of stimulus frequency. We reinvestigated response lateralization and tonotopic organization employing an active listening task and an adapted fMRI-paradigm. In our analysis, cortico-spatial preferences for stimulus lateralization and frequency were determined and visualized at the level of individual voxels, for instance allowing a distinction to be made between primary and secondary processing centers. Furthermore, by investigating both response lateralization and tonotopic organization in a single study, potential interactions between these two characteristics could be investigated.

§5.2 *Materials and Methods*

5.2.1 Subjects

Ten healthy subjects (7 male, 3 female; 7 right-handed, 2 left-handed, 1 ambidextrous (Oldfield, 1971); age 23–43 years, mean 31 years) were recruited on the basis of written informed consent, in approved accordance with the requirements of the medical ethical committee of the institute. All subjects were screened by an audiologist using standard pure tone audiometry (Katz, 2001) and had thresholds better than 20 dB hearing level in the range of 250–8000 Hz.

5.2.2 Auditory task

During the fMRI sessions subjects performed an active listening task, because this was presumed to enhance the recorded response (Hall et al., 2000). A tone-in-noise detection task with band-limited stimuli was chosen. Noise bursts had a central frequency equal to either 125, 500, 2000 or 8000 Hz, a bandwidth of $\frac{1}{4}$ octave, a duration of 1.5 s, and a rise/fall time of 10 ms; pure tones had frequencies identical to the central frequencies of the noise, a duration of 0.5 s, and a rise/fall time of 10 ms. Noise bursts were spaced by 0.5 s from each other and preceded the following MR-acquisition by 0.75 s. The tone presentations coincided with either the first or the second

noise burst in a pseudo-random fashion. The onset of the tone occurred 0.5 s after the onset of the noise stimulus (see Fig. 5.1). Subjects were instructed to indicate whether the additional tone coincided with the first or the second noise stimulus by pressing one of two fiber-optic button devices that were held in each hand. The results were used only to monitor that the subjects understood the task and stayed alert during the fMRI session.

Tones and noise were presented at an intensity of 60 dB sensation level by means of modified electrostatic head phones (Sennheiser HE 60, Sennheiser) built into a noise protection headset that were connected to a PC setup equipped with a sound card with 16-bit DAC output capability (Sound Blaster Live! Platinum 5.1, Creative). Stimulus waveform data were

Figure 5.1 Schematic representation of the experimental paradigm. Functional echo planar imaging (EPI) acquisitions were performed in pairs (labeled 'A' and 'B') at time intervals (t) of 16 s using a sparse acquisition paradigm. Preceding each acquisition pair, two 1.5-s band-passed noise stimuli were delivered with a center frequency of 125, 500, 2000 or 8000 Hz. At 0.5 s after the onset of one of the noise bursts, an additional 0.5-s pure tone stimulus was presented. The tone frequency was equal to the center frequency of the noise. Subjects were instructed to indicate by means of a button press whether the tone coincided with the first or the second noise burst. The stimulus and acquisition timing was chosen such that the expected hemodynamic response to the noise stimuli (solid curves) was approximately maximal at the time of the EPI acquisitions, whereas the response to the acoustic noise of the previous MRI acquisitions (dashed curve) had faded to a negligible level.

generated in advance using digital Fourier transformation techniques, and were stored in the form of standard 16-bit uncompressed sound files with a sampling frequency of 44.1 kHz.

5.2.3 MR acquisitions

Subjects were placed in supine position in the bore of a clinical 1.5-T MRI system (Philips Intera, Philips Medical Systems), which was equipped with a standard quadrupolar head receiver coil. A sagittal 3D T_1 -weighted fast-field echo scan was acquired for anatomical orientation (repetition time TR 10.20 ms; echo time TE 3.5 ms; flip angle 15°; matrix 256×256×75; voxel dimensions 1.0×1.0×2.0 mm). On these high-resolution images a central imaging plane was positioned approximately parallel to the Sylvian fissure of both hemispheres, containing the superior temporal gyri of both temporal lobes. An additional 3D T_1 -weighted fast-field echo scan (TR 10.34 ms; TE 3.5 ms; flip angle 15°; matrix 256×256×12; voxel dimensions 1.0×1.0×2.0 mm) was acquired parallel to this imaging plane and served as anatomical reference. The functional scans were acquired covering the same brain volume, and consisted of a dynamic series of 2.5-s single-shot T_2^* -sensitive echo planar imaging (EPI) sequences with twelve 2.0-mm thick adjacent slices (TE 50 ms; flip angle 90°; matrix 192×192; field of view 192×192 mm).

A sparse clustered volume scanning paradigm was employed to reduce the influence of acoustic scanner noise. The standard paradigm (Edmister et al., 1999; Hall et al., 1999; Talavage et al., 1999; Yang et al., 2000; Le et al., 2001; Amaro et al., 2002; Backes et al., 2002) was modified to include two contiguous acquisitions, henceforth referred to as 'A' and 'B'. These pairs of acquisitions were performed every 16 s such that the TR alternated between 13.5 s and 2.5 s for acquisitions A and B respectively (see Fig. 5.1). Auditory stimuli were presented during the 11-s gap of scanner silence between acquisition pairs.

For each subject the fMRI session consisted of six runs of 20 acquisition pairs each, excluding initial start-up scans. Between successive scan pairs, the four different stimulus frequencies and a silent baseline condition were cycled. This allowed each of the stimulus frequencies as well as the silent condition to be presented four times per run. In the first and fourth functional run, all stimuli were presented binaurally; in the second and fifth run, stimuli were presented to the left ear only; and in the third and sixth run, stimuli were presented to the right ear only.

5.2.4 Image analysis

In the data analysis we made use of the MatLab programming environment (The MathWorks Inc.) and routines from the SPM2 software package (Wellcome Department of Cognitive Neurology). The functional image volumes were sorted into two sets, respectively containing the images from acquisitions A and B. Both sets were separately corrected for motion effects. Image set B was coregistered with respect to image set A using rigid body transformations on the basis of the mean images of each set, and subsequently the images of all subjects were spatially normalized with respect to each other using 12-parameter affine transformations. To improve signal to noise ratio (SNR) characteristics, spatial smoothing was performed using an isotropic 3-mm Gaussian kernel and images were interpolated to voxel dimensions of $2.0 \times 2.0 \times 2.0$ mm.

General linear models were set up for each voxel (Friston et al., 1999), both for individual subjects and for the group of all 10 subjects collectively. All models had a similar design and involved the 13 stimulus conditions (i.e., four stimulus frequencies, each presented binaurally or monaurally left or right; plus a single silent baseline condition). To correct for drifts of the scanner signal, the baseline signal was modeled by a first order polynomial for each of the functional runs separately. The models were evaluated using multiple linear regression.

Initially, one set of voxels that showed significant responses ($p < 0.05$, corrected for multiple comparisons) was determined on the basis of an *F*-test involving the contrasts of all individual stimulus conditions with the silent baseline. SNR characteristics were averaged over all voxels in this single set for various weightings between the A and B sets of images. The outcomes were compared to optimize the signal weighting, where optimality should be interpreted in the sense that the achieved SNR values were largest such that the assignment of voxels to the set of active voxels could be done with the largest degree of statistical certainty. As a result, image set A was weighted twice as strong as set B for all further analyses (for detailed outcomes, see the results section).

5.2.5 Activation characterization

Estimates were derived for the activation in the form of statistical parametric maps (SPMs) by contrasting conditions (or combinations thereof) with the silent baseline condition using *T*-test statistics (Friston et al., 1995). The responses of all voxels to specific stimulus characteristics were determined

by grouping the conditions that shared a particular characteristic, and contrasting these conditions against the silent baseline condition. For instance, the activation to left ear stimulation was determined by contrasting the four conditions corresponding with left ear stimulation (comprising the four different stimulus frequencies) against the baseline. Similarly, the activation to stimuli with a particular frequency was determined by grouping over the three corresponding conditions with different stimulus lateralizations. Activation levels (ALs) for each voxel were defined as the average signal change percentage in a particular contrast relative to the baseline signal intensity. A mean AL was determined by averaging the voxel ALs over a subject-dependent region of interest (ROI) consisting of the 625 voxels (i.e., 5.00 cm³ of brain tissue) in the superior surfaces of the left and right temporal lobes with the most significant activation in the contrast involving all twelve stimulus conditions.

To demonstrate potential topographic mappings in the cortex, the ALs of individual voxels were corrected for overall differences in sensitivity of the whole brain to different stimulus conditions. To this end, the responses to the three different stimulus lateralizations and the four different stimulus frequencies were normalized by dividing the ALs of individual voxels by the mean AL of all voxels in the previously determined ROI. Each active voxel was assigned an index that corresponded with the stimulus lateralization or frequency that led to the largest normalized AL. For the lateralization mapping, this index was in the range 1 to 3 (1 = monaural left; 2 = bilateral; 3 = monaural right) whereas for the frequency mapping the range was 1 to 4 (1 = 125 Hz; 2 = 500 Hz; 3 = 2 kHz; 4 = 8 kHz). The average index of all active voxels that projected onto the same location in a transverse anatomical slice was displayed using a red-to-yellow color code for the entire range of index values.

§5.3 Results

5.3.1 Sparse acquisition weighting

To optimize signal detection, the relative weighting of the two image sets (A:B) was varied between a weighting that effectively uses only set A (1:0) to a weighting that uses only set B (0:1). In this analysis, all conditions were weighted equally with respect to stimulus lateralization and stimulus frequency, and data were averaged over the set of voxels in the upper

surfaces of the temporal lobes that showed significant responses according to the general *F*-test.

The results with respect to the average hemodynamic response signal, expressed as a percentage relative to the baseline signal level, are shown in Fig. 5.2a. For set A the signal was stronger than for set B, with values of 0.76 % and 0.45 % respectively. For intermediate weightings that combine both sets, the signal magnitude shows a linear transition. In Fig. 5.2b the corresponding mean standard error of the voxel signals has been plotted as a function of the weighting employed. For set A standard errors were lower than for set B, equaling 0.11 % and 0.14 % respectively. However, by combining both sets standard errors as low as 0.09 % could be obtained (for a 3:2 weighting). Fig. 5.2c shows the resulting average *T*-values, calculated per voxel as the ratio of the signal to the standard error. Following from the attained *T*-values, the activation according to image set A ($T = 6.9$) was more significant than that according to set B ($T = 3.3$). Optimal signal detection was achieved for weightings near 2:1 ($T = 7.7$).

The activation volume (AV) and activation level (AL) were determined using the various set weightings, and the outcomes are plotted in Figs. 5.2d-e. The AV that followed from set A only (9.4 cm³) was an order of magnitude larger than that from set B (0.94 cm³). A maximum in the number of detected voxels was achieved for intermediate weightings, amounting to 11.0 cm³ near a 2:1 weighting. For the calculation of the mean AL, included in the average were all voxels that were significantly active according to the particular weighting under consideration; this is in contrast with the calculations leading to Fig. 5.2a, for which one invariant set of voxels was considered. The AL showed less variation, and equaled 0.72 % for set A and 0.74 % for set B. For a 2:3 weighting, the AL decreased to 0.55 %.

On the basis of the previous results, a weighting equal to 2:1 was chosen for all following analyses to ensure optimal signal detection (i.e., highest significance according to *T*-values, see Fig. 5.2c). In Fig. 5.2f, the group activation pattern is shown according to set A only, set B only, and both sets using the 2:1 weighting. Activation maxima were projected on a transverse plane and overlaid on a corresponding anatomical image. Extensive activation was mainly clustered in the two temporal lobes, especially in the regions from Heschl's gyrus to the temporal plane. In the right temporal lobe, the group activation was organized into elongated substructures with stripe-like activation clusters in the (postero)medial to (antero)lateral direction, somewhat parallel but mostly posterior to the Heschl's gyrus. In the left hemisphere, activation was more diffuse and most pronounced in

the most lateral regions. Some additional motor activation in the thalamic region was observed.

5.3.2 Activation characteristics

The activation in the auditory cortex in response to various stimuli was determined using *T*-test statistics. Fig. 5.3 shows the activation patterns that were found for individual subjects as well as for the group as a whole. The thresholds for voxels to be considered significantly active were adjusted for each subject such that precisely the 625 most active voxels were detected, resulting in individual thresholds for significance in a range that corresponded with *p*-values of $8 \cdot 10^{-5}$ to 0.016 with a median of $6 \cdot 10^{-4}$ (uncorrected for multiple comparisons). For the group analysis, the 625-voxel threshold amounted to $p = 2 \cdot 10^{-9}$.

Activation was found in both temporal lobes for all subjects (Fig. 5.3a), and was distributed over Heschl's gyrus (HG), Heschl's sulcus (HS), the planum temporale (PT) and in some subjects the planum polare (PP). The activation was clustered in stripe-like patterns, typically in a direction parallel to HG and HS (see e.g. subjects #3, #6 and #8). The stripe-like patterns that were found in the individual subjects were also discernable in the group data (Fig. 5.3b). Especially in the right hemisphere, a dominant elongated cluster of activation was found along HS, bordering on the posterolateral side of HG. In the left hemisphere, activation appeared more diffuse and was concentrated towards the lateral side of the brain.

Figure 5.2 (next page) Signal-to-noise and activation characteristics as a function of the weighting of both repeated acquisitions ('A' and 'B' in Fig. 5.1). (a) The signal magnitude of a constant reference volume of voxels was highest in the set of first acquisitions ('A') and showed a linear decrease as a function of the weighting. (b) The mean standard error in the reference voxels was smaller in set 'A' than in set 'B', but minimal errors were obtained using a weighted average. (c) The achieved *T*-values, obtained as the ratio of the signal to the standard error, were highest for an approximate 2:1 weighting of the two image sets. (d) The detectable volume of significantly active voxels showed a similar qualitative dependence upon the weighting as the *T*-values, but variations were more extreme. (e) The average activation level of the activated volume according to the weighting used was generally smaller for a weighted combination of both image sets than for both sets separately, indicating that lower activation levels are detectable using a weighted combination. (f) Activation patterns according to image set 'A', image set 'B', and the 2:1 weighting, projected on a transverse slice. Activation was most extensive using the optimized weighting, but the location of the detected activation maxima was independent of the weighting.

Fig. 5.3c shows group activation patterns split up with respect to the stimulus conditions. The 625 most active voxels according to the respective contrasts have been shown. Regarding the stimulus lateralization, activation following monaural stimulation was dominantly present in the hemisphere contralateral to the stimulated ear. For binaural stimulation the activation appeared more symmetrical, although it was more parcellated in the right hemisphere and diffuse in the left. Regarding the stimulus frequency, the

activation in the elongated cluster in the right hemisphere was located more (postero)medially for the lower frequencies than for the highest frequency. The general activation patterns seemed similar for all conditions in the sense that activation clusters were found at similar locations.

The mean AL across the individual subjects is summarized in Fig. 5.4 for the left and the right hemispheres separately. The average AL in response to all stimulus conditions was in the order of 1.5 %. Differences between the left and the right hemisphere were not significant. Weak but systematic trends were observed as a function of the stimulus lateralization: in both hemispheres responses were larger for contralateral stimulation than for ipsilateral stimulation, while for binaural stimulation, responses were intermediate. Results for the AL as a function of stimulus frequency showed that responses were strongest for 500-Hz stimuli and weakest for 8-kHz stimuli.

5.3.3 Lateralization maps

For all conditions, the voxel ALs were divided by the ROI mean as depicted in Fig. 5.4 in order to correct for differences in the overall brain sensitivity. Topographic activation maps were constructed by determining the stimulus condition that led to a maximal scaled voxel response. Fig. 5.5a shows the projected maps with regard to the preferred stimulus lateralization in individual subjects. For most of the subjects, responses were strongest in the hemisphere contralateral to the stimulated ear (see e.g. subjects #5, #6 and #8). Joint results from the group model (Fig. 5.5b) clearly confirmed the

Figure 5.3 (previous page) The activation in the 625 most significantly active voxels in the upper surface of the temporal lobes is projected on an anatomical slice by means of a 'hot' color code that represents the *T*-value of the activation using a subject-dependent threshold. (a) Activation in the ten individual subjects. All stimulated conditions were contrasted with silence. The activation was clustered in the temporal lobes, forming stripe-like patterns in the (antero)lateral to (postero)medial direction approximately parallel to Heschl's gyrus. (b) The data of all subjects combined into a single group model. In the right hemisphere the activation was strongly clustered along Heschl's gyrus (HG) and Heschl's sulcus (HS), whereas in the left hemisphere the activation was more diffuse. Some additional activation was found in the planum polare (PP) and the planum temporale (PT). (c,d) The activation patterns, segregated with regard to stimulus lateralization and frequency, respectively. Patterns were similar in the sense that the activation was located at approximately the same locations. However, inter-hemispheric differences are visible, as activation to monaural stimulation was found most strongly in the contralateral hemisphere. In the right hemisphere activation to low-frequency stimuli was located more laterally than that to high-frequency stimuli.

Figure 5.4 The mean activation levels AL of all significantly active voxels were determined for various stimulus conditions. Averaged over all stimulus conditions (*All*), ALs were in the order of 1.5 % and did not differ significantly between both hemispheres. For left monaural (*L*), binaural (*B*) and right monaural (*R*) stimulus lateralizations, systematic trends were found. For both hemispheres the activation to contralateral stimulation was stronger than that to ipsilateral stimulation, while activation to binaural stimulation was intermediate. Regarding the stimulus frequency, the strongest responses were evoked by 500-Hz stimuli and responses were least strong following 8-kHz stimulation. The mean activation levels of all active voxels were used to normalize the activation data prior to determining which stimulus conditions led to the strongest activation for individual voxels. The error bars in the figure indicate the standard error of the mean across subjects.

contralateral preference for a majority of the voxels: 53 % responded most strongly to contralateral stimulus presentation, while only 16 % responded most strongly to ipsilateral stimuli. Especially in the regions of the auditory cortex corresponding with HG and HS a contralateral preference was evident. In contrast, preferences with regard to stimulus lateralization were more erratic and randomly organized in the anteriorly located PP and the posteriorly located PT.

Mappings of the preferred stimulus lateralization in the group model for the four stimulus frequencies individually (Fig. 5.5c) were similar to each other and to those in Fig. 5.5b and only minor differences were observable.

5.3.4 Tonotopic maps

A similar analysis was performed with regard to preferences for stimulus frequency. For some individual subjects (Fig. 5.6a), tonotopic gradients were found stretching out along the elongated activation clusters that border HG, from low to high frequencies in the anterolateral to posteromedial direction (see e.g. subjects #3, #6 and #8). However, there was considerable variability between the subjects. For the group as a whole (Fig. 5.6b), low-to-high tonotopic gradients could be observed in an anterolateral-to-posteromedial direction along HG and HS. Although such a gradient was present in both hemispheres, it was less prominent in the left hemisphere than in the right. In addition, some voxels that were situated on the extreme lateral end of the active regions in HG and HS in both hemispheres showed preferential responses to high-frequency stimuli. In the PP and PT local variations in response characteristics were visible, but on a larger scale systematic frequency gradients could not be identified in the individual subjects or in the subjects as a group.

Fig. 5.6c shows tonotopic mappings for each of the three different stimulus lateralization conditions. These mappings are highly similar, and tonotopic organization did not depend on the lateralization of the stimulus.

§5.4 Discussion

In this study cortical activation was detected in response to monaurally and binaurally presented stimuli of various frequencies using an adapted sparse acquisition paradigm. The response characteristics of individual voxels were calculated to determine their relative preferences regarding stimulus lateralization and frequency, and the results were visualized to show the presence of topographic mappings.

5.4.1 Activation patterns

Our results were all obtained using an active listening paradigm, which involved two noise bursts and a tone. Because both the noise bursts and the tones were presented preceding each MRI acquisition, the activation corresponds to the combined response to these stimuli. Since the tones were spectrally centered in the $\frac{1}{4}$ -octave noise bursts, all derived activation patterns (Figs. 5.3, 5.5, and 5.6) represent the brain response to narrowband stimuli.

The activation showed multiple stripe-like regions of activation parallel to the Heschl's gyrus (HG) and Heschl's sulcus (HS) of both hemispheres. Because the activation from all slices was projected on a single image, the

Figure 5.5 Topographic cortical mappings of stimulus lateralization. (a) Individual subjects. Considerable variability in response preference was found across individual subjects, although in most subjects there was a general tendency for the auditory cortices to be more responsive to stimulation of the contralateral ear. (b) Group model. Systematic contralateral response behavior was found in the Heschl's gyrus (HG) and Heschl's sulcus (HS). In the Planum polare (PP) and planum temporale (PT), response lateralization was more erratic and not systematically contralateral to the ear of stimulation. (c) Preferential lateralization for each of the four stimulus frequencies. The response characteristics were very similar to those in panel b. All stimulus lateralization maps were determined after a normalization procedure (see text).

clusters cannot be an artificial result of displaying only a cross-section through a 3D activation volume.

Therefore, multiple active regions with an elongated shape running in the

Figure 5.6 Topographic cortical mappings of stimulus frequency. (a) Individual subjects. Frequency mappings appeared highly variable across subjects. However, in several subjects topographic gradients were found stretching out along the length of the stripe-like activated regions. (b) Group model. Preferential responses to high-frequency stimuli were found in the medial parts of the activated regions in the Heschl's gyrus (HG) and Heschl's sulcus (HS), while responses shifted towards the lateral side for lower frequencies. A distinct tonotopic gradient is visible especially in the right hemisphere. There are no clear tonotopic gradients in the planum polare (PP) and the planum temporale (PT). (c) Tonotopic mappings for three stimulus lateralization conditions. These maps are similar to that in panel b. Apparently, tonotopic organization does not depend on stimulus lateralization. All stimulus frequency maps were determined after a normalization procedure (see text).

(antero)lateral to (postero)medial direction were shown in both hemispheres. Such regions have been described in other studies as well (Scheich et al., 1998; Di Salle et al., 2001; Brechmann et al., 2002) and might reflect the location of different functional areas in the cortex. Alternatively, they may be related to the morphology of the local vasculature, since the hemodynamic response in fMRI has a vascular origin and the courses of the veins on the cortical surface follow the cerebral sulci to some extent.

For the group model, response patterns were more diffuse in the left hemisphere than in the right. This might be caused by functional or anatomical inter-hemispheric asymmetries in individuals. However, because the stripe-like patterns were observed in comparable fashion for both hemispheres in individual subjects, it seems more likely to be the result of larger variations in the left hemisphere regarding the spatial location of functional and anatomical features across subjects (Leonard et al., 1998; Rademacher et al., 2001). Nevertheless, multiple activated sub-regions could be identified in the group analysis, especially in the right hemisphere, indicating that the stripe-like features are similar and reproducible across subjects.

The spatial response patterns to all types of stimuli were similar in the sense that the location and shape of the activated regions were similar. However, a preferential hemispheric response to stimulation of the contralateral ear was evident. In addition, in the right hemisphere some evidence for the presence of a tonotopic gradient was observed as activation to high-frequency stimuli was found more posteromedially than that to low-frequency stimuli. Because differences between activation images are not easy to interpret, additional color-coded mappings were constructed on the basis of which *stimulus conditions led to the largest activation levels for individual voxels* after correction for the average response characteristics of the active brain regions as a whole.

5.4.2 Response lateralization

For unilateral stimulation the activation was found to be strongest in the contralateral hemisphere, in spite that most processing centers of the central auditory system are known to receive information from both ears. It has been shown in literature that at least from the inferior colliculi onwards activation is predominantly - although not exclusively - contralateral in nature (Fitzpatrick et al., 1980; Pantev et al., 1986; Scheffler et al., 1998). Our findings for the auditory cortex are in general agreement with these observations.

Relative preferences for left or right stimulus presentation were also analyzed and visualized at the level of individual voxels. The observation that activation occurs most strongly in response to stimulation of the contralateral ear was especially valid for the anterior regions of the auditory cortex, comprising the medial two-thirds of the HG where the primary processing centers of the auditory core regions are presumed to be located. Because the primary auditory cortex cannot be demarcated with exact precision on the basis of anatomical and functional data (Leonard et al., 1998; Morosan et al., 2001; Rademacher et al., 2001), some of the directly surrounding areas (e.g. those located in the HS) may show a contralateral preference as well.

In contrast, contralateral preferences were not systematically observed for activation situated in the planum temporale (PT) and the planum polare (PP). These parts of the upper temporal lobe contain areas involved in higher auditory processing. Here, a more random distribution of voxels responding preferentially to the left and/or right ear was found in the group analysis. Although this might reflect a fine parcellation into various specialized sub-areas, it can also be the result of an absence of a clear preference regarding stimulus lateralization.

5.4.3 Tonotopic organization

A low-to-high frequency gradient in the (antero)lateral to (postero)medial direction in the primary auditory cortex was prominent. The frequency dependence of brain activation patterns that was found in this study is compatible with previous investigations by others (Romani et al., 1982; Howard et al., 1996; Wessinger et al., 1997; Bilecen et al., 1998; Schmid et al., 1998; Yetkin et al., 2004b). Tonotopic gradients in the left primary auditory cortex were less pronounced than in the right, but still noticeable.

Hemispheric differences have been found before in relation to frequency selectivity and tonotopic organization (Wessinger et al., 1997; Rosburg et al., 1998; Yetkin et al., 2004b), possibly related to an asymmetry in the anatomy and function of both hemispheres. It has been suggested that the right hemisphere is most responsive to acoustic sound features like pitch, such as found e.g. in music, whereas the left hemisphere is more or less specialized in the processing of temporal dynamics, like the phonological aspects of speech-like stimuli (Binder et al., 1997; Shapleske et al., 1999; Hickok et al., 2000; Gage et al., 2002; Palomaki et al., 2002; Josse et al., 2003; Specht et al., 2003; Tervaniemi et al., 2003). A compatible interpretation would be that the right half of the brain specializes on spectral resolution and the left half on

temporal resolution (Zatorre et al., 2002). Our results are consistent with an emphasis on spectral analysis in the right hemisphere as opposed to the left hemisphere.

On a more detailed level it is much more difficult to discern the relevant substructures in the brain, to objectify their existence on the basis of functional criteria, and to establish their organizational principles. Some reports show multiple sub-areas with mirror-symmetric tonotopic representations in humans and other mammals (Ehret, 1997; Kaas et al., 1999; Formisano et al., 2003; Scarff et al., 2004). Our results are not inconsistent with such findings since some of the most laterally located voxels did show a response preference for high frequencies. However, definite conclusions about the presence and nature of a second mirrored gradient on the far medial side of Heschl's gyrus do not seem justifiable. Other recent results in literature report the existence of various low- and high-frequency endpoints posterior to the primary auditory cortex, e.g. on the PT (Talavage et al., 2004b). Our study shows small-scale variations in optimal stimulus frequency in the PT that are reminiscent of these findings. However, we are unable to assess whether these patterns involve gradual tonotopic gradients (Schönwiesner et al., 2002) and whether they reflect a systematic organizational principle.

5.4.4 Inter-subject variability

The significance of the detected activation showed some variation across subjects. Such variation may arise from numerous causes (e.g. task-related attention levels, perceptual stimulus loudness, subject head motion, susceptibility artifacts, individual cerebro-vascular characteristics, etc.) and will lead to variable volumes of activation if the statistical threshold for significance is held constant. Because this would complicate a comparison of activation patterns and topographic mappings across subjects in this study, statistical thresholds were adjusted to obtain equal activation volumes across subjects. In all cases the resulting thresholds ($p < 0.02$ or better) were sufficiently stringent to restrict the influence of falsely positive voxels to negligibly low levels.

To determine general response characteristics, a group model was designed involving the data of all subjects simultaneously. Because error levels according to this model were substantially lower due to the larger amount of available data, a much more stringent statistical threshold could be obtained. At the same time, the estimated activation level AL equaled the average of the ALs of all subjects and was therefore representative for the average

response behavior in the subject group as a whole. By averaging out inter-subject variability in this way, the interpretation of activation images and the detection of lateralization preferences and tonotopic gradients were greatly simplified. However, conclusions applying to our subject group as a whole obviously do not necessarily extend to each individual subject. For this reason, in this chapter both individual and group results were listed throughout.

5.4.5 Sparse acquisition weighting

Imaging was performed in pairs to investigate a potential improvement of data acquisition efficiency. The SNR of all second images (B) was expected to be lower than that of all first images (A) because of multiple reasons. Firstly, the signal according to image set B may be contaminated with auditory responses to the scanner noise of the preceding scan. This will elevate the baseline signal level, and because of physiological saturation effects this may decrease the amplitude of the response that can be measured (Elliott et al., 1999; Talavage et al., 1999). Secondly, set B will have a different T_1 -relaxation contrast due to the shorter TR. Because the magnetization at the onset of the acquisitions has not yet completely relaxed back to its steady state level, the image intensity will be lower than that in set A. Although some contributions to the noise that is not explained by the model scale proportionally with the signal magnitude (e.g. physiological fluctuations), the relative importance of other noise contributions may become larger when the MR signal decreases (e.g. thermal noise), deteriorating the overall SNR. Thirdly, the hemodynamic response to the stimuli peaks after approximately 4-5 s, and will decline afterwards (Aguirre et al., 1998; Friston et al., 1998b; Belin et al., 1999; Backes et al., 2002; Pfeuffer et al., 2003). These dynamics can lead to differences in the response magnitude between two consecutive acquisitions as well, depending upon the precise timing in the paradigm relative to the response duration.

Our results indicate that the hemodynamic response signal was higher for image set A than for set B. This can be explained by the saturating influence of acoustic scanner noise and by hemodynamic response dynamics. For intermediate weightings, the transition is expected to be linear because the signal levels themselves are weighted accordingly. Furthermore, the noise in set A was smaller than in set B, which is probably caused by the effects of T_1 -relaxation. For intermediate weightings noise levels are smaller because an averaging of (partially) uncorrelated noise signals leads to a decrease in noise power.

Differences in the activation volumes AV as determined on the basis of image sets A or B were much more extreme than those in the activation level AL. Whereas an increase in the AL directly reflects changes in local blood oxygenation as a result of the interaction between neuronal metabolism and vascular hemodynamics, changes in the AV could be related to physiological processes like neuronal recruitment or increased response contributions from draining veins, but also to data processing steps like smoothing. Moreover, the magnitude of variations in the AV will depend on the distribution of *T*-values near the statistical threshold in the SPM. For instance, if many active voxels reach a significance level only just above the statistical threshold, a relatively small increase in the overall noise level (or, equivalently, a small decrease in activation level) may result in a considerable decrease in the number of activated voxels, as many will suddenly fall below the threshold. The latter mechanism provides a likely explanation why the AV shows large differences between sets A and B.

Apart from differences in the significance and level of the activation, the spatial response patterns according to all weightings were the same in the sense that the locations of the activation peaks were highly similar. The activation showed multiple stripe-like regions of activation along the Heschl's gyrus and in the planum temporale of both hemispheres, especially in the right hemisphere. Such patterns are a common finding and are proposed to reflect the location of different functional areas in the cortex (Scheich et al., 1998; Di Salle et al., 2001; Brechmann et al., 2002). The fact that responses in the left hemisphere were more diffuse may be the result of an inter-hemispheric asymmetry within subjects, but can also be caused by a larger variability in our subject population regarding the spatial location of functional and anatomical features of the left auditory cortex (Leonard et al., 1998; Rademacher et al., 2001; Tervaniemi et al., 2003).

Signal as well as noise levels indicate that set A should be weighted more strongly than set B in this experiment. If combined, the SNR characteristics lead to *T*-values that are optimal for intermediate weightings near 2:1, where optimality should be interpreted in the sense that the assignment of voxels to the set of active voxels is done with the largest degree of statistical certainty. Although achieving large activation clusters is not a goal in itself, such an optimization automatically leads to a maximal number of active voxels if the threshold for significance is kept constant. With the 2:1 weighting, more voxels were detected, with an activation level that was lower on average. Alternatively, the statistical threshold could have been raised, keeping the activation volume constant but decreasing the probability for the occurrence of false positives. Therefore, this weighting

was used for all analyses concerning response lateralization and frequency selectivity.

Although the actual SNR gain due to the acquisition of multiple consecutive functional scans will depend on the detailed parameters of the experiment, in our setup a suitably chosen weighting was shown to lead to a notably improved detection of hemodynamic responses at the cost of only a limited increase in session duration. Especially in experiments with short scans (e.g., single slice acquisitions), long repetition times (e.g., sparse paradigms), intense stimuli and/or silent scan sequences, the acquisition of multiple contiguous images promises to be beneficial.

§5.5 Conclusion

In summary, we found that at the lowest cortical processing levels both hemispheres respond most strongly to the contralateral ear. Also, we showed a tonotopic organization in the primary auditory cortices. Both organizational principles can also be found at lower levels in the central auditory system, for instance in the inferior colliculi in the midbrain and in the medial geniculate nuclei in the thalamus (Ehret et al., 1997; Yost, 2000). We interpret this as evidence that the auditory system, up to and including the level of the primary auditory cortex, primarily processes the physical properties (e.g. frequency and lateralization) of sound.

At higher processing levels in the auditory cortex (i.e., in secondary and higher cortical areas in the PP and PT) no systematic preferences with regard to stimulus lateralization and frequency were found. Apparently, the information regarding spectral content and sound source location is rearranged into a neural code that differs from the topographic principles that were found at lower levels. Presumably, the sensory input is split up into a number of streams associated with different 'objects' in the auditory environment that can be categorized and integrated with object information from other sensory modalities (Binder et al., 2004; Griffiths et al., 2004; Nelken, 2004; Zatorre et al., 2004). Our data supports the hypothesis that the transformation of the physical characteristics of sound into information about real-world objects takes place in the transition from primary to secondary auditory cortex (Kaas et al., 2000; Rauschecker et al., 2000; Arnott et al., 2004).

LATERALIZATION, CONNECTIVITY AND PLASTICITY

Chapter

6

Published in modified form:
Neuroimage 2005;28(2):490-499
doi:10.1016/j.neuroimage.2005.06.024

Abstract

Although it is known that responses in the auditory cortex are evoked predominantly contralateral to the side of stimulation, the lateralization of responses at lower levels in the human central auditory system has hardly been studied. Furthermore, little is known on the functional interactions between the involved processing centers. In this study, functional MRI was performed using sound stimuli of varying left and right intensities. In normal hearing subjects, contralateral activation was consistently detected in the temporal lobe, thalamus and midbrain. Connectivity analyses showed that auditory information crosses to the contralateral side in the lower brainstem followed by ipsilateral signal conduction towards the auditory cortex, similar to the flow of auditory signals in other mammals. In unilaterally deaf subjects, activation was more symmetrical for the cortices but remained contralateral in the midbrain and thalamus. Input connection strengths were different only at cortical levels, and there was no evidence for plastic reorganization at subcortical levels.

§6.1 Introduction

The structure and function of the central auditory system in humans is still relatively poorly understood as compared to e.g. the visual and somatosensory systems. The flow of information from the ears to the cortex has been studied intensively in animals using electrophysiological techniques, but because of their invasive nature similar studies can hardly be performed on humans. Instead, non-invasive techniques like functional magnetic resonance imaging (fMRI) have been used. However, most of these studies are confined to the auditory cortex, and only a minority encompasses the various subcortical processing centers (Guimaraes et al., 1998; Giraud et al., 2000; Melcher et al., 2000; Hesselmann et al., 2001; Backes et al., 2002; Budd et al., 2003; Yetkin et al., 2004a; Hawley et al., 2005; Krumbholz et al., 2005a).

The anatomical structure of the central auditory system is outlined schematically in Fig. 6.1. It comprises several processing centers that are distributed over the lower brainstem, midbrain, thalamus, and temporal lobes, and that perform extensive spectral and temporal processing (Nieuwenhuys, 1984; Moore, 1991; Ehret et al., 1997; Toga et al., 2000; Yost, 2000; Eggermont, 2001). Signals from the stimulated ear first enter the ipsilateral cochlear nucleus via the cochlear nerve, to subsequently continue via multiple pathways towards nuclei on both sides of the brain. Binaural information is received first in the superior olivary nuclei and is integrated to construct a representational mapping of the spatial auditory environment in the inferior colliculi. Hereafter, information is transmitted via the medial geniculate nuclei in the thalamus towards the primary auditory cortices in both cerebral hemispheres.

Although the human auditory cortex eventually receives sensory input from both ears, it is excited most strongly by stimulation of the contralateral ear (Woldorff et al., 1999; Lipschutz et al., 2002; Suzuki et al., 2002; Alho et al., 2003; Kaneko et al., 2003; Krumbholz et al., 2005b). However, in unilaterally deaf people cortical activation is known to shift towards a more symmetrical pattern (Scheffler et al., 1998; Bilecen et al., 2000; Tschopp et al., 2000; Ponton et al., 2001). An important issue concerns the extent to which such (re)distributions reflect processes that occur in the cerebral cortex itself or in subcortical structures. In animal studies evidence for changes in preferred stimulus lateralization has been demonstrated in the lower brainstem, midbrain and thalamus following unilateral loss of hearing. However, it is yet unclear at what level changes are most prominent, and reports vary from

the olivary nuclei and the colliculi to the medial geniculate nuclei (Popelar et al., 1994; McAlpine et al., 1997; Gabriele et al., 2000; Illing et al., 2000; Mossop et al., 2000; Kral et al., 2001; Kamke et al., 2003). Moreover, in humans information regarding the response lateralization at these levels is sparse, and it is yet unknown whether similar shifts exist for the subcortical nuclei in subjects with unilateral deafness.

Changes in response lateralization can presumably be attributed to changes in the strengths of functional relationships between various processing centers. Functional MRI has traditionally been used to localize the structures that are involved in the processing of auditory stimuli with certain characteristics, and to determine their response characteristics. However, most studies do not address the relations between the various centers in the processing network. The fact that multiple processing centers respond to a given stimulus supplies evidence that these centers are part of a distributed system, but provides no information on the mutual interactions. To quantify connectivity levels between the involved centers in such a network, correlations between (induced) blood oxygen level dependent (BOLD) signal fluctuations can be determined (Friston, 1994; Horwitz et al., 1999; Buchel et al., 2000). Connectivity measures are independent of activation measures in the sense that centers that show significant responses to the same stimuli do not necessarily have to show high levels of connectivity, while centers that show activation to different stimuli potentially can. Connectivity analysis is expected to supplement the information regarding stimulus activation and response lateralization with more insights concerning the signal flow in the central auditory system.

The goal of our study was to investigate the human central auditory system as a whole, using fMRI to characterize the responses in various centers, and to consider the relations between these centers building on prior knowledge gained from animal studies. Auditory stimuli were varied systematically in left and right presentation intensities to determine the lateralization characteristics in both normal hearing and unilaterally deaf subjects.

Figure 6.1 The major processing centers of the central auditory system. The ascending projections involved in the signal conduction from a single ear are shown in simplified form. Neuronal impulses that are generated in the inner ear enter the lower brainstem via the cochlear nerve at the location of the ipsilateral cochlear nucleus (CN). From here, information is fed into at least two auditory pathways respectively focusing on the content and the location of sound sources. In the monaural contralateral pathway that originates in the dorsal subdivision of the CN, spectral features of the incoming sound are enhanced and fed directly to the inferior colliculus (IC) on the contralateral side. In parallel, nerve fibers of the binaural pathway project from the anterior ventral subdivision of the CN towards the ipsilateral as well as the contralateral superior olivary complexes (SO), which are the first centers to receive direct binaural input. Its medial and lateral subdivisions are involved in the detection of differences in time-delay and sound intensity between the left and the right ear. The outgoing signals are transmitted upwards towards the inferior colliculi. The neurons in the inferior colliculus receive mostly excitatory input from the contralateral ear, and inhibitory input from the ipsilateral ear. Some of the subdivisions of the IC are thought to integrate spectral and temporal information for the construction of a spatial auditory map. Next, signals leave the mid-brain and proceed via the medial geniculate nuclei (MG) in the thalamus to arrive at the primary auditory cortices (AC) in the upper surfaces of both temporal lobes. Further information exchange between the left and right hemispheres is possible via the corpus callosum. Simultaneously, corticofugal connections (not shown) provide feedback mechanisms that influence the types of filtering and processing at lower levels.

§6.2 *Materials and Methods*

6.2.1 **Subjects**

Eight healthy subjects with normal hearing were recruited (6 male, 2 female; 7 right-handed, 1 left-handed (Oldfield, 1971); age 21-59 years, mean 30 years). In addition, five subjects with unilateral deafness (3 male, 2 female; all right-handed; age 22-59 years, mean 38 years) were selected from a patient database of the audiology department at the Maastricht University Hospital (Maastricht, the Netherlands). All subjects gave written informed consent to participate in this study, in approved accordance with the requirements of the medical ethical committee of the institute.

Subjects were examined by an audiologist prior to the fMRI exams using pure tone audiometry (Katz, 2001) to assess current hearing thresholds in a range of 250-8000 Hz. The normal hearing subjects had bilateral thresholds below 20 dB hearing level (HL), while the subjects with unilateral deafness had thresholds of at least 100 dB HL in the affected ear (1 right, 4 left) and normal thresholds (<20 dB HL) for the unaffected other ear. Unilateral deafness was acquired 5-43 years (mean 21 years) before the time of the experiment at an onset age of 6-36 years (mean 18 years).

6.2.2 **MR acquisitions**

During the MRI-sessions, subjects were placed in supine position in the bore of a clinical 1.5-T MRI system (Philips Intera, Philips Medical Systems), which was equipped with a standard quadrupolar head receiver coil. They wore a noise protection headset with built-in modified electrostatic speakers (Sennheiser HE 60, Sennheiser) and were provided with fiber-optic button-devices in both hands to register their responses to an auditory task.

A sagittal 3D T_1 -weighted fast-field echo scan was acquired for anatomical orientation (repetition time TR 10.20 ms; echo time TE 3.5 ms; flip angle 15°; matrix dimensions 256×256×75; voxel size 1.0×1.0×2.0 mm). On these high-resolution images, an imaging volume was positioned in an oblique coronal orientation, such that it enclosed the left and right cochlear nuclei (CN), superior olivary nuclei (SO), inferior colliculi (IC), medial geniculate nuclei (MG), and a substantial part of the superior surface of both temporal lobes containing the auditory cortices (AC). In this imaging volume, an additional 3D T_1 -weighted fast-field echo scan (TR 10.34 ms; TE 3.5 ms; flip angle 15°; matrix dimensions 256×256×12; voxel size 1.0×1.0×2.5 mm) was conducted to serve as anatomical reference. All functional scans were acquired covering

the same volume, and consisted of 2.5-s single-shot T_2^* -sensitive echo planar imaging (EPI) sequences with twelve contiguous 2.5-mm thick slices (TR 11 s; TE 50 ms; flip angle 90°; matrix dimensions 192×192; field of view 192 mm). To reduce the influence of acoustic scanner noise, a sparse clustered volume paradigm (Edmister et al., 1999; Hall et al., 1999; Talavage et al., 1999; Amaro et al., 2002) was employed in which auditory stimuli were presented during an 8.5-s gap of scanner silence between successive acquisitions. For each subject the functional session consisted of six runs of 30 acquisitions each, excluding initial start-up scans that were discarded.

6.2.3 Stimuli

Auditory stimuli were delivered by means of the head phones connected to a PC setup equipped with a sound card with 16-bit DAC output capability (Sound Blaster Live! Platinum 5.1, Creative Labs). The stimuli consisted of either band-passed pink noise or dynamic ripples; the latter are identical to pink noise in their average spectral content, but contain modulations in the temporal and spectral domain. All stimuli had a duration of 8.0 s and a spectral range of 125-8000 Hz; the dynamic ripples had a spectral modulation density of $\frac{1}{2}$ cycle per octave and a temporal modulation rate of 4 cycles per second. These broad-band stimuli were originally selected because of their expected effectiveness in evoking brain responses. Although the spectrally complex dynamic ripple stimuli turned out to evoke stronger responses in the AC than the noise stimuli, for the processing centers in the brainstem responses to both types of stimuli were approximately equal. In the following analyses no further distinction was made between the two types of stimuli and the corresponding data were included equally to optimize fMRI contrast-to-noise characteristics.

Stimuli were presented at intensities of 0, 40, or 70 dB above the normal hearing subjects' average threshold for the respective stimulus. Stimulus intensities were independently varied for both ears, resulting in nine different stimulus conditions that were delivered in a pseudo-random order. Each stimulus condition was presented three times per functional run, except for the pseudo-silent baseline condition (i.e., 0 dB bilaterally) which was presented six times. In the first and fourth functional run, pink noise stimuli were presented; in the four other runs, dynamic ripples were presented with an upward (second and fifth run) or downward (third and sixth run) spectral drift. Subjects were instructed to respond by left, right or bilateral button presses whenever they perceived an audible stimulus in the corresponding ear(s). The results were used only to monitor that the subjects remained focused on the stimuli during the fMRI session.

6.2.4 Preprocessing

MRI images were analyzed using the MatLab (The MathWorks Inc.) and SPM2 (Wellcome Department of Cognitive Neurology) software packages. The functional images were corrected for motion and were spatially smoothed with an isotropic 3-mm Gaussian kernel. A log-transformation was applied to the raw voxel data to obtain signal variations as a percentage relative to the baseline level.

For every subject, ten regions of interest (ROIs) corresponding with the CN, SO, IC, MG and AC on the left and right sides were broadly defined in the anatomical reference images on the basis of a brain atlas (Hanaway et al., 1998). For the single subject with an affected right ear results were inspected individually, but – apart from a reversed lateralization – these revealed no divergence from the group data. For this subject, the lateralization of the stimulus conditions and the definitions of the ROIs were mirrored in the data analysis such that all impaired subjects could effectively be treated as being deaf in the left ear.

6.2.5 Regression analysis

A general linear model (Friston et al., 1995) was set up involving a design matrix with three covariates of interest, described by the equation

$$Y_j = Y_0 + b_L \cdot L_j + b_R \cdot R_j + b_S \cdot S_j + e_j. \quad [6.1]$$

The vector Y_j represents the preprocessed intensity data of a single voxel as a function of the acquisition number ($j = 1-180$). The baseline signal Y_0 was modeled as a first-degree polynomial for each of the six functional runs separately to correct for linear drifts in the scanner signal. The values of the coefficients b were determined by minimizing the sum of squares of the errors e_j using multiple linear regression. L_j and R_j are covariate vectors containing the left and right stimulus intensities (in dB) corresponding with the consecutive acquisitions. The coefficients b_L and b_R can thus be interpreted as the measured increase in signal level as a function of the left or right intensity separately. To allow for non-linear saturation and interaction effects, a quadratic term S_j was included in the model. It was defined as the component of $(L_j + R_j)^2$ orthogonal to both L_j and R_j to exclude interference with the other response terms.

Initially, the model according to Eq. 6.1 was applied to the data of all individual voxels. Separate T -test significance levels on the coefficients b_L and b_R were color-coded to visualize the activation patterns in individual subjects. In addition, F -test statistics were calculated for every voxel to

determine the combined significance of all response terms in the model. The 10 % of the voxels in each ROI that responded most strongly according to this F -test were selected, and their data were averaged to obtain mean ROI response strengths (b) and mean ROI signals (Y_j) for each of the ten processing centers per subject.

Group statistics regarding the significance of the responses were calculated using one-sample T -tests. Responses in the corresponding left and right ROIs in response to stimulation of either ear were compared using paired-samples T -tests, and similar comparisons were made for responses to left and right ear stimulation in each ROI.

6.2.6 Connectivity analysis

In the context of fMRI, connectivity measures express the extent to which the measured signals in various processing centers are similar, under the assumption that centers that are functionally closely related will show highly similar responses (Friston, 1994; Goncalves et al., 2001; Horwitz, 2003). Unlike conventional activation measures that relate voxel signals to the stimulus paradigm, connectivity measures relate voxel signals to the signals from other brain regions that the brain tissue in the voxel is assumed to receive input from.

Figure 6.2 Schematic lay-out of the structural connectivity model, including left and right stimulus presentation intensities (L and R) and the mean ROI signals corresponding with the bilateral inferior colliculi (IC), medial geniculate nuclei (MG) and auditory cortices (AC). The cochlear nuclei and superior olives were not included in this model (dashed rectangle), since the corresponding ROI signals were relatively weak. Single-headed straight arrows indicate the ascending connections that were taken into account to determine effective connectivity levels using path analysis, allowing for correlations between the pairs of left and right ROI counterparts (double-headed curved arrows). Correlation analysis was used to calculate functional connectivity levels between the connected model components. The ϵ represent error terms.

A distinction can be made between two types of connectivity measures (Friston et al., 1997). (i) Functional connectivity considers pairs of processing centers without taking into consideration their functional surroundings, and is usually calculated as the correlation between the signals of the respective centers. (ii) Effective connectivity is calculated by optimizing the connection strengths in a structural model comprising all relevant centers such that as much as possible of the (co)variances in the measured signals are explained, using techniques like structural equation modeling. In contrast with the model-free functional connectivity measure that quantifies connectivity irrespective of the direct or indirect causal relationship between activated areas, the model-based effective connectivity measure differentiates between the influences of multiple inputs to a certain area. However, effective connectivity calculations hereby have to rely upon prior assumptions or knowledge regarding what signals and what connections to include in the model (Buchel et al., 1997; Penny et al., 2004; Ramnani et al., 2004). To combine their respective advantages, both connectivity measures were calculated in this study.

A basic hierarchical structural model was designed comprising the left and right ears and the bilateral IC, MG and AC (see Fig. 6.2). On the basis of existing literature (Nieuwenhuys, 1984; Moore, 1991; Ehret et al., 1997; Yost, 2000), both contralateral (crossing) and ipsilateral (non-crossing) feed-forward connections were assumed to exist between successive levels of the processing hierarchy (see also Fig. 6.1). The input from the ears was modeled by the left and right sound presentation intensities (L_j and R_j). The signals of the involved brain areas were estimated by the observed ROI signal Y_j (see above). All individual signals were standardized to zero mean and unit variance by an affine transformation, after which the signals of all subjects were concatenated to calculate group results. Functional connectivity levels were determined by calculating the correlation coefficients ρ between pairs of signals. Also, effective connectivity levels were determined by relating the signal of a ROI to the two inputs from the next lower level using linear regression, resulting in standardized regression coefficients β . The latter analysis is commonly referred to as path analysis and is a form of structural equation modeling (Hox et al., 1998).

§6.3 Results

6.3.1 Activation and lateralization

The significance of the hemodynamic BOLD responses to auditory stimuli that were presented to the left and the right ear was visualized by means of T-test statistic parametric maps (Friston et al., 1995). Fig. 6.3 illustrates the activation patterns in single subjects from both the normal hearing and the unilaterally deaf subject group.

Activation in the superior olives (SO) and cochlear nuclei (CN) could on average be detected in approximately half of the normal hearing subjects. Significantly active voxels to left or right stimulus presentation ($p < 0.001$) were present in 1 out of 8 subjects in the left SO, 6/8 in the right SO, 0/8 in the left CN, and 6/8 in the right CN.

Activation was very consistently found in the auditory cortices (AC), medial geniculate nuclei (MG) and inferior colliculi (IC). Active voxels were found

Table 6.1 Activation summary statistics.

Subject		ROI					
		IC _L	IC _R	MG _L	MG _R	AC _L	AC _R
NH	<i>n</i>	1 ± 1	22 ± 6*	1 ± 0	6 ± 2*	99 ± 51	372 ± 89*
		21 ± 4*	2 ± 1	10 ± 4	2 ± 2	535 ± 71**	111 ± 29*
	<i>b</i>	6 ± 2*	23 ± 3**	4 ± 1*	17 ± 4*	8 ± 3*	24 ± 4**
		22 ± 3**	4 ± 3	16 ± 3*	8 ± 5	38 ± 4**	11 ± 2**
	<i>z</i>	1.6 ± 0.3**	5.1 ± 0.9**	1.3 ± 0.2**	2.8 ± 0.3**	2.1 ± 0.3**	4.6 ± 0.6**
		4.8 ± 0.4**	1.2 ± 0.4*	3.2 ± 0.4**	1.5 ± 0.3*	6.1 ± 0.6**	2.2 ± 0.2**
UD	<i>n</i>	0 ± 0	0 ± 0	1 ± 1	1 ± 1	74 ± 47	74 ± 45
		8 ± 4	0 ± 0	8 ± 4	0 ± 0	281 ± 96*	114 ± 51
	<i>b</i>	8 ± 3*	4 ± 4	4 ± 3	0 ± 3	7 ± 4	17 ± 9
		14 ± 2*	7 ± 2*	13 ± 7	0 ± 3	27 ± 10*	18 ± 6*
	<i>z</i>	1.5 ± 0.4*	1.4 ± 0.3*	1.4 ± 0.1**	1.7 ± 0.2**	2.2 ± 0.5*	2.0 ± 0.5*
		3.2 ± 0.5*	1.4 ± 0.3*	3.0 ± 0.5*	1.5 ± 0.3*	4.2 ± 0.9*	2.5 ± 0.5*

For the normal hearing (NH) or unilaterally deaf (UD) subjects, and for each of the defined ROIs corresponding with the left (L) and right (R) inferior colliculi (IC), medial geniculate nuclei (MG) and auditory cortices (AC), the response to left (upper entry) and right (lower entry) ear stimulation is shown. Responses are quantified by means of (i) the number of voxels (*n*) in the ROI that were significantly active ($p < 0.001$), and (ii) the average activation level (*b* [10^{-3} %/dB], see Eq. 1) and (iii) the mean *z*-value of the 10 % most active voxels in the ROI. Numbers indicate mean ± standard error across the groups (*: $p < 0.05$; **: $p < 0.001$).

in 6 out of 8 subjects in the right MG, while for the left MG as well as for the bilateral IC and AC activation was found in 8/8 subjects. Because response levels were insignificant across subjects for the CN and SO, these nuclei will not be discussed further.

In Fig. 6.4 the mean signals Y_i of the 10 % most active voxels in each ROI in the normal hearing subject group are plotted as a function of the stimulus

Figure 6.3 Activation patterns to left and right ear stimulus presentation in a normal hearing subject. In order to show the left and right cochlear nuclei (CN), superior olivary nuclei (SO), inferior colliculi (IC), medial geniculate nuclei (MG) and auditory cortices (AC) in a single image, this cross section is composed of multiple oblique coronal slices (outlined by black lines). Activation appeared mainly ipsilateral in the CN, bilateral in the SO, and predominantly contralateral in the IC, MG and AC. (This figure continues on the next page.)

condition with respect to left and right stimulus intensity. In general, an increase of the signal with increasing stimulus intensity was found. For the AC, MG and IC the strongest responses were evoked by contralateral stimulus presentation; for ipsilateral stimuli, responses were smaller. Non-linear response behavior was evident and significant in the AC only, and was likely due to saturation or inhibition processes (Krumbholz et al.,

2005a).

Table 6.1 lists the summary statistics for the volume, level, and voxel-wise significance of the responses to stimulation of either ear in the bilateral IC, MG and AC. In these brain regions, activation was most extensive (i.e. larger n) contralateral to the side of stimulation for normal hearing subjects. The average responses levels were also strongest (i.e., larger b) contralateral to

Figure 6.4 The average ROI signal data Y_i for all normal hearing subjects are averaged and plotted as a function of the left (L) and right (R) stimulus intensities relative to a mean baseline level. Responses for the ROIs corresponding with the left and right auditory cortex (AC), medial geniculate nuclei (MG), and inferior colliculi (IC) are shown. For all ROIs, responses were predominantly contralateral in nature. Furthermore, in the AC non-linear response saturation is observable. (*This figure continues on the next page.*)

the ear of stimulation in this group.

Comparisons of the activation in the left and right IC, MG and AC, respectively, showed that the brain area contralateral to the ear of stimulation showed the strongest response ($p < 0.05$); e.g. for right ear stimulation the b_R of the contralateral left IC is larger than that of the ipsilateral right IC. Correspondingly, for each of these brain areas, the

Figure 6.4 (continued) Similarly, responses were plotted for the unilaterally deaf subjects. The response characteristics for ROIs on the left side, contralateral to the unaffected ear, were similar to those in normal hearing subjects for the AC, the MG as well as the IC. In the right AC responses were detectable to stimulation of either ear, possibly mediated by cross hearing and a stronger input from the unaffected ear at this level. In contrast, responses in the right IC and MG had largely vanished, indicating that plastic reorganization to obtain stronger input from the ipsilateral ear is negligible.

average responsiveness to the ear on the opposite side of the head is significantly larger than that to stimulation of the ear on the same side of the head ($p < 0.05$); e.g. for the left IC, b_R is significantly larger than b_L . The average z -values of the 10 % most active voxels that were selected in each ROI reflect the same predominantly contralateral activation patterns.

In unilaterally deaf subjects, activation patterns in response to stimulation of the unaffected right ear were similar to those in normal hearing subjects (see Fig. 6.3, Fig. 6.4 and Table 6.1). The response of the midbrain and thalamus (i.e., the IC and MG) to stimulation of the hearing ear was dominantly contralateral and similar to that in the normal hearing subject group. In contrast, activation of the AC was more symmetrical as compared to normal hearing subjects.

Stimulation of the deaf ear caused weak brain responses. In the IC and the MG, these responses were strongest ipsilateral to the deaf ear. Thus, the responses are consistent with weak cross talk to the hearing ear (Tschopp et al., 2000; Suzuki et al., 2002).

6.3.2 Connectivity measures

Two types of connectivity measures were determined for the connections in the structural connectivity model that is pictured in Fig. 6.2. Table 6.2 lists the functional connectivity levels ρ as well as the effective connectivity levels β . All effective connectivity levels were smaller than the corresponding functional connectivity levels. This is related to the fact that the bilateral counterparts of all ROIs showed signals with a positive mutual correlation. The significance of the functional connectivity levels was generally very high, and exceeded those of the corresponding effective connectivity levels. The model-free functional connectivity approach and the model-based effective connectivity approach strongly corroborated each other, as both measures showed the same trends. Therefore, in the following discussion both connectivity measures will be considered together.

At the lowest hierarchical level in the model, the contralateral (crossing) connections from the level of the actual stimulus presentation to the IC were evidently stronger than the ipsilateral (non-crossing) connections in normal hearing subjects. In the impaired subjects, the connection from the unaffected right ear to the contralateral IC was strongest, while the contralateral connection emanating from the deaf left ear was decreased in strength most notably.

For the connections from the IC to the MG, ipsilateral connections were stronger than contralateral ones in the normal hearing subjects, although the

Table 6.2 Functional and effective connectivity levels.

Connection	Normal hearing		Impaired hearing	
	ρ	β	ρ	β
L $\leftrightarrow$ R	0.09**	-	0.09*	-
L $\rightarrow$ IC _L	0.14**	0.10*	0.17**	0.14*
L $\rightarrow$ IC _R	0.43**	0.43**	0.07*	0.06
R $\rightarrow$ IC _L	0.42**	0.41**	0.32**	0.30**
R $\rightarrow$ IC _R	0.09**	0.06	0.12**	0.12*
IC _L $\leftrightarrow$ IC _R	0.36**	-	0.24**	-
IC _L $\rightarrow$ MG _L	0.34**	0.30**	0.29**	0.25**
IC _L $\rightarrow$ MG _R	0.17**	0.09*	0.14**	0.13*
IC _R $\rightarrow$ MG _L	0.23**	0.12*	0.22**	0.16*
IC _R $\rightarrow$ MG _R	0.24**	0.21**	0.08*	0.05
MG _L $\leftrightarrow$ MG _R	0.24**	-	0.21**	-
MG _L $\rightarrow$ AC _L	0.36**	0.34**	0.23**	0.23**
MG _L $\rightarrow$ AC _R	0.16**	0.12*	0.29**	0.25**
MG _R $\rightarrow$ AC _L	0.15**	0.06*	0.05	0.00
MG _R $\rightarrow$ AC _R	0.21**	0.18**	0.22**	0.17*
AC _L $\leftrightarrow$ AC _R	0.59**	-	0.43**	-

ROI signal correlation coefficients ρ and standardized regression coefficients β , respectively expressing the functional and the effective connectivity levels for the structural model of Fig. 6.2. The significance of the coefficients is indicated with asterisks (*, $p < 0.05$; **, $p < 0.001$).

connection on the right side was less strong than its left counterpart. In the impaired subjects, the ipsilateral connection on the side contralateral to the hearing ear remained relatively strong, while that on the other side was negligible in strength.

In normal hearing subjects, ipsilateral connections between the MG and AC were also stronger than contralateral ones, particularly on the left side. In the impaired subjects, the connections emanating from the MG contralateral to the deaf ear were weakest. In this subject group, the AC on this side seemed to dominantly receive information through connections from the contralateral MG.

In summary, according to the connectivity analysis each ear was mostly connected to the contralateral IC in normal hearing subjects, and from the IC onwards connections remained on the same side of the brain. In unilaterally deaf subjects, the connectivity scheme of the normal hearing ear was similar

Figure 6.5 Schematic information flow in the central auditory system in normal hearing and unilaterally deaf subjects according to the connectivity analysis. To highlight the flow of auditory information in normal hearing and impaired subjects respectively, only the connections for which the effective connectivity levels reached a significance $p < 0.001$ are shown (see also Fig. 6.2 and Table 6.2). In the normal hearing subjects, information appears to cross below the level of the IC and continues ipsilaterally hereafter. In the impaired subjects, the signal flow from the normal right ear remains intact but that from the deaf left ear has vanished. Only at a cortical level were changes in input connection strengths apparent, leading to more symmetrical cortical activation in this subject group.

to that in normal hearing subjects, although the MG contralateral to this ear was connected to both AC. In this group, the signal flow from the deaf ear had essentially vanished as compared to the normal hearing subjects. The general picture that emerged is illustrated in Fig. 6.5 for the two subject groups. In order to illustrate only the most important features regarding signal flow, only the most significant connections from the structural model are shown in these figures by thresholding the significance of the effective connectivity levels β at $p < 0.001$.

§6.4 Discussion

6.4.1 Normal hearing

Using a standard 1.5-T fMRI system, activation could be detected in multiple processing centers of the central auditory system, including the cochlear nuclei (CN) and superior olivary nuclei (SO) in the lower brainstem, the inferior colliculi (IC) in the midbrain, the medial geniculate nuclei (MG) in the thalamus, and the auditory cortices (AC) in the cerebral temporal lobes. Activation in the IC, MG, and AC was detected most consistently and was quantified for all subjects.

For the AC, activation was strongest on the side contralateral to the stimulation, with activation levels being approximately three times as high as those on the ipsilateral side. Regarding the activation volume (i.e., the number of active voxels) even more extreme ratios were obtained. These values are in good agreement with findings from literature (Scheffler et al., 1998; Woldorff et al., 1999; Suzuki et al., 2002). Quite similar ratios between left and right activation levels were observed for the IC and the MG. Few previous fMRI studies have considered these subcortical nuclei, and their results regarding response lateralization are contradictory. For instance, activation to monaural stimulation in the IC is reported to be ipsilateral (Giraud et al., 2000), contralateral (Melcher et al., 2000), as well as bilateral (Yetkin et al., 2004a). Our findings strongly suggest that the lateralized response in the cortex directly reflects that in the thalamus and midbrain.

By performing connectivity analyses we were able to confirm that auditory signals dominantly cross to the contralateral side below the level of the IC. In our model, connections from the level of the stimulus delivery that crossed to the contralateral IC were notably stronger than those that remained on the same side. At levels from the IC to the MG and from the MG to the AC the converse was true: ipsilaterally conducted signals appeared more influential than signals crossing to the contralateral side.

To relate our results to the more detailed electrophysiological measurements in animal studies, the relationship between fMRI activation and neural activity must be taken into consideration (Horwitz, 2003; Buxton et al., 2004; Logothetis et al., 2004a). The BOLD signal reflects vascular changes that result from variations in metabolic activity in the surrounding brain tissue. These are induced by increases or decreases in neural activity, which comprises excitatory and inhibitory synaptic events as well as neuronal discharges. Although the relative importance of the latter contributions to the total metabolic demand is still a topic of debate, it appears that synaptic

processes play a more important role than spike activity (Arthurs et al., 2002; Buxton et al., 2004). The energy demands related to excitatory input in particular may account for most of the variations in brain metabolism and are assumed to be the major factor contributing to the BOLD fMRI signal (Waldvogel et al., 2000; Logothetis et al., 2004a).

The IC receives excitatory and inhibitory input from both ears. In non-human mammals, the majority of neurons in the IC is excited by a contralateral stimulus and inhibited by an ipsilateral stimulus (Ehret et al., 1997). We found that the BOLD response is largest in the IC contralateral to the ear of stimulation. Since the BOLD response is presumably dominated by excitatory synaptic input, our results are consistent with those in non-human mammals: excitation in the colliculus is dominated by stimulation from the contralateral ear. A similar conclusion holds at higher levels for the dominantly contralateral excitatory input to the MG and AC (Imig et al., 1977; Rodrigues-Dagaëff et al., 1989), consistent with the contralateral responses in this study.

6.4.2 Unilateral deafness

In unilaterally deaf subjects, stimulation of the unaffected ear led to similar activation patterns as in normal hearing subjects. However, the response of the left and right AC was more symmetrical than in normal hearing subjects. This is in agreement with other reports on the response lateralization in the auditory cortex (Scheffler et al., 1998; Bilecen et al., 2000; Tschopp et al., 2000; Ponton et al., 2001). Extending these earlier studies, we showed that for the MG and the IC activation was largely contralateral in nature. Thus, in the IC and MG, the response to stimulation of the hearing ear is comparable to that in normal hearing subjects. These results suggest that plastic reorganization of response lateralization takes place in the cortex, and not in the brainstem or thalamus. Possibly, such reorganization is mediated by callosal connections with the contralateral hemisphere, as the corpus callosum also transmits information on behalf of the integration of binaural auditory signals (Lessard et al., 2002).

It is somewhat surprising that the right IC and MG that would normally respond most strongly to the left ear show no evidence for an increase in response to the remaining input from the unaffected right side when the left ear has become deaf. Given the consistency with which activation of the IC and MG could be detected in this experiment, some detectable activity would have been expected if plastic reorganization played a significant role.

Some animal studies have proven the possibility of adaptation processes at and below the level of the IC leading to both short- and long-term increased collicular responses to ipsilateral stimuli (Popelar et al., 1994; McAlpine et al., 1997; Illing et al., 2000; Mossop et al., 2000). The changes in response characteristics might be a result of unmasking of normally inhibited inputs or dynamic processes of reorganization (Irvine et al., 2003). However, other studies suggest a dominant involvement of the thalamus (MG) and cortex (AC) in plastic reorganization (Kamke et al., 2003), which is more consistent with the shifts in activation in our study.

The discrepancy between the observation of reorganization at the thalamic and collicular levels in some animal studies and the lack of changes in our results may also be related to the fact that the BOLD response reflects mainly the synaptic input activity and not the neuronal discharge rate (Waldvogel et al., 2000; Arthurs et al., 2002; Buxton et al., 2004; Logothetis et al., 2004a). Changes in response characteristics of neurons to given signals could therefore remain undetected using fMRI if the input to these neurons remains the same. At the same time, the spike rates that would be found using electrophysiological measurements would be affected by such changes. In the case of the IC, the measured hemodynamic responses would reflect the activity in synaptic inputs from the CN and the SO. Thus, the absence of shifts in fMRI activity in the IC of impaired subjects might merely reflect a lack of plastic reorganization at lower levels in the central auditory system. However, following this line of reasoning, plastic changes occurring in the midbrain and thalamus could affect the input activity at higher levels in the processing chain and induce detectable changes at a cortical level such as found in this study. Therefore, our findings do not necessarily preclude plastic changes in neuronal firing behavior to given input at the level of the IC and possibly the MG.

It has also been suggested that the monaural contralateral pathway from the CN to the IC is best preserved in deaf subjects (Alwatban et al., 2002). Since this pathway is not known to contain ipsilateral connections below the level of the IC, input from the ear on the same side of the head cannot be received directly through this pathway. Although plastic changes following acquired deafness could affect the strength of existing connections, it is less likely to lead to the formation of new fiber tracts. Therefore, the monaural contralateral pathway does not provide any means for the deprived IC to obtain input from the normal hearing ear.

In contrast, the binaural pathway does provide ipsilateral input, but plastic reorganization may remain absent in this pathway because its primary function presumably concerns spatial sound source localization and this

hardly benefits from newly acquired input from one ear only. Thus, we speculate that since there may not be a way that the brain benefits from renewed binaural connections, there is no developmental drive to reuse the connection from the hearing ear to the ipsilateral IC, and there is no gain from plastic reorganization at the level of the brainstem.

Finally, the absence of plastic reorganizations in the midbrain in our study may be related to the fact that for all subjects unilateral deafness was acquired several years after birth, while maturation of the auditory pathways in the brainstem takes place in the first two years of life (Salamy et al., 1985; Inagaki et al., 1987). It has been claimed that although neonatal unilateral hearing loss leads to a rearrangement of binaural connections in the auditory midbrain and to changes in the physiology of collicular neurons in response to stimulation of the unaffected ear, the same hearing loss in adulthood does not produce such changes (Moore, 1991). Additional fMRI research similar to this study but involving a group of subjects with congenital unilateral hearing loss may be able to resolve this issue.

§6.5 Conclusion

In summary, we found that the activation in the midbrain and thalamus is predominantly contralateral in nature, similar to the lateralization of the responses to auditory stimuli in the cerebral cortex. In unilaterally deaf subjects activation in the cortex is more symmetrical as compared to normal hearing subjects, but at the mentioned lower processing levels responses remain strongly contralateral. This was observable in the activation patterns, and was furthermore confirmed by connectivity analyses. These findings suggest that plastic reorganization processes play a more important role at the cortical level than at sub-cortical levels.

GENERAL DISCUSSION AND CONCLUSIONS

Chapter

7

Abstract

In this thesis, possible improvements of the experimental paradigm for functional magnetic resonance imaging of the central auditory system were investigated. Also, a number of essential characteristics of the central auditory system were explored in more detail. Furthermore, comparisons were made with subjects that suffered from various types of hearing loss that were expected to affect these characteristics. In the following paragraphs these topics will be discussed and further speculated on in relation to the results that were obtained and described in the previous chapters.

§7.1 Functional Magnetic Resonance Imaging

7.1.1 Acoustic scanner noise

Functional magnetic resonance imaging (fMRI) is a powerful technique for purposes of neuroimaging in humans due to its non-invasive nature and its superior spatial resolution. However, investigations of the central auditory system (CAS) are hindered by the presence of intense acoustic scanner noise. In the past, sparse acquisition paradigms have been developed to reduce the influence of acoustic scanner noise. In this thesis variations of this method were considered to investigate the merit of potential improvements hereupon.

In Chapter 2, the scan-to-scan and the stimulus-to-scan intervals were varied systematically. A method was shown to overcome variations in baseline signal levels due to longitudinal relaxation effects. This allowed a distinction to be made between the undisturbed response to auditory stimuli and the interaction effects with the background scanner noise. It was shown that the acoustic scanner noise of only a single previous scan can severely reduce the detectable hemodynamic response levels. A simple model was used to predict the influence of acoustic scanner noise and the magnitude of the interaction effects. Although some deviations from the model were found, especially for auditory stimuli with spectral resemblance to the acoustic scanner noise, the model prediction that interaction effects are maximal near the peaks of the responses to the previous stimulus and the previous scanner noise was supported by the results. It follows that the scan-to-scan interval (and thus the repetition time TR) should be chosen in the order of the total duration of the hemodynamic response itself, i.e. in the order of 10 s, to achieve optimal sensitivity while still accomodating a maximal number of scans in a session with given duration. At the same time the duration of the acquisitions should be chosen well below the time-to-peak of the hemodynamic response, being in the order of 4-5 s. These findings corroborate earlier results and give a better foundation to recommendations regarding scan-to-scan interval and scan duration (Talavage et al., 1999).

In Chapter 5 a variation of the standard sparse acquisition scheme was tried in which not only the acquisition of multiple slices was clustered, but also the acquisition of multiple image volumes. It was shown that in spite of the increasing importance of acoustic scanner noise and in spite of the detrimental effect of a small TR on the magnetization and signal magnitude, some improvement in signal-to-noise ratio (SNR) characteristics can be achieved by the use of repeated acquisitions. Although in the presented

study the gain was similar to that what would have been obtained by the increased acquisition density in the absence of repeated acquisitions, the method promises to be more advantageous for various other experimental settings. For instance, for shorter acquisitions times (e.g., single slice acquisitions), longer silent periods between clustered scans, less noisy acquisition types, or for stimuli with stronger or longer lasting responses, the inclusion of multiple consecutive repeated acquisitions is likely to be beneficial for the attained sensitivity to hemodynamic response signals.

Driven by considerations regarding patient comfort, future developments in scanner technology will lead to the production and clinical availability of inherently more silent configurations of magnetic resonance scanners. However, at the same time stronger main magnetic field and gradient strengths and higher gradient switching speeds will be used to improve on spatial resolution and SNR, causing mechanical vibrations and sound intensities to rise (Di Salle et al., 2003b). Therefore, it does not seem likely that acoustic scanner noise problems will be resolved by generally available technological measures in the near future. Although much can be gained in setups specifically dedicated to research, particularly in combination with passive and active noise attenuation and quieter acquisition sequences, acoustic scanner noise will remain present in the near future and will interfere with the perception of auditory stimuli in fMRI experiments on the CAS. Apart from the separation in time of the hemodynamic responses to auditory stimuli and scanner noise, sparse clustered volume acquisition paradigms have the great advantage that stimulus presentation takes place during virtually silent periods of scanner inactivity. Such paradigms will remain advantageous as compared to traditional acquisition schemes with continuous scanning. Thus, studies on the dynamics and interactions of auditory hemodynamic responses, including responses to acoustic scanner noise in particular, and the development of models to account for such effects, will remain of practical importance in the future of auditory fMRI.

7.1.2 Response linearity

Implicit assumptions regarding the linearity of the hemodynamic response to sensory stimulation are of utmost importance for the interpretation of fMRI results, and for the comparison of response magnitudes to various stimuli within and between fMRI studies.

In Chapter 2 it was shown that acoustic scanner noise has a detrimental influence on the amplitude of the detectable responses to other, independently presented, auditory stimuli. Also, response interactions were

most complicated for sound stimuli that resembled each other. This is clear evidence for the presence and importance of saturation effects and non-linear response behavior. In contrast, in Chapter 3 it was demonstrated that non-linear effects play only a minor role when the intensity of auditory stimuli was varied even over a large range, from barely audible stimulation near threshold to loud stimuli around 70 dB (stepping over the fact that the dB-scale in itself expresses Weber's law and is highly non-linear with respect to physical sound attributes). Although non-linearities were significant for some types of stimuli in some subjects, they were not very influential overall. Only at stimulus intensities near threshold did non-linear response characteristics potentially arise. In Chapter 6, saturation effects were observed, but these were significant only in the auditory cortex. Thus, findings regarding the importance of non-linearities varied.

The brain is a highly complex organ showing very intricate response behavior. Already at the level of individual neurons, responses are highly non-linear. Adding that metabolic and vascular variations have a limited dynamic range, it should come as no surprise that the brain as a whole displays non-linear hemodynamic response characteristics. Nevertheless, non-linear systems can show linear responses to subsets of input signals. Especially for small variations in input signals, linear response behavior dominates. Therefore, the finding of linearly varying responses under certain circumstances does not yet form evidence for the conclusion that parts of the brain, and the auditory system in particular, perform linear processing. It is a much more difficult question that needs answering: under which conditions does the auditory system behave in a (pseudo-)linear way? Although this question is too broad to be answered in this thesis, some important trends can be derived from the presented studies.

First, non-linearities in the vascular dynamics may have influence on the fMRI signal at high stimulus intensities. This once more stresses that the hemodynamic response is an indirect measure of brain activation and that the interpretation of fMRI results should be carried out with this in mind.

Second, non-linear threshold effects may arise at low stimulus intensities. These are most likely of neural or possibly cochleomechanical origin, and can manifest themselves as a strong initial increase in activation as a function of stimulus intensity just above the auditory threshold, followed by a more gradual rise at higher sound levels.

Third, non-linearities emerge most clearly when various stimuli interact. Although variation of the intensity of a single stimulus leads to essentially linear response variations, the presentation of multiple stimuli can lead to strong non-linear response contributions. Non-linearities seem most

important if stimuli with widely varying characteristics are superimposed. This might be related to the identification of multiple auditory objects at cortical levels in the central auditory system.

Fourth, non-linear effects seem to play a more important role at high levels of auditory processing (e.g., the auditory cortex) than at lower levels (e.g., the midbrain). Although non-linear signal processing is taking place in the brainstem, overall responses as measured using fMRI are more or less linearly dependent upon stimulus intensity. Also, it was suggested that with regard to stimulus intensity, the primary auditory cortex might show a more linear dependence than the secondary cortex. Such trends are likely related to the more complex types of processing that take place at high levels, and to the fact that the representation of the auditory environment at high levels is less directly related to the sound waveform.

7.1.3 Activation and percept

In Chapter 3, responses in normal-hearing subjects to stimuli with varying intensities were compared with those in subjects with moderately severe hearing impairment. It was shown that the response characteristics were different between the two groups if expressed as a function of physical attributes (i.e., intensity), but did not differ significantly if expressed as a function of perceptive attributes (i.e., loudness). This explicitly shows that the activation in response to auditory stimulation is more closely related to the (subjective) percept of the stimulus than to its (objective) physical characteristics. Vice versa, this forms proof that the subjective percept of a stimulus directly reflects the level of the brain activation that is induced by the stimulus.

Also, in Chapter 4 the sensitivity of the auditory cortex to spectro-temporal modulations was systematically measured. The results corresponded well with other findings that were obtained using psychophysical measurements, both in qualitative and in quantitative sense (Chi et al., 1999). This again suggests that cortical activation is closely related to the subjective percept of a stimulus.

It is not trivial that the brain response in the auditory cortex should reflect the subjective percept. The neural correlate of a subject's experience of a stimulus is hardly known, but it seems reasonable that the subjective percept is somehow the end result of all auditory processing, both at sensory and cognitive levels. Only a fraction of the information that is available in the stimulus waveform is transduced by the peripheral ear and transmitted to the brain in the form of neuronal discharge patterns. During the various

processing stages in the lower brainstem, midbrain, thalamus, temporal lobe, and multimodal cortical areas, information is re-encoded and compacted further. Therefore, the auditory signal is transformed in various ways, and only the end product of this processing chain is expected to be directly related to the subjective experience of the stimulus. In pathological conditions, the way in which auditory information is processed at various levels may change markedly, possibly along with the subject's perception. Although more research is necessary, the finding in this thesis that the brain response in the temporal lobe relates well to the subjective experience suggests that the encoding of auditory information has reached a final form in the auditory cortex as far as basic sensory aspects of sound stimuli like loudness and spectro-temporal modulations are concerned (for stimuli with e.g. language content the situation might of course be different).

Also, these findings support the hypothesis that the fMRI response in the cortex not only can be considered to be a measure of the underlying physiological processes, but also reasonably well describes the intensity of the experience that the subject is having. Thus fMRI might provide a means for quantifying subjective perception in a more objective way.

7.1.4 Technological and methodological prospects

Functional MRI is a relatively novel technique and has been used in research environments for approximately ten years now. The application of fMRI to the central auditory system in particular generally results in unorthodox experimental setups. Although many promising developments of basic importance are still taking place, in Chapter 6 some relatively new types of analyses related to connectivity were already applied in the context of a non-methodological study.

In the field of data analysis many noticeable developments are going on. Although regression analysis on the basis of a general linear model will probably remain the working horse in the field of fMRI, especially for routine experiments (from a technical viewpoint), other techniques that rely less on prior assumptions are gaining strongly in importance. Independent component analysis (ICA) is becoming more fashionable, although the interpretation of its results remains difficult and requires validated methods that are not yet available. Yet, this and related techniques will open up the road to experiments in which the detailed stimulus conditions are of lesser importance, eliminating the prior formulation of contrasts and enabling the identification of resting state networks in the auditory system for example. Somewhat related, various types of connectivity analyses are starting to

supplement pure activation data, allowing the brain to be analyzed as an interactive network rather than as a collection of response areas. In parallel, cortical flattening is getting more common in the field of analysis, registration and visualization. Although still burdened with its own difficulties, it forms a valuable complement to volumetric and cross-sectional techniques and is increasingly worthy of consideration.

In addition, an area for potential improvements is the blood oxygenation level dependent (BOLD) contrast mechanism itself. BOLD fMRI suffers from small contrasts between the activated and resting states as compared to e.g. electroencephalography (EEG). Also, the temporal resolution of the BOLD effect is very poor, and there can be a spatial mismatch between the sites of neural activation and the hemodynamic response. As research is focusing on more detailed aspects of cortical processing, these shortcomings become more and more restrictive. By combining T_2 and T_2^* images, a distinction can be made between signal contributions that originate from large and from small venules. By detecting signals from the capillary bed only, a better coregistration of the detected hemodynamic response with the actual site of neural activity may be obtained, leading to a gain in spatial resolution. Completely different contrast mechanisms might further improve on BOLD fMRI. However, this requires much more insight into the precise dynamics that give rise to the hemodynamic response. As techniques have come available to independently measure cerebral blood flow, volume, and oxygenation levels, new insights into these response mechanisms should be expected in the near future. Possibly, this will give rise to improved experimental paradigms by combining BOLD imaging with arterial spin labeling, vascular space occupancy, or related techniques. Along a different route, other recent developments of importance comprise the combination of fMRI with EEG measurements, or with fiber tracking using diffusion tensor imaging.

§7.2 *The Auditory System*

7.2.1 Hierarchical processing

The central auditory system is known to consist of multiple processing centers, distributed across the lower brainstem, midbrain, thalamus, and temporal cerebral cortex, forming a hierarchy of processing levels. At each level, input is obtained from lower levels, various types of processing take place, and the resulting output signals are transmitted to higher levels. Input

is also obtained from higher levels and transmitted to lower levels via efferent connections, providing important feedback mechanisms.

In the majority of the studies that were performed, imaging was limited to the upper temporal cortical surface. This was done because of practical limitations (e.g., restrictions on acquisition times) or because of inherent interest in the organization of the cortex only. Also, functional imaging of the brainstem is notoriously difficult due to the increased influence of motion artifacts and because the sub-cortical nuclei are small in size (Komisaruk et al., 2002). Nevertheless, in Chapter 6 all major processing centers of the central auditory system were included. Activation in the lower brainstem was detectable in approximately half of the subjects, but was insignificant for the subject groups as a whole. The nuclei in the midbrain and thalamus could be detected very reliably.

Functional MRI is sensitive to brain activation, but it is much harder to derive information about causal relationships and the flow of information between the various centers. However, by introducing a priori knowledge or assumptions (e.g., based on animal research), the relationship between regions can be established using connectivity analysis. In Chapter 6 such analyses were performed, focusing on response properties regarding stimulus lateralization. It was found that signals from the two ears cross at low levels in the brainstem, and dominantly remain on the same side of the brain hereafter. This complemented the findings of predominantly contralateral activation in the inferior colliculi, medial geniculate body, and auditory cortex. The lateralization of the activation in the cortex thus directly reflects that at immediately lower levels.

However, evidence was found for a hierarchical organization within the auditory cortex itself as well. In Chapter 5 it was shown that the internal representation of sound signals differ between the primary and secondary auditory cortices. In the primary auditory cortex, strong representations of sound lateralization and tonotopic organization were found. The hemispheres were most responsive to sounds from their contralateral hemifield and signals from the contralateral ear. In addition they showed a tonotopic progression from low to high sound frequencies in the lateral-to-medial direction. However, in the secondary regions these organizational principles were much weaker or absent. These findings show that a clear functional distinction can be made between the primary and secondary regions regarding their position in the processing chain.

In Chapter 4 gradients were found with respect to the sensitivity to the spectral density and temporal frequency of modulations in dynamic rippled noise stimuli. It was found that the auditory cortex is relatively more

responsive to stimuli with broad and slow spectral and temporal dynamics towards the posterior and lateral sides respectively. In other words, the primary auditory cortex on the anteromedial side was most responsive to fine-scale spectral dynamics, while the surrounding secondary regions were more strongly excited by large-scale spectral characteristics. Under the assumption that large-scale acoustic information is obtained from the spectral and temporal integration of small-scale features, these results form evidence for a hierarchical organization where the secondary areas form a higher processing layer that integrates the signals from the primary regions.

7.2.2 Topographic cortical mappings

In various studies in this thesis evidence was described for three independent gradual topographic mappings on the superior cortical surface of the temporal lobe. (i) In Chapter 4 a gradient was shown in the lateral-to-medial direction with respect to the optimal sensitivity to temporal modulations, with lateral regions responding relatively well to slow modulations and medial regions responding to fast modulations. (ii) Also in Chapter 4, a posterior-to-anterior gradient was similarly shown for spectral modulations, with posterior regions responding relatively well to broad modulations and anterior regions responding to dense modulations. (iii) In Chapter 5, a lateral-to-medial gradient was shown with respect to the optimal stimulus frequency (i.e., tonotopy), with lateral regions responding relatively well to low frequencies and medial regions responding to high frequencies. In the latter chapter preferences with regard to stimulus lateralization were also mapped, but these did not result in gradual gradients.

Gradients, and in particular tonotopic gradients, have been described many times in animal studies and form a means for the distinction between various cortical sub-areas on the basis of functional arguments. The presence of such gradients may also provide pointers to the identification of the type of functional processing that is taking place in the respective areas. In humans, the subdivisions of the primary and secondary auditory cortex have not yet been revealed convincingly, let alone has their function been identified. Tonotopic organizations have been studied quite extensively, but apart from the large-scale gradient in the lateral-to-medial direction along Heschl's gyrus that has been described relatively consistently, results are rather unspecific. Furthermore, it can be questioned whether the gradient that has been identified in humans using fMRI corresponds with the gradients in functional sub-areas in animal studies, given the variations in scale and form. Certainly, the 'modulotopic' gradients that were identified in

Chapter 4 should not be interpreted as the result of a functional characteristic of some sub-area, but should rather be seen as a large-scale feature of the auditory cortex as a whole or as a shift in function from sub-area to sub-area. Although a direct interpretation of the observed gradients in the context of local function and organization is thus not without ambiguity, they provide valuable guidelines for the understanding of cortical auditory processing and may prove to be essential to detect functional abnormalities in patients with various types of hearing disorders.

7.2.3 Plasticity

Neural plasticity is the ability of neural circuits to undergo changes in function or organization due to previous activity. This comprises both short-term effects like neural facilitation and habituation, and long-term effects like learning and memory, or development and maturation. One potential cause that may induce plastic effects is the occurrence of long-lasting systematic changes in sensory input, for instance as a result of disorders in the peripheral hearing organ like partial deafness or hearing loss.

In this thesis two studies were described involving subjects with hearing disorders. In Chapter 3 patients with moderately severe high-frequency hearing loss were studied. It was found that the relation between sound intensity and neural activation was disturbed as compared to normal hearing subjects. The perceptual loudness of sound stimuli was found to increase disproportionately strongly with sound intensity (i.e., loudness recruitment), as was the hemodynamic response activation. This phenomenon can be related to changes in the physical transfer characteristics of the inner ear following the loss of hair cells in the cochlea, but changes in response characteristics at a neural level are also likely to play a role.

In Chapter 6 patients that suffered from unilateral deafness were studied. It was found that responses at the level of the brainstem and thalamus hardly differed from those in normal hearing subjects in qualitative sense. However, at the level of the cortex a more symmetrical distribution was found. This forms evidence that in this subject group plastic changes have taken place at the level of the auditory cortex or in the strength of its input connections from the bilateral medial geniculate nuclei in the thalamus. In more detail, it was found that the deprived nuclei in the midbrain and thalamus did not seem to obtain new input from the unaffected contralateral ear, while the deprived cortical hemisphere did show some increase in response to stimulation of that ear.

These results seem to suggest that the connections in the brainstem are more 'hard-wired' than in the cortex and show less flexibility towards changes in connectivity levels. However, from animal studies it is known that the nuclei in the brainstem can show the potential to easily change their response characteristics. The discrepancy with the results in this thesis may for instance be due to the limited sensitivity of fMRI, but may also be related to the physiological mechanism behind the hemodynamic response. However, the overall conclusion seems justified that high processing levels play a dominant role in the realization of plastic changes in response behavior, at least in unilaterally deaf subjects.

7.2.4 Practical applications and clinical relevance

The use of MRI in a clinical environment, and in an audiology department in particular, is currently mostly limited to the imaging of purely anatomical features (Casselmann, 2002). For instance, MRI is very suitable for the detection of brain tumors (e.g., acoustic nerve schwannomas). Other emerging applications in the field of brain imaging in general are diffusion weighted imaging for the detection of atrophic brain regions, diffusion tensor imaging to trace fiber tracts, and perfusion imaging to monitor the treatment of tumors for example. The role of fMRI in a clinical context is yet limited, although it is increasingly important for mapping brain functions in individuals (especially for demarcating speech related regions) prior to surgical treatment of brain tumors. As more detailed knowledge is obtained about the various functions of the brain, more clinical applications of fMRI may emerge (Hennig et al., 2003). In the field of auditory processing a number of potential applications can be thought of for future clinical use.

Tinnitus is a disorder which is likely to have both peripheral and central causes (Eggermont et al., 2004). The understanding of the mechanism behind tinnitus is severely limited by the restricted knowledge about the central functioning of the auditory system in general (Weisz et al., 2005). It has been shown that the presence of tinnitus does lead to changes in the functioning of the central auditory system that can be detected with fMRI (Melcher et al., 2000). Also, it is thought that plasticity may play a major role in the development of tinnitus (Rauschecker, 1999), and fMRI is a technique that enables plastic functional changes to be imaged and localized in relative detail. Therefore, fMRI seems a potentially important tool for the understanding of tinnitus. Although much more research will be necessary, fMRI might serve as a diagnostic method to classify various types of tinnitus. Possibly multiple pathophysiological mechanisms can lead to the generation of a tinnitus percept, and fMRI might be able to distinguish

between these and pinpoint the location of dysfunctional areas in the brain. Alternatively, disorders that result in tinnitus may be followed by functional reorganization and adaptation processes in the brain that differ between individuals, and fMRI might be used to map such individual changes. Also, fMRI may help to guide and follow up on interventions to eliminate the tinnitus percept. Similarly, fMRI might be used in the future as a tool to diagnose other yet poorly understood hearing related disorders with a central component, like cortical deafness, word deafness, auditory agnosia, and auditory processing deficits related to e.g. specific language impairment or dyslexia, etcetera (Weismer et al., 2005).

Another field of interest for clinical applications of fMRI is the use of cochlear implants (CIs) or related devices and hearing aids. CIs function by direct galvanic stimulation of the cochlear nerve to overcome hearing loss or deafness due to loss of hair cells. However, this method relies on the intactness and excitability of the cochlear nerve and will not lead to improved hearing if the cochlear nerve itself has ceased to function. At higher levels, speech-related areas must have been developed and must have remained intact in order to achieve an improved recognition of speech. Although the measurement of brainstem auditory evoked potentials (BAEPs) can ascertain the intactness of the cochlear nerve, fMRI has the advantage that all levels of auditory processing are in principle accessible for evaluation. Thus, fMRI may form a valuable tool for clinical selection of patients that should be considered for CI implantation or related therapies. Moreover, fMRI may be used in the follow-up of such patients to monitor the reaction of the brain to the newly acquired input (Kral et al., 2001; Seghier et al., 2005). Although fMRI experiments on patients with CIs are still difficult, in the future MRI-compatible devices are likely to become available. Finally, fMRI already plays a role in research towards the understanding of speech-processing in the brain, which is of importance for the design of efficient encoding algorithms in hearing aids in the first place.

Nevertheless, given that fMRI is a relatively expensive, time-consuming and labor-intensive technique, it seems unlikely that fMRI in its current form will become a routinely applied technique in the near future, obviously with the exception of a number of cases. Nevertheless, fMRI will remain one of the most essential tools for brain research purposes, certainly with all the new technological and methodological developments that have only started to emerge. Therefore, fMRI will probably play an important role in the development of new techniques for clinical application in the near future.

ANNEX

Chapter

8

References
Curriculum Vitae
Financial Support
Postface

§8.1 References

- [1] Aguirre, G. K., E. Zarahn and M. D'Esposito (1998). "The variability of human, BOLD hemodynamic responses." *Neuroimage* 8(4): 360-9.
- [2] Alho, K., V. A. Vorobyev, S. V. Medvedev, S. V. Pakhomov, M. S. Roudas, M. Tervaniemi, T. van Zuijen and R. Naatanen (2003). "Hemispheric lateralization of cerebral blood-flow changes during selective listening to dichotically presented continuous speech." *Brain research. Cognitive brain research* 17(2): 201-11.
- [3] Alwatban, A. Z., C. N. Ludman, S. M. Mason, G. M. O'Donoghue, A. M. Peters and P. G. Morris (2002). "A method for the direct electrical stimulation of the auditory system in deaf subjects: a functional magnetic resonance imaging study." *Journal of magnetic resonance imaging* 16(1): 6-12.
- [4] Amaro, E., Jr., S. C. Williams, S. S. Shergill, C. H. Fu, M. MacSweeney, M. M. Picchioni, M. J. Brammer and P. K. McGuire (2002). "Acoustic noise and functional magnetic resonance imaging: Current strategies and future prospects." *Journal of magnetic resonance imaging* 16(5): 497-510.
- [5] Arnott, S. R., M. A. Binns, C. L. Grady and C. Alain (2004). "Assessing the auditory dual-pathway model in humans." *Neuroimage* 22(1): 401-8.
- [6] Arthurs, O. J. and S. Boniface (2002). "How well do we understand the neural origins of the fMRI BOLD signal?" *Trends in Neurosciences* 25(1): 27-31.
- [7] Backes, W. H. and P. van Dijk (2002). "Simultaneous sampling of event-related BOLD responses in auditory cortex and brainstem." *Magnetic resonance in medicine* 47(1): 90-6.
- [8] Bandettini, P. A., A. Jesmanowicz, J. Van Kylen, R. M. Birn and J. S. Hyde (1998). "Functional MRI of brain activation induced by scanner acoustic noise." *Magnetic resonance in medicine* 39(3): 410-6.
- [9] Behne, N., H. Scheich and A. Brechmann (2005). "Contralateral White Noise Selectively Changes Right Human Auditory Cortex Activity Caused by a FM-Direction Task." *Journal of Neurophysiology* 93(1): 414-23.
- [10] Belin, P., R. J. Zatorre, R. Hoge, A. C. Evans and B. Pike (1999). "Event-related fMRI of the auditory cortex." *Neuroimage* 10(4): 417-29.
- [11] Belin, P., R. J. Zatorre, P. Lafaille, P. Ahad and B. Pike (2000). "Voice-selective areas in human auditory cortex." *Nature* 403(6767): 309-12.
- [12] Belin, P., R. J. Zatorre and P. Ahad (2002). "Human temporal-lobe response to vocal sounds." *Brain research. Cognitive brain research* 13(1): 17-26.
- [13] Bilecen, D., K. Scheffler, N. Schmid, K. Tschopp and J. Seelig (1998). "Tonotopic organization of the human auditory cortex as detected by BOLD-fMRI." *Hearing research* 126(1-2): 19-27.
- [14] Bilecen, D., E. Seifritz, E. W. Radu, N. Schmid, S. Wetzel, R. Probst and K. Scheffler (2000). "Cortical reorganization after acute unilateral hearing loss traced by fMRI." *Neurology* 54(3): 765-7.
- [15] Bilecen, D., E. Seifritz, K. Scheffler, J. Henning and A. C. Schulte (2002). "Amplitopicity of the human auditory cortex: an fMRI study." *Neuroimage* 17(2): 710-8.
- [16] Binder, J. R., S. M. Rao, T. A. Hammeke, J. A. Frost, P. A. Bandettini and J. S. Hyde (1994a). "Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex." *Brain research. Cognitive brain research* 2(1): 31-8.
- [17] Binder, J. R., S. M. Rao, T. A. Hammeke, F. Z. Yetkin, A. Jesmanowicz, P. A. Bandettini, E. C. Wong, L. D. Estkowski, M. D. Goldstein, V. M. Houghton and et al. (1994b). "Functional magnetic resonance imaging of human auditory cortex." *Annals of neurology* 35(6): 662-72.

- [18] Binder, J. R., J. A. Frost, T. A. Hammeke, R. W. Cox, S. M. Rao and T. Prieto (1997). "Human brain language areas identified by functional magnetic resonance imaging." *Journal of neuroscience, The* 17(1): 353-62.
- [19] Binder, J. R., J. A. Frost, T. A. Hammeke, P. S. Bellgowan, J. A. Springer, J. N. Kaufman and E. T. Possing (2000). "Human temporal lobe activation by speech and nonspeech sounds." *Cerebral Cortex* 10(5): 512-28.
- [20] Binder, J. R., E. Liebenthal, E. T. Possing, D. A. Medler and B. D. Ward (2004). "Neural correlates of sensory and decision processes in auditory object identification." *Nature Neuroscience* 7(3): 295-301.
- [21] Birn, R. M. and P. A. Bandettini (2005). "The effect of stimulus duty cycle and "off" duration on BOLD response linearity." *Neuroimage*.
- [22] Bohning, D. E., A. Shastri, M. P. Lomarev, J. P. Lorberbaum, Z. Nahas and M. S. George (2003). "BOLD-fMRI response vs. transcranial magnetic stimulation (TMS) pulse- train length: Testing for linearity." *Journal of magnetic resonance imaging* 17(3): 279-90.
- [23] Bonham, B. H., S. W. Cheung, B. Godey and C. E. Schreiner (2004). "Spatial organization of frequency response areas and rate/level functions in the developing AI." *Journal of neurophysiology* 91(2): 841-54.
- [24] Brechmann, A., F. Baumgart and H. Scheich (2002). "Sound-Level-Dependent Representation of Frequency Modulations in Human Auditory Cortex: A Low-Noise fMRI Study." *Journal of neurophysiology* 87(1): 423-33.
- [25] Brechmann, A. and H. Scheich (2005). "Hemispheric shifts of sound representation in auditory cortex with conceptual listening." *Cerebral Cortex* 15(5): 578-87.
- [26] Brodmann, K. (1909). "Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues". Leipzig, Barth.
- [27] Brunetti, M., P. Belardinelli, M. Caulo, C. Del Gratta, S. Della Penna, A. Ferretti, G. Lucci, A. Moretti, V. Pizzella, A. Tartaro, K. Torquati, M. Olivetti Belardinelli and G. L. Romani (2005). "Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study." *Human Brain Mapping* 26(4): 251-61.
- [28] Buchel, C. and K. J. Friston (1997). "Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI." *Cerebral Cortex* 7(8): 768-78.
- [29] Buchel, C., A. P. Holmes, G. Rees and K. J. Friston (1998). "Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments." *Neuroimage* 8(2): 140-8.
- [30] Buchel, C. and K. Friston (2000). "Assessing interactions among neuronal systems using functional neuroimaging." *Neural networks* 13(8-9): 871-82.
- [31] Budd, T. W., D. A. Hall, M. S. Goncalves, M. A. Akeroyd, J. R. Foster, A. R. Palmer, K. Head and A. Q. Summerfield (2003). "Binaural specialisation in human auditory cortex: an fMRI investigation of interaural correlation sensitivity." *Neuroimage* 20(3): 1783-94.
- [32] Buracas, G. T. and G. M. Boynton (2002). "Efficient design of event-related fMRI experiments using M-sequences." *Neuroimage* 16(3 Pt 1): 801-13.
- [33] Buxton, R. B. (2002). "Introduction to functional magnetic resonance imaging". New York, Cambridge University Press.
- [34] Buxton, R. B., K. Uludag, D. J. Dubowitz and T. T. Liu (2004). "Modeling the hemodynamic response to brain activation." *Neuroimage* 23 Suppl 1: S220-33.
- [35] Calhoun, B. M. and C. E. Schreiner (1998). "Spectral envelope coding in cat primary auditory cortex: linear and non-linear effects of stimulus characteristics." *European journal of neuroscience, The* 10(3): 926-40.
- [36] Cansino, S., A. Ducorps and R. Ragot (2003). "Tonotopic cortical representation of periodic complex sounds." *Human Brain Mapping* 20(2): 71-81.

- [37] Casselman, J. W. (2002). "Diagnostic imaging in clinical neuro-otology." *Current Opinion in Neurology* 15(1): 23-30.
- [38] Chambers, J., M. A. Akeroyd, A. Q. Summerfield and A. R. Palmer (2001). "Active control of the volume acquisition noise in functional magnetic resonance imaging: method and psychoacoustical evaluation." *Journal of the Acoustical Society of America*, The 110(6): 3041-54.
- [39] Chi, T., Y. Gao, M. C. Guyton, P. Ru and S. Shamma (1999). "Spectro-temporal modulation transfer functions and speech intelligibility." *Journal of the Acoustical Society of America*, The 106(5): 2719-32.
- [40] Cho, Z. H., S. H. Park, J. H. Kim, S. C. Chung, S. T. Chung, J. Y. Chung, C. W. Moon, J. H. Yi, C. H. Sin and E. K. Wong (1997). "Analysis of acoustic noise in MRI." *Magnetic Resonance Imaging* 15(7): 815-22.
- [41] Cohen, M. S. (1997). "Parametric analysis of fMRI data using linear systems methods." *Neuroimage* 6(2): 93-103.
- [42] Crinion, J. T., M. A. Lambon-Ralph, E. A. Warburton, D. Howard and R. J. Wise (2003). "Temporal lobe regions engaged during normal speech comprehension." *Brain* 126(Pt 5): 1193-201.
- [43] Dale, A. M. (1999). "Optimal experimental design for event-related fMRI." *Human Brain Mapping* 8(2-3): 109-14.
- [44] de Zwart, J. A., P. van Gelderen, P. Kellman and J. H. Duyn (2002). "Reduction of gradient acoustic noise in MRI using SENSE-EPI." *Neuroimage* 16(4): 1151-5.
- [45] deCharms, R. C., D. T. Blake and M. M. Merzenich (1998). "Optimizing sound features for cortical neurons." *Science* 280(5368): 1439-43.
- [46] Demonet, J. F., F. Chollet, S. Ramsay, D. Cardebat, J. L. Nespoulous, R. Wise, A. Rascol and R. Frackowiak (1992). "The anatomy of phonological and semantic processing in normal subjects." *Brain* 115(Pt 6): 1753-68.
- [47] Depireux, D. A., J. Z. Simon, D. J. Klein and S. A. Shamma (2001). "Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex." *Journal of neurophysiology* 85(3): 1220-34.
- [48] Di Salle, F., E. Formisano, E. Seifritz, D. E. Linden, K. Scheffler, C. Saulino, G. Tedeschi, F. E. Zanella, A. Pepino, R. Goebel and E. Marciano (2001). "Functional fields in human auditory cortex revealed by time-resolved fMRI without interference of EPI noise." *Neuroimage* 13(2): 328-38.
- [49] Di Salle, F., F. Esposito, T. Scarabino, E. Formisano, E. Marciano, C. Saulino, S. Cirillo, R. Elefante, K. Scheffler and E. Seifritz (2003a). "fMRI of the auditory system: understanding the neural basis of auditory gestalt." *Journal of magnetic resonance imaging* 21(10): 1213-24.
- [50] Di Salle, F., F. Esposito, A. Elefante, T. Scarabino, A. Volpicelli, S. Cirillo, R. Elefante and E. Seifritz (2003b). "High field functional MRI." *European Journal of Radiology* 48(2): 138-45.
- [51] Eden, G. F., J. E. Joseph, H. E. Brown, C. P. Brown and T. A. Zeffiro (1999). "Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: the behavior interleaved gradients technique." *Magnetic resonance in medicine* 41(1): 13-20.
- [52] Edmister, W. B., T. M. Talavage, P. J. Ledden and R. M. Weisskoff (1999). "Improved auditory cortex imaging using clustered volume acquisitions." *Human brain mapping* 7(2): 89-97.
- [53] Eggermont, J. J. (2001). "Between sound and perception: reviewing the search for a neural code." *Hearing Research* 157(1-2): 1-42.
- [54] Eggermont, J. J. and L. E. Roberts (2004). "The neuroscience of tinnitus." *Trends in neurosciences* 27(11): 676-82.
- [55] Ehret, G. (1997). "The auditory cortex." *Journal of comparative physiology. Sensory, neural, and behavioral physiology* 181(6): 547-57.

-
- [56] Ehret, G. and R. Romand (1997). "The central auditory system". New York, Oxford University Press, Inc.
 - [57] Elliott, M. R., R. W. Bowtell and P. G. Morris (1999). "The effect of scanner sound in visual, motor, and auditory functional MRI." *Magnetic Resonance in Medicine* 41(6): 1230-5.
 - [58] Engelien, A., Y. Yang, W. Engelien, J. Zonana, E. Stern and D. Silbersweig (2002). "Physiological Mapping of Human Auditory Cortices with a Silent Event-Related fMRI Technique." *Neuroimage* 16(4): 944.
 - [59] Fitzpatrick, K. A. and T. J. Imig (1980). "Auditory cortico-cortical connections in the owl monkey." *Journal of comparative neurology, The* 192(3): 589-610.
 - [60] Formisano, E., D. S. Kim, F. Di Salle, P. F. van de Moortele, K. Ugurbil and R. Goebel (2003). "Mirror-symmetric tonotopic maps in human primary auditory cortex." *Neuron* 40(4): 859-69.
 - [61] Friman, O., J. Cedefamn, P. Lundberg, M. Borga and H. Knutsson (2001). "Detection of neural activity in functional MRI using canonical correlation analysis." *Magnetic resonance in medicine* 45(2): 323-30.
 - [62] Friston, K. (2002). "Beyond phrenology: what can neuroimaging tell us about distributed circuitry?" *Annual Reviews of Neuroscience* 25: 221-50.
 - [63] Friston, K. J. (1994). "Functional and effective connectivity in neuroimaging: a synthesis." *Human brain mapping* 2(1): 56-78.
 - [64] Friston, K. J., A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith and R. S. J. Frackowiak (1995). "Statistical Parametric Maps in Functional Imaging: A General Linear Approach." *Human brain mapping* 2: 189-210.
 - [65] Friston, K. J., C. Buchel, G. R. Fink, J. Morris, E. Rolls and R. J. Dolan (1997). "Psychophysiological and modulatory interactions in neuroimaging." *Neuroimage* 6(3): 218-29.
 - [66] Friston, K. J., P. Fletcher, O. Josephs, A. Holmes, M. D. Rugg and R. Turner (1998a). "Event-related fMRI: characterizing differential responses." *Neuroimage* 7(1): 30-40.
 - [67] Friston, K. J., O. Josephs, G. Rees and R. Turner (1998b). "Nonlinear event-related responses in fMRI." *Magnetic resonance in medicine* 39(1): 41-52.
 - [68] Friston, K. J., A. P. Holmes, C. J. Price, C. Buchel and K. J. Worsley (1999). "Multisubject fMRI studies and conjunction analyses." *Neuroimage* 10(4): 385-96.
 - [69] Friston, K. J., L. Harrison and W. Penny (2003). "Dynamic causal modelling." *Neuroimage* 19(4): 1273-302.
 - [70] Fujioka, T., B. Ross, H. Okamoto, Y. Takeshima, R. Kakigi and C. Pantev (2003). "Tonotopic representation of missing fundamental complex sounds in the human auditory cortex." *European journal of neuroscience, The* 18(2): 432-40.
 - [71] Gabriele, M. L., J. K. Brunso-Bechtold and C. K. Henkel (2000). "Plasticity in the development of afferent patterns in the inferior colliculus of the rat after unilateral cochlear ablation." *Journal of neuroscience, The* 20(18): 6939-49.
 - [72] Gage, N. M., T. P. Roberts and G. Hickok (2002). "Hemispheric asymmetries in auditory evoked neuromagnetic fields in response to place of articulation contrasts." *Brain research. Cognitive brain research* 14(2): 303-6.
 - [73] Galaburda, A. and F. Sanides (1980). "Cytoarchitectonic organization of the human auditory cortex." *Journal of comparative neurology, The* 190(3): 597-610.
 - [74] Gasiorowicz, S. (1974). "Quantum physics". New York, John Wiley & Sons, Inc.
 - [75] Giraud, A. L., C. Lorenzi, J. Ashburner, J. Wable, I. Johnsrude, R. Frackowiak and A. Kleinschmidt (2000). "Representation of the temporal envelope of sounds in the human brain." *Journal of neurophysiology* 84(3): 1588-98.
 - [76] Glover, G. H. (1999). "Deconvolution of impulse response in event-related BOLD fMRI." *Neuroimage* 9(4): 416-29.
-

- [77] Goncalves, M. S., D. A. Hall, I. S. Johnsrude and M. P. Haggard (2001). "Can meaningful effective connectivities be obtained between auditory cortical regions?" *Neuroimage* **14**(6): 1353-60.
- [78] Griffiths, T. D. and J. D. Warren (2004). "What is an auditory object?" *Nature Reviews Neuroscience* **5**(11): 887-92.
- [79] Guimaraes, A. R., J. R. Melcher, T. M. Talavage, J. R. Baker, P. Ledden, B. R. Rosen, N. Y. Kiang, B. C. Fullerton and R. M. Weisskoff (1998). "Imaging subcortical auditory activity in humans." *Human brain mapping* **6**(1): 33-41.
- [80] Haacke, E. M., R. W. Brown, M. R. Thompson and R. Venkatesan (1999). "Magnetic Resonance Imaging: Physical Principles and Sequence Design". New York, John Wiley & Sons, Inc.
- [81] Hall, D. A., M. P. Haggard, M. A. Akeroyd, A. R. Palmer, A. Q. Summerfield, M. R. Elliott, E. M. Gurney and R. W. Bowtell (1999). "'Sparse' temporal sampling in auditory fMRI." *Human brain mapping* **7**(3): 213-23.
- [82] Hall, D. A., M. P. Haggard, M. A. Akeroyd, A. Q. Summerfield, A. R. Palmer, M. R. Elliott and R. W. Bowtell (2000). "Modulation and task effects in auditory processing measured using fMRI." *Human brain mapping* **10**(3): 107-19.
- [83] Hall, D. A., M. P. Haggard, A. Q. Summerfield, M. A. Akeroyd, A. R. Palmer and R. W. Bowtell (2001). "Functional magnetic resonance imaging measurements of sound-level encoding in the absence of background scanner noise." *Journal of the Acoustical Society of America*, **The** **109**(4): 1559-70.
- [84] Hanaway, J., T. A. Woolsey, M. H. Gado and M. P. Roberts (1998). "The brain atlas, a visual guide to the human central nervous system." Bethesda, Maryland, Fitzgerald Science Press, Inc.
- [85] Harms, M. P., J. J. Guinan, Jr., I. S. Sigalovsky and J. R. Melcher (2005). "Short-term sound temporal envelope characteristics determine multisecond time patterns of activity in human auditory cortex as shown by fMRI." *Journal of Neurophysiology* **93**(1): 210-22.
- [86] Hart, H., A. Palmer and D. Hall (2002). "Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex." *Hearing research* **171**(1-2): 177.
- [87] Hart, H. C., A. R. Palmer and D. A. Hall (2003a). "Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex." *Cerebral cortex* **13**(7): 773-81.
- [88] Hart, H. C., D. A. Hall and A. R. Palmer (2003b). "The sound-level-dependent growth in the extent of fMRI activation in Heschl's gyrus is different for low- and high-frequency tones." *Hearing research* **179**(1-2): 104-12.
- [89] Hawley, M. L., J. R. Melcher and B. C. Fullerton (2005). "Effects of sound bandwidth on fMRI activation in human auditory brainstem nuclei." *Hearing Research* **204**(1-2): 101-10.
- [90] Heinz, M. G. and E. D. Young (2004). "Response growth with sound level in auditory-nerve fibers after noise-induced hearing loss." *Journal of neurophysiology* **91**(2): 784-95.
- [91] Hennel, F. (2001). "Fast spin echo and fast gradient echo MRI with low acoustic noise." *Journal of Magnetic Resonance Imaging* **13**(6): 960-6.
- [92] Hennig, J., O. Speck, M. A. Koch and C. Weiller (2003). "Functional magnetic resonance imaging: a review of methodological aspects and clinical applications." *Journal of magnetic resonance imaging* **18**(1): 1-15.
- [93] Hesselmann, V., C. Wedekind, H. Kugel, O. Schulte, B. Krug, N. Klug and K. J. Lackner (2001). "Functional magnetic resonance imaging of human pontine auditory pathway." *Hearing research* **158**(1-2): 160-4.
- [94] Hickok, G. and D. Poeppel (2000). "Towards a functional neuroanatomy of speech perception." *Trends in Cognitive Science* **4**(4): 131-138.
- [95] Hochberg, Y. (1988). "A sharper Bonferroni procedure for multiple tests of significance." *Biometrika* **75**: 800-802.

- [96] Horwitz, B., M. A. Tagamets and A. R. McIntosh (1999). "Neural modeling, functional brain imaging, and cognition." *Trends in Cognitive Science* 3(3): 91-98.
- [97] Horwitz, B. (2003). "The elusive concept of brain connectivity." *Neuroimage* 19(2 Pt 1): 466-70.
- [98] Howard, M. A., 3rd, I. O. Volkov, P. J. Abbas, H. Damasio, M. C. Ollendieck and M. A. Granner (1996). "A chronic microelectrode investigation of the tonotopic organization of human auditory cortex." *Brain Research* 724(2): 260-4.
- [99] Hox, J. J. and T. M. Bechger (1998). "An introduction to structural equation modeling." *Family Science Review* 11: 354-373.
- [100] Hyvarinen, A. and E. Oja (2000). "Independent component analysis: algorithms and applications." *Neural Networks* 13(4-5): 411-30.
- [101] Illing, R. B., K. S. Kraus and S. A. Michler (2000). "Plasticity of the superior olivary complex." *Microscopy Research and Technique* 51(4): 364-81.
- [102] Imig, T. J. and H. O. Adrian (1977). "Binaural columns in the primary field (A1) of cat auditory cortex." *Brain Research* 138(2): 241-57.
- [103] Inagaki, M., Y. Tomita, S. Takashima, K. Ohtani, G. Andoh and K. Takeshita (1987). "Functional and morphometrical maturation of the brainstem auditory pathway." *Brain and Development* 9(6): 597-601.
- [104] Inan, S., T. Mitchell, A. Song, J. Bizzell and A. Belger (2004). "Hemodynamic correlates of stimulus repetition in the visual and auditory cortices: an fMRI study." *Neuroimage* 21(3): 886-93.
- [105] Inui, K., H. Okamoto, K. Miki, A. Gunji and R. Kakigi (2005). "Serial and Parallel Processing in the Human Auditory Cortex: A Magnetoencephalographic Study." *Cerebral Cortex*.
- [106] Irvine, D. R., R. Rajan and S. Smith (2003). "Effects of restricted cochlear lesions in adult cats on the frequency organization of the inferior colliculus." *Journal of comparative neurology*, *The* 467(3): 354-74.
- [107] Jäncke, L., N. J. Shah, S. Posse, M. Grosse Ryken and H. W. Müller Gartner (1998). "Intensity coding of auditory stimuli: an fMRI study." *Neuropsychologia* 36(9): 875-83.
- [108] Jäncke, L., T. Buchanan, K. Lutz, K. Specht, S. Mirzazade and N. J. Shah (1999). "The time course of the BOLD response in the human auditory cortex to acoustic stimuli of different duration." *Brain research. Cognitive brain research* 8(2): 117-24.
- [109] Jäncke, L., T. Wustenberg, K. Schulze and H. J. Heinze (2002a). "Asymmetric hemodynamic responses of the human auditory cortex to monaural and binaural stimulation." *Hearing research* 170(1-2): 166-78.
- [110] Jäncke, L., T. Wustenberg, H. Scheich and H. J. Heinze (2002b). "Phonetic perception and the temporal cortex." *Neuroimage* 15(4): 733-46.
- [111] Jezzard, P., P. M. Matthews and S. M. Smith (2001). "Functional MRI". New York, Oxford University Press, Inc.
- [112] Johansson, M. S. and S. D. Arlinger (2003). "Prevalence of hearing impairment in a population in Sweden." *International journal of audiology* 42(1): 18-28.
- [113] Josse, G., B. Mazoyer, F. Crivello and N. Tzourio-Mazoyer (2003). "Left planum temporale: an anatomical marker of left hemispheric specialization for language comprehension." *Brain research. Cognitive brain research* 18(1): 1-14.
- [114] Kaas, J. H., T. A. Hackett and M. J. Tramo (1999). "Auditory processing in primate cerebral cortex." *Current opinion in neurobiology* 9(2): 164-70.
- [115] Kaas, J. H. and T. A. Hackett (2000). "Subdivisions of auditory cortex and processing streams in primates." *Proceedings of the National Academy of Sciences of the United States of America* 97(22): 11793-9.
- [116] Kamke, M. R., M. Brown and D. R. Irvine (2003). "Plasticity in the tonotopic organization of the medial geniculate body in adult cats following restricted unilateral cochlear lesions." *Journal of comparative neurology*, *The* 459(4): 355-67.

- [117] Kaneko, K., N. Fujiki and R. Hari (2003). "Binaural interaction in the human auditory cortex revealed by neuromagnetic frequency tagging: no effect of stimulus intensity." *Hearing research* 183(1-2): 1-6.
- [118] Karlsmose, B., T. Lauritzen and A. Parving (1999). "Prevalence of hearing impairment and subjective hearing problems in a rural Danish population aged 31-50 years." *British journal of audiology* 33(6): 395-402.
- [119] Katz, J. (2001). "Handbook of Clinical Audiology". Philadelphia, Lippincott Williams and Wilkins.
- [120] Kelly, J. B., P. W. Judge and D. P. Phillips (1986). "Representation of the cochlea in primary auditory cortex of the ferret (*Mustela putorius*).\" *Hearing Research* 24(2): 111-5.
- [121] Klein, D. J., D. A. Depireux, J. Z. Simon and S. A. Shamma (2000). "Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design." *Journal of computational neuroscience* 9(1): 85-111.
- [122] Komisaruk, B. R., K. M. Mosier, W. C. Liu, C. Criminale, L. Zaborszky, B. Whipple and A. Kalnin (2002). "Functional localization of brainstem and cervical spinal cord nuclei in humans with fMRI." *American Journal of Neuroradiology* 23(4): 609-17.
- [123] Konig, R., P. Heil, E. Budinger and H. Scheich (2005). "The auditory cortex; a synthesis of human and animal research". London, Lawrence Erlbaum Associates, Inc.
- [124] Kowalski, N., H. Versnel and S. A. Shamma (1995). "Comparison of responses in the anterior and primary auditory fields of the ferret cortex." *Journal of Neurophysiology* 73(4): 1513-23.
- [125] Kowalski, N., D. A. Depireux and S. A. Shamma (1996). "Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra." *Journal of neurophysiology* 76(5): 3503-23.
- [126] Kral, A., R. Hartmann, J. Tillein, S. Heid and R. Klinke (2001). "Delayed maturation and sensitive periods in the auditory cortex." *Audiology and Neurootology* 6(6): 346-62.
- [127] Krumbholz, K., M. Schonwiesner, R. Rubsamen, K. Zilles, G. R. Fink and D. Y. von Cramon (2005a). "Hierarchical processing of sound location and motion in the human brainstem and planum temporale." *European Journal of Neuroscience* 21(1): 230-8.
- [128] Krumbholz, K., M. Schonwiesner, D. Y. von Cramon, R. Rubsamen, N. J. Shah, K. Zilles and G. R. Fink (2005b). "Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe." *Cerebral Cortex* 15(3): 317-24.
- [129] Langner, G., M. Sams, P. Heil and H. Schulze (1997). "Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography." *Journal of comparative physiology. Sensory, neural, and behavioral physiology* 181(6): 665-76.
- [130] Lauter, J. L., P. Herscovitch, C. Formby and M. E. Raichle (1985). "Tonotopic organization in human auditory cortex revealed by positron emission tomography." *Hearing Research* 20(3): 199-205.
- [131] Le, T. H., S. Patel and T. P. Roberts (2001). "Functional MRI of human auditory cortex using block and event-related designs." *Magnetic resonance in medicine* 45(2): 254-60.
- [132] Lee, C. C., K. Imaizumi, C. E. Schreiner and J. A. Winer (2004). "Concurrent tonotopic processing streams in auditory cortex." *Cerebral Cortex* 14(4): 441-51.
- [133] Leonard, C. M., C. Puranik, J. M. Kuldau and L. J. Lombardino (1998). "Normal variation in the frequency and location of human auditory cortex landmarks. Heschl's gyrus: where is it?" *Cerebral cortex* 8(5): 397-406.
- [134] Lessard, N., F. Lepore, J. Villemagne and M. Lassonde (2002). "Sound localization in callosal agenesis and early callosotomy subjects: brain reorganization and/or compensatory strategies." *Brain* 125(Pt 5): 1039-53.

- [135] Liang, L., T. Lu and X. Wang (2002). "Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates." *Journal of neurophysiology* 87(5): 2237-61.
- [136] Liebenthal, E., J. R. Binder, S. M. Spitzer, E. T. Possing and D. A. Medler (2005). "Neural substrates of phonemic perception." *Cerebral Cortex* 15(10): 1621-31.
- [137] Liegeois-Chauvel, C., K. Giraud, J. M. Badier, P. Marquis and P. Chauvel (2001). "Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex." *Annals of the New York Academy of Sciences* 930: 117-32.
- [138] Liegeois-Chauvel, C., C. Lorenzi, A. Trebuchon, J. Regis and P. Chauvel (2004). "Temporal envelope processing in the human left and right auditory cortices." *Cerebral Cortex* 14(7): 731-40.
- [139] Linden, J. F., R. C. Liu, M. Sahani, C. E. Schreiner and M. M. Merzenich (2003). "Spectro-temporal structure of receptive fields in areas AI and AAF of mouse auditory cortex." *Journal of neurophysiology* 90(4): 2660-75.
- [140] Lipschutz, B., R. Kolinsky, P. Damhaut, D. Wikler and S. Goldman (2002). "Attention-dependent changes of activation and connectivity in dichotic listening." *Neuroimage* 17(2): 643-56.
- [141] Liu, T. T. (2004). "Efficiency, power, and entropy in event-related fMRI with multiple trial types. Part II: design of experiments." *Neuroimage* 21(1): 401-13.
- [142] Lockwood, A. H., R. J. Salvi, M. L. Coad, S. A. Arnold, D. S. Wack, B. W. Murphy and R. F. Burkard (1999). "The functional anatomy of the normal human auditory system: responses to 0.5 and 4.0 kHz tones at varied intensities." *Cerebral cortex* 9(1): 65-76.
- [143] Logothetis, N. K. and B. A. Wandell (2004a). "Interpreting the BOLD signal." *Annual Review of Physiology* 66: 735-69.
- [144] Logothetis, N. K. and J. Pfeuffer (2004b). "On the nature of the BOLD fMRI contrast mechanism." *Magnetic Resonance Imaging* 22(10): 1517-1531.
- [145] London, M. and M. Hausser (2005). "Dendritic computation." *Annual Reviews of Neuroscience* 28: 503-32.
- [146] Lu, H., X. Golay, J. J. Pekar and P. C. Van Zijl (2003). "Functional magnetic resonance imaging based on changes in vascular space occupancy." *Magnetic resonance in medicine* 50(2): 263-74.
- [147] Martin, F. N. (1996). "Introduction to audiology". Needham Heights, MA, A Viacom Company.
- [148] McAlpine, D., R. L. Martin, J. E. Mossop and D. R. Moore (1997). "Response properties of neurons in the inferior colliculus of the monaurally deafened ferret to acoustic stimulation of the intact ear." *Journal of Neurophysiology* 78(2): 767-79.
- [149] McDermott, H. J., M. Lech, M. S. Kornblum and D. R. Irvine (1998). "Loudness perception and frequency discrimination in subjects with steeply sloping hearing loss: possible correlates of neural plasticity." *Journal of the Acoustical Society of America, The* 104(4): 2314-25.
- [150] McIntosh, A. R. and N. J. Lobaugh (2004). "Partial least squares analysis of neuroimaging data: applications and advances." *Neuroimage* 23 Suppl 1: S250-63.
- [151] Melcher, J. R., I. S. Sigalovsky, J. J. Guinan, Jr. and R. A. Levine (2000). "Lateralized tinnitus studied with functional magnetic resonance imaging: abnormal inferior colliculus activation." *Journal of Neurophysiology* 83(2): 1058-72.
- [152] Mendelson, J. R., C. E. Schreiner, M. L. Sutter and K. L. Grasse (1993). "Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps." *Experimental Brain Research* 94(1): 65-87.
- [153] Merzenich, M. M. and J. F. Brugge (1973). "Representation of the cochlear partition of the superior temporal plane of the macaque monkey." *Brain Research* 50(2): 275-96.

- [154] Millen, S. J., V. M. Haughton and Z. Yetkin (1995). "Functional magnetic resonance imaging of the central auditory pathway following speech and pure-tone stimuli." *Laryngoscope, The* **105**(12 Pt 1): 1305-10.
- [155] Miller, L. M., M. A. Escabi, H. L. Read and C. E. Schreiner (2002). "Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex." *Journal of neurophysiology* **87**(1): 516-27.
- [156] Moelker, A. and P. M. Pattynama (2003). "Acoustic noise concerns in functional magnetic resonance imaging." *Human Brain Mapping* **20**(3): 123-41.
- [157] Moelker, A., R. A. Maas, M. W. Vogel, M. Ouhlous and P. M. Pattynama (2005). "Importance of bone-conducted sound transmission on patient hearing in the MR scanner." *Journal of Magnetic Resonance Imaging* **22**(1): 163-9.
- [158] Mohr, C. M., W. M. King, A. J. Freeman, R. W. Briggs and C. M. Leonard (1999). "Influence of speech stimuli intensity on the activation of auditory cortex investigated with functional magnetic resonance imaging." *Journal of the Acoustical Society of America, The* **105**(5): 2738-45.
- [159] Moore, B. C. (1996). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids." *Ear and Hearing* **17**(2): 133-61.
- [160] Moore, B. C. (2003). "An introduction to the psychology of hearing". London, Academic Press.
- [161] Moore, B. C. and B. R. Glasberg (2004). "A revised model of loudness perception applied to cochlear hearing loss." *Hearing research* **188**(1-2): 70-88.
- [162] Moore, D. R. (1991). "Anatomy and physiology of binaural hearing." *Audiology* **30**(3): 125-34.
- [163] Morita, T., Y. Naito, T. Nagamine, N. Fujiki, H. Shibasaki and J. Ito (2003). "Enhanced activation of the auditory cortex in patients with inner-ear hearing impairment: a magnetoencephalographic study." *Clinical Neurophysiology* **114**(5): 851-9.
- [164] Morosan, P., J. Rademacher, A. Schleicher, K. Amunts, T. Schormann and K. Zilles (2001). "Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system." *Neuroimage* **13**(4): 684-701.
- [165] Mossop, J. E., M. J. Wilson, D. M. Caspary and D. R. Moore (2000). "Down-regulation of inhibition following unilateral deafening." *Hearing Research* **147**(1-2): 183-7.
- [166] Muhlcnickcl, W., T. Elbert, E. Taub and H. Flor (1998). "Reorganization of auditory cortex in tinnitus." *Proceedings of the National Academy of Sciences of the United States of America* **95**(17): 10340-3.
- [167] Mulert, C., L. Jager, S. Propp, S. Karch, S. Stormann, O. Pogarell, H. J. Moller, G. Juckel and U. Hegerl (2005). "Sound level dependence of the primary auditory cortex: Simultaneous measurement with 61-channel EEG and fMRI." *Neuroimage*.
- [168] Nair, D. G. (2005). "About being BOLD." *Brain research. Brain research reviews* **50**(2): 229-43.
- [169] Nelken, I. and H. Versnel (2000). "Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex." *European journal of neuroscience, The* **12**(2): 549-62.
- [170] Nelken, I. (2004). "Processing of complex stimuli and natural scenes in the auditory cortex." *Current Opinion in Neurobiology* **14**(4): 474-80.
- [171] Nichols, T. and S. Hayasaka (2003). "Controlling the familywise error rate in functional neuroimaging: a comparative review." *Statistical Methods in Medical Research* **12**(5): 419-46.
- [172] Nieuwenhuys, R. (1984). "Anatomy of the auditory pathways, with emphasis on the brain stem." *Advances in Otorhinolaryngology* **34**: 25-38.
- [173] Norman, G. R. and D. L. Streiner (2000). "Biostatistics, the bare essentials". Hamilton, Ontario, B. C. Becker, Inc.

-
- [174] Novitski, N., I. Anourova, S. Martinkauppi, H. J. Aronen, R. Naatanen and S. Carlson (2003). "Effects of noise from functional magnetic resonance imaging on auditory event-related potentials in working memory task." *Neuroimage* **20**(2): 1320-8.
- [175] Oldfield, R. C. (1971). "The assessment and analysis of handedness: the Edinburgh inventory." *Neuropsychologia* **9**(1): 97-113.
- [176] Oxenham, A. J. and S. P. Bacon (2003). "Cochlear compression: perceptual measures and implications for normal and impaired hearing." *Ear & Hearing* **24**(5): 352-66.
- [177] Palomaki, K. J., H. Tiitinen, V. Makinen, P. May and P. Alku (2002). "Cortical processing of speech sounds and their analogues in a spatial auditory environment." *Brain research. Cognitive brain research* **14**(2): 294-9.
- [178] Pantev, C., B. Lutkenhoner, M. Hoke and K. Lehnertz (1986). "Comparison between simultaneously recorded auditory-evoked magnetic fields and potentials elicited by ipsilateral, contralateral and binaural tone burst stimulation." *Audiology* **25**(1): 54-61.
- [179] Pantev, C., L. E. Roberts, T. Elbert, B. Ross and C. Wienbruch (1996). "Tonotopic organization of the sources of human auditory steady-state responses." *Hearing Research* **101**(1-2): 62-74.
- [180] Parkes, L. M., J. V. Schwarzbach, A. A. Bouts, R. H. Deckers, P. Pullens, C. M. Kerskens and D. G. Norris (2005). "Quantifying the spatial resolution of the gradient echo and spin echo BOLD response at 3 Tesla." *Magnetic Resonance in Medicine* **54**(6): 1465-72.
- [181] Penny, W. D., K. E. Stephan, A. Mechelli and K. J. Friston (2004). "Comparing dynamic causal models." *Neuroimage* **22**(3): 1157-72.
- [182] Pfeuffer, J., J. C. McCullough, P. F. Van de Moortele, K. Ugurbil and X. Hu (2003). "Spatial dependence of the nonlinear BOLD response at short stimulus duration." *Neuroimage* **18**(4): 990-1000.
- [183] Pickles, J. O. (1988). "An introduction to the physiology of hearing." London, Academic.
- [184] Ponton, C. W., J. P. Vasama, K. Tremblay, D. Khosla, B. Kwong and M. Don (2001). "Plasticity in the adult human central auditory system: evidence from late-onset profound unilateral deafness." *Hearing Research* **154**(1-2): 32-44.
- [185] Popelar, J., J. P. Erre, J. M. Aran and Y. Cazals (1994). "Plastic changes in ipsi-contralateral differences of auditory cortex and inferior colliculus evoked potentials after injury to one ear in the adult guinea pig." *Hearing Research* **72**(1-2): 125-34.
- [186] Purves, D., G. J. Augustine, D. Fitzpatrick, L. C. Katz, A. S. LaMantia and J. O. McNamara (1997). "Neuroscience". Sunderland, MA, Sinauer Associates, Inc.
- [187] Rademacher, J., P. Morosan, T. Schormann, A. Schleicher, C. Werner, H. J. Freund and K. Zilles (2001). "Probabilistic mapping and volume measurement of human primary auditory cortex." *Neuroimage* **13**(4): 669-83.
- [188] Ramnani, N., T. E. Behrens, W. Penny and P. M. Matthews (2004). "New approaches for exploring anatomical and functional connectivity in the human brain." *Biological Psychiatry* **56**(9): 613-9.
- [189] Rauschecker, J. P., B. Tian and M. Hauser (1995). "Processing of complex sounds in the macaque nonprimary auditory cortex." *Science* **268**(5207): 111-4.
- [190] Rauschecker, J. P. (1999). "Auditory cortical plasticity: a comparison with other sensory systems." *Trends in neurosciences* **22**(2): 74-80.
- [191] Rauschecker, J. P. and B. Tian (2000). "Mechanisms and streams for processing of "what" and "where" in auditory cortex." *Proceedings of the National Academy of Sciences of the United States of America* **97**(22): 11800-6.
- [192] Ravicz, M. E. and J. R. Melcher (2001). "Isolating the auditory system from acoustic noise during functional magnetic resonance imaging: examination of noise conduction through the ear canal, head, and body." *Journal of the Acoustical Society of America*, *The* **109**(1): 216-31.
-

- [193] Reale, R. A. and T. J. Imig (1980). "Tonotopic organization in auditory cortex of the cat." *Journal of comparative neurology, The* **192**(2): 265-91.
- [194] Recanzone, G. H., D. C. Guard and M. L. Phan (2000). "Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey." *Journal of neurophysiology* **83**(4): 2315-31.
- [195] Rees, G., A. Howseman, O. Josephs, C. D. Frith, K. J. Friston, R. S. Frackowiak and R. Turner (1997). "Characterizing the relationship between BOLD contrast and regional cerebral blood flow measurements by varying the stimulus presentation rate." *Neuroimage* **6**(4): 270-8.
- [196] Robson, M. D., J. L. Dorosz and J. C. Gore (1998). "Measurements of the temporal fMRI response of the human auditory cortex to trains of tones." *Neuroimage* **7**(3): 185-98.
- [197] Rodrigues-Dageaff, C., G. Simm, Y. De Ribaupierre, A. Villa, F. De Ribaupierre and E. M. Rouiller (1989). "Functional organization of the ventral division of the medial geniculate body of the cat: evidence for a rostro-caudal gradient of response properties and cortical projections." *Hearing Research* **39**(1-2): 103-25.
- [198] Romani, G. L., S. J. Williamson and L. Kaufman (1982). "Tonotopic organization of the human auditory cortex." *Science* **216**(4552): 1339-40.
- [199] Rosburg, T., I. Kreitschmann-Andermahr, E. Emmerich, H. Nowak and H. Sauer (1998). "Hemispheric differences in frequency dependent dipole orientation of the human auditory evoked field component N100m." *Neuroscience Letters* **258**(2): 105-8.
- [200] Salamy, A., L. Eldredge and A. Wakeley (1985). "Maturation of contralateral brain-stem responses in preterm infants." *Electroencephalography and Clinical Neurophysiology* **62**(2): 117-23.
- [201] Scarff, C. J., J. C. Dort, J. J. Eggermont and B. G. Goodyear (2004). "The effect of MR scanner noise on auditory cortex activity using fMRI." *Human Brain Mapping* **22**(4): 341-9.
- [202] Scheffler, K., D. Bilecen, N. Schmid, K. Tschopp and J. Seelig (1998). "Auditory cortical responses in hearing subjects and unilateral deaf patients as detected by functional magnetic resonance imaging." *Cerebral Cortex* **8**(2): 156-63.
- [203] Scheich, H., F. Baumgart, B. Gaschler-Markefski, C. Tegeler, C. Tempelmann, H. J. Heinze, F. Schindler and D. Stiller (1998). "Functional magnetic resonance imaging of a human auditory cortex area involved in foreground-background decomposition." *European journal of neuroscience, The* **10**(2): 803-9.
- [204] Schmid, N., K. Tschopp, C. Schillinger, D. Bilecen, K. Scheffler and J. Seelig (1998). "[Visualization of central auditory processes with functional magnetic resonance tomography]." *Laryngorhinootologie* **77**(6): 328-31.
- [205] Schmithorst, V. J. and S. K. Holland (2004). "Event-related fMRI technique for auditory processing with hemodynamics unrelated to acoustic gradient noise." *Magnetic resonance in medicine* **51**(2): 399-402.
- [206] Schönwiesner, M., D. Y. von Cramon and R. Rubsamen (2002). "Is it tonotopy after all?" *Neuroimage* **17**(3): 1144-61.
- [207] Schreiner, C. E. and J. V. Urbas (1988). "Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields." *Hearing research* **32**(1): 49-63.
- [208] Schreiner, C. E. (1998). "Spatial distribution of responses to simple and complex sounds in the primary auditory cortex." *Audiology and Neuro-Otology* **3**(2-3): 104-22.
- [209] Scott, S. K., C. C. Blank, S. Rosen and R. J. Wise (2000). "Identification of a pathway for intelligible speech in the left temporal lobe." *Brain* **123**(12): 2400-6.
- [210] Seghier, M. L., C. Boex, F. Lazeyras, A. Sigrist and M. Pelizzzone (2005). "fMRI evidence for activation of multiple cortical regions in the primary auditory cortex of deaf subjects users of multichannel cochlear implants." *Cerebral Cortex* **15**(1): 40-8.

-
- [211] Seifritz, E., F. Esposito, F. Hennel, H. Mustovic, J. G. Neuhoﬀ, D. Bilecen, G. Tedeschi, K. Scheﬄer and F. Di Salle (2002). "Spatiotemporal pattern of neural processing in the human auditory cortex." *Science* **297**(5587): 1706-8.
 - [212] Shah, N. J., L. Jäncke, M. L. Grosse Ruyken and H. W. Muller Gartner (1999). "Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination." *Journal of magnetic resonance imaging* **9**(1): 19-25.
 - [213] Shah, N. J., S. Steinhoff, S. Mirzazade, O. Zafiris, M. L. Grosse-Ruyken, L. Jancke and K. Zilles (2000). "The effect of sequence repeat time on auditory cortex stimulation during phonetic discrimination." *Neuroimage* **12**(1): 100-8.
 - [214] Shamma, S. A., J. W. Fleshman, P. R. Wiser and H. Versnel (1993). "Organization of response areas in ferret primary auditory cortex." *Journal of neurophysiology* **69**(2): 367-83.
 - [215] Shapleske, J., S. L. Rossell, P. W. Woodruff and A. S. David (1999). "The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance." *Brain research. Brain research reviews* **29**(1): 26-49.
 - [216] Shtyrov, Y., T. Kujala, S. Palva, R. J. Ilmoniemi and R. Naatanen (2000). "Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres." *Neuroimage* **12**(6): 657-63.
 - [217] Sininger, Y. S. and B. Cone-Wesson (2004). "Asymmetric cochlear processing mimics hemispheric specialization." *Science* **305**(5690): 1581.
 - [218] Specht, K. and J. Reul (2003). "Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: an auditory rapid event-related fMRI-task." *Neuroimage* **20**(4): 1944-54.
 - [219] Stefanovic, B., J. M. Warnking and G. B. Pike (2004). "Hemodynamic and metabolic responses to neuronal inhibition." *Neuroimage* **22**(2): 771-8.
 - [220] Stiebler, I., R. Neulist, I. Fichtel and G. Ehret (1997). "The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation." *Journal of comparative physiology. Sensory, neural, and behavioral physiology* **181**(6): 559-71.
 - [221] Supin, A., V. V. Popov, O. N. Milekhina and M. B. Tarakanov (1999). "Ripple depth and density resolution of rippled noise." *Journal of the Acoustical Society of America*, **The** **106**(5): 2800-4.
 - [222] Suzuki, M., H. Kitano, T. Kitanishi, R. Itou, A. Shiino, Y. Nishida, Y. Yazawa, F. Ogawa and K. Kitajima (2002). "Cortical and subcortical activation with monaural monosyllabic stimulation by functional MRI." *Hearing research* **163**(1-2): 37-45.
 - [223] Taberner, A. M. and M. C. Liberman (2005). "Response properties of single auditory nerve fibers in the mouse." *Journal of Neurophysiology* **93**(1): 557-69.
 - [224] Talavage, T. M., W. B. Edmister, P. J. Ledden and R. M. Weisskoff (1999). "Quantitative assessment of auditory cortex responses induced by imager acoustic noise." *Human brain mapping* **7**(2): 79-88.
 - [225] Talavage, T. M., P. J. Ledden, R. R. Benson, B. R. Rosen and J. R. Melcher (2000). "Frequency-dependent responses exhibited by multiple regions in human auditory cortex." *Hearing research* **150**(1-2): 225-44.
 - [226] Talavage, T. M. and W. B. Edmister (2004a). "Nonlinearity of fMRI responses in human auditory cortex." *Human Brain Mapping* **22**(3): 216-28.
 - [227] Talavage, T. M., M. I. Sereno, J. R. Melcher, P. J. Ledden, B. R. Rosen and A. M. Dale (2004b). "Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity." *Journal of neurophysiology* **91**(3): 1282-96.
 - [228] Tervaniemi, M. and K. Hugdahl (2003). "Lateralization of auditory-cortex functions." *Brain research. Brain research reviews* **43**(3): 231-46.
 - [229] Thivard, L., P. Belin, M. Zilbovicius, J. B. Poline and Y. Samson (2000). "A cortical region sensitive to auditory spectral motion." *Neuroreport* **11**(13): 2969-72.
-

- [230] Tian, B. and J. P. Rauschecker (1998). "Processing of frequency-modulated sounds in the cat's posterior auditory field." *Journal of neurophysiology* 79(5): 2629-42.
- [231] Tian, B. and J. P. Rauschecker (2004). "Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey." *Journal of Neurophysiology* 92(5): 2993-3013.
- [232] Toga, A. W., J. C. Mazziotta and R. S. J. Frackowiak (2000). "Brain Mapping". San Diego, Academic Press.
- [233] Tschopp, K., C. Schillinger, N. Schmid, M. Rausch, D. Bilecen and K. Scheffler (2000). "[Detection of central auditory compensation in unilateral deafness with functional magnetic resonance tomography]." *Laryngorhinootologie* 79(12): 753-7.
- [234] Turner, R. (2002). "How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes." *Neuroimage* 16(4): 1062.
- [235] van de Ven, V. G., E. Formisano, D. Prvulovic, C. H. Roeder and D. E. Linden (2004). "Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest." *Human Brain Mapping* 22(3): 165-78.
- [236] Van Wanrooij, M. M. and A. J. Van Opstal (2005). "Relearning sound localization with a new ear." *Journal of Neuroscience* 25(22): 5413-24.
- [237] Versnel, H., N. Kowalski and S. Shamma (1995). "Ripple Analysis in Ferret Primary Auditory Cortex. III. Topographic Distribution of Ripple Response Parameters." *Auditory Neuroscience* 1(3): 271-286.
- [238] Versnel, H. and S. A. Shamma (1998). "Spectral-ripple representation of steady-state vowels in primary auditory cortex." *Journal of the Acoustical Society of America*, The 103(5 Pt 1): 2502-14.
- [239] Viviani, R., G. Gron and M. Spitzer (2005). "Functional principal component analysis of fMRI data." *Human Brain Mapping* 24(2): 109-29.
- [240] Wakefield, G. H. and N. F. Viemeister (1990). "Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise." *Journal of the Acoustical Society of America*, The 88(3): 1367-73.
- [241] Waldvogel, D., P. van Gelderen, W. Muellbacher, U. Ziemann, I. Immisch and M. Hallett (2000). "The relative metabolic demand of inhibition and excitation." *Nature* 406(6799): 995-8.
- [242] Weismer, S. E., E. Plante, M. Jones and J. B. Tomblin (2005). "A functional magnetic resonance imaging investigation of verbal working memory in adolescents with specific language impairment." *Journal of Speech, Language, and Hearing Research* 48(2): 405-25.
- [243] Weisz, N., C. Wienbruch, S. Hoffmeister and T. Elbert (2004). "Tonotopic organization of the human auditory cortex probed with frequency-modulated tones." *Hearing Research* 191(1-2): 49-58.
- [244] Weisz, N., C. Wienbruch, K. Dohrmann and T. Elbert (2005). "Neuromagnetic indicators of auditory cortical reorganization of tinnitus." *Brain* 128(Pt 11): 2722-31.
- [245] Wessinger, C. M., M. H. Buonocore, C. L. Kussmaul and G. R. Mangun (1997). "Tonotopy in human auditory cortex examined with functional magnetic resonance imaging." *Human brain mapping* 5(1): 18-25.
- [246] Wessinger, C. M., J. VanMeter, B. Tian, J. Van Lare, J. Pekar and J. P. Rauschecker (2001). "Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging." *Journal of cognitive neuroscience* 13(1): 1-7.
- [247] Wise, R. J., S. K. Scott, S. C. Blank, C. J. Mummery, K. Murphy and E. A. Warburton (2001). "Separate neural subsystems within 'Wernicke's area'." *Brain* 124(Pt 1): 83-95.
- [248] Woldorff, M. G., C. Tempelmann, J. Fell, C. Tegeler, B. Gaschler-Markefski, H. Hinrichs, H. J. Heinz and H. Scheich (1999). "Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography." *Human brain mapping* 7(1): 49-66.

-
- [249] Wong, D., D. B. Pisoni, J. Learn, J. T. Gandour, R. T. Miyamoto and G. D. Hutchins (2002). "PET imaging of differential cortical activation by monaural speech and nonspeech stimuli." *Hearing Research* **166**(1-2): 9-23.
- [250] Yang, Y., A. Engelien, W. Engelien, S. Xu, E. Stern and D. A. Silbersweig (2000). "A silent event-related functional MRI technique for brain activation studies without interference of scanner acoustic noise." *Magnetic resonance in medicine* **43**(2): 185-90.
- [251] Yang, Y., H. Gu and E. A. Stein (2004). "Simultaneous MRI acquisition of blood volume, blood flow, and blood oxygenation information during brain activation." *Magnetic Resonance in Medicine* **52**(6): 1407-17.
- [252] Yetkin, F. Z., P. S. Roland, D. B. Mendelsohn and P. D. Purdy (2004a). "Functional magnetic resonance imaging of activation in subcortical auditory pathway." *Laryngoscope* **114**(1): 96-101.
- [253] Yetkin, F. Z., P. S. Roland, W. F. Christensen and P. D. Purdy (2004b). "Silent functional magnetic resonance imaging (fMRI) of tonotopicity and stimulus intensity coding in human primary auditory cortex." *Laryngoscope* **114**(3): 512-8.
- [254] Yost, W. A. (2000). "Fundamentals of hearing". London, Academic Press.
- [255] Zatorre, R. J., A. C. Evans, E. Meyer and A. Gjedde (1992). "Lateralization of phonetic and pitch discrimination in speech processing." *Science* **256**(5058): 846-9.
- [256] Zatorre, R. J., P. Belin and V. B. Penhune (2002). "Structure and function of auditory cortex: music and speech." *Trends in Cognitive Science* **6**(1): 37-46.
- [257] Zatorre, R. J., M. Bouffard and P. Belin (2004). "Sensitivity to auditory object features in human temporal neocortex." *Journal of neuroscience, The* **24**(14): 3637-42.
- [258] Zeng, F. G., K. Nie, G. S. Stickney, Y. Y. Kong, M. Vongphoe, A. Bhargave, C. Wei and K. Cao (2005). "Speech recognition with amplitude and frequency modulations." *Proceedings of the National Academy of Sciences of the United States of America* **102**(7): 2293-8.

§8.2 Curriculum Vitae

■ Dave Langers was born in Stramproy (Limburg, the Netherlands) on July 9, 1973. From 1985 to 1991 he followed pre-university education (VWO) at the Philips van Horne comprehensive school in Weert. Final subjects comprised mathematics, physics, chemistry, biology, economics, and the Dutch, English and French languages. He passed his finals with outstanding results. After that, he went to the Eindhoven University of Technology to study Applied Physics. After passing the Propaedeutic and Bachelor phases, he studied transport phenomena in vortex configurations in thin stratified fluid layers, and his graduation project concerned the topic 'chaotic advection in quasi two-dimensional flows'. After receiving his Master degree in 1997, he completed further training as a fully qualified physics teacher at this same university. In 1999 and 2000 he was employed as a researcher at the R&D facilities of Océ Technologies

■ Dave Langers werd geboren op 9 juli 1973 in Stramproy (Limburg, NL). Van 1985 tot 1991 doorliep hij zijn middelbare schooltijd (VWO) op de Philips van Horne Scholengemeenschap te Weert. Zijn vakkenpakket omvatte wiskunde B, natuurkunde, scheikunde, biologie, economie, Nederlands, Engels en Frans. Hij haalde zijn eindexamens met uitstekende resultaten. Daarna begon hij met zijn studie Technische Natuurkunde aan de Technische Universiteit te Eindhoven. Na het propaedeutisch examen te hebben afgesloten, bestudeerde hij transportverschijnselen in wervelconfiguraties in dunne gestratificeerde vloeistoflagen en betrof zijn afstudeerproject het onderwerp 'chaotische advectie in quasi tweedimensionale stromingen'. In aansluiting op het behalen van zijn ingenieurstitel in 1997 volbracht hij een opleiding tot 1e graads docent natuurkunde aan diezelfde universiteit. In 1999 en 2000 was hij werk-

in Venlo, an international manufacturer of digital printers and copiers. Here he modeled the cooling behavior of hot-melt inkjet droplets and modulation transfer characteristics of fiber optic arrays in printheads. For half a year, he studied intellectual property law and worked as a junior patent officer in this same company. In February 2001 he started a research project as a PhD student in the field of auditory fMRI within the context of a collaboration between the Maastricht University, the Maastricht University Hospital, and the Eindhoven University of Technology. Various studies were performed on acquisition paradigms in auditory fMRI, the function and organization of the central auditory system in humans, and the presence of functional reorganization in the central auditory system under pathological conditions. The project culminated in the public defense of a thesis that included a selection of results for the acquisition of the PhD-title in 2006. Contiguously, he accepted a post-doc position at the Groningen University where, in direct continuation of this project, he will carry on doing research on the central auditory system using functional MRI.

zaam als onderzoeker op de O&O-afdeling van Océ Technologies te Venlo, een internationale producent van digitale printers en kopieerapparatuur. Hier modelleerde hij het afkoelgedrag van hot-melt inkjet druppels en modulatie transfer karakteristieken van optische fiber arrays in printkoppen. Gedurende een half jaar studeerde hij intellectueel eigendomsrecht en werkte hij als octrooigemachtigde i.o. bij ditzelfde bedrijf. In februari 2001 begon hij een promotieonderzoek op het gebied van auditieve fMRI vanuit een samenwerking tussen de Universiteit Maastricht, het Academisch Ziekenhuis Maastricht, en de Technische Universiteit Eindhoven. Verscheidene studies werden uitgevoerd naar acquisitie-paradigma's in auditieve fMRI, de functie en organisatie van het centrale auditieve systeem in de mens, en de aanwezigheid van functionele reorganisatie in het centrale auditieve systeem onder pathologische omstandigheden. Het project resulteerde in de openbare verdediging van een proefschrift met een selectie van resultaten en de verkrijging van de doctorsgraad in 2006. Aansluitend heeft hij een positie als post-doc geaccepteerd aan de universiteit van Groningen, waar hij in het verlengde van dit project verder onderzoek doet aan het centrale auditieve systeem middels functionele MRI.

§8.3 *Financial Support*

This research project was funded by:

- azM Profileringsfonds.

The publication of this thesis was kindly contributed to by:

- Atze Spoor Fonds;
- Philips Medical Systems, Best;
- KNO vereniging Nederland;
- Maastricht University Faculty of Medicine.

§8.4 Postface

When I started a research project in the field of fMRI of the central auditory system as a physicist some five years ago, I could hardly surmise what this would eventually lead to. As you read this, the investigations have been successfully completed and written down, and the results lie before you. Although I realize that a thesis like this cannot ever be complete, I am very proud that I've been able to cover a lot of ground while maintaining a high standard at the same time.

I would like to conclude by thanking those people who have played an indispensable role in the realization of this piece of work. First and foremost of course Walter and Pim, because they provided me with a flying start in this field of research, and for being able to profit from their knowledge and experience. Furthermore, I am indebted to five students from the department of biomedical engineering who have occupied themselves with auditory fMRI in the context of an apprenticeship or graduation project; Suzanne, Lambert, Anne, Karolien and Esther, I think that you did a great job and it was a pleasure to work with you! I also thank Jos and Hans, whom I refrained from involving more closely, but who I know were there in the background.

I owe some thanks to particularly

Toen ik zo'n vijf jaar geleden aan een onderzoeksproject begon op het gebied van fMRI van het centrale auditieve systeem kon ik niet voorzien waar het allemaal toe zou leiden. Terwijl u dit leest zijn de onderzoeken afgerond en uitgewerkt, en liggen de resultaten voor u. Hoewel ik besef dat een proefschrift als dit nooit volledig kan zijn ben ik toch trots dat veel gebieden bestreken zijn en dat het niveau daarbij hoog gehouden is.

Ik wil afsluiten met de mensen te bedanken die een onmisbare rol hebben gespeeld in de totstandkoming van dit werk. Allereerst natuurlijk Walter en Pim, omdat ik van hen een vliegende start heb gekregen en omdat ik constant van hun kennis en ervaring heb kunnen meeprofiteren. Verder ben ik dank verschuldigd aan vijf studenten van de faculteit biomedische technologie die zich bezig hebben gehouden met auditieve fMRI in het kader van een stage- of afstudeeropdracht; Suzanne, Lambert, Anne, Karolien en Esther, ik vind dat jullie prima werk hebben afgeleverd en het was een genoegen om met jullie samen te werken. Ook bedank ik Jos en Hans, die ik nagelaten heb wat nauwer bij het een en ander te betrekken, maar waarvan ik weet dat ze altijd op de achtergrond aanwezig zijn geweest.

Ik bedank ook m.n. Yvonne van de

Yvonne from the audiology department, and her charming Flemish assistants that were tremendously flexible in scheduling and performing audiometric exams during patient inclusion. Also, the people from the IDEE department and the radiological engineering department were essential in the construction of a number of electronic gadgets that were necessary in the course of the actual research. Also Marc, thanks for helping out with all kinds of computer troubles that I have suffered from now and then. Of course, I thank the members of the reading committee who judged and approved the manuscript. And I won't forget to mention the department's MR-operators and secretaries for helping out occasionally, for instance when I didn't know which buttons to push on the MR-scanner or fax-machine, respectively.

Although the volunteers that participated in either of the experiments will remain anonymous (to the reader, but most of them not to the author), it goes without saying that they should be mentioned in these acknowledgements as well. Moreover, I thank my national and international colleagues in related fields for many insightful discussions, interesting meetings, conference talks and colloquia, and enjoyable lab visits.

All the many people that I have worked with in the research department in the former MRI-2 and

afdeling audiologie, en haar charmante Vlaamse assistentes, die altijd enorm flexibel waren in het inplanen en uitvoeren van audiometrie gedurende de patiënteninclusie. De mensen van de afdeling IDEE en de röntgentechnische dienst waren essentieel bij de fabricage van een aantal elektronische apparaatjes die noodzakelijk waren voor de feitelijke uitvoering van de onderzoeken. Ook Marc, bedankt voor de hulp bij de vele soorten computerproblemen die ik bij tijd en wijle ben tegengekomen. Natuurlijk dank ik de leden van de beoordelingscommissie die het manuscript beoordeeld en goedgekeurd hebben. En ik zal niet vergeten ook de MR-laboranten en de dames van het secretariaat van de afdeling te noemen omdat ze me soms uit de brand geholpen hebben, bijvoorbeeld als ik niet wist welk knopje ik moest indrukken op respectievelijk de scanner of de fax.

Hoewel de vrijwilligers die hebben deelgenomen aan de experimenten anoniem zullen blijven (voor de lezer, maar het merendeel niet voor mij), spreekt het voor zich dat ook zij niet in dit dankwoord mogen ontbreken. Verder bedank ik mijn nationale en internationale collega's in aanverwante onderzoeksgebieden voor de vele inzichtelijke discussies, interessante bijeenkomsten, conferentiepraatjes en colloquia, en de plezierige bezoeken aan verscheidene labs.

De vele mensen waar ik mee heb

the later MRI-3 'basements' as colleagues and companions deserve special mention; in particular my fellow PhD-students Quido and Marianne for being such pleasant room mates for three years, and Jaap and Vera for your intrinsic interest in what I was trying to express during our 'trio-meetings'. Equally important, the audiology researchers that were somewhat sparser and more or less scattered but that I certainly had quite some pleasure sharing time with, especially Bas whose defense I really hated to miss. Thank you all for the enjoyable atmosphere and the pleasant conversations about lots of scientific and less scientific topics!

Also thanks to my dedicated 'home front' for being happy to see me again all those evenings when I arrived home late, or each weekend when I visited Stramproy without being empty-handed. I realize that all of you probably won't have understood many things in this thesis, but I also know that this is not all that counts.

And finally to all those people that I haven't forgotten about but for whom this thesis just proves a bit too short: thanks a bunch!

samen mogen werken op de onderzoeksafdeling in de voormalige MRI-2 en de latere MRI-3 'kelders' als collega's en vrienden verdienen een speciale vermelding; in het bijzonder mijn collega-AiO's (of soortgelijk) Quido en Marianne omdat ze zulke leuke kamergenoten waren gedurende drie jaar, en Jaap en Vera voor hun inhoudelijke belangstelling tijdens het 'trio-overleg' voor de dingen die ik zoal uitvogelde. Net zo belangrijk, de onderzoekers bij audilogie die weliswaar wat dunner gezaaid waren en wat meer her en der verspreid, maar waar ik zeker prettige tijden mee heb doorgebracht, in het bijzonder Bas wiens verdediging ik zo jammerlijk moest missen. Dank jullie wel allemaal voor de genoeglijke atmosfeer en de amusante gesprekken over tal van wetenschappelijke en minder wetenschappelijke zaken!

Aûch bedântj aan mien toegewiedje thoësfrônt det altiêd bliej waås om mich weer trûgk te zeên as ich weer ins eine oâvendj laât thoës kwoom, of wan ich oppe zonnig mèt de heinj vól dings in Rooj kwoom aanzette. Ich besef detj gae wersjienlik mer weinig van dit proefschrift zultj snappe, mer ich weît aûch det det neet alles is waât teltj.

En tenslotte aan al die mensen die ik niet vergeten ben, maar waarvoor dit proefschrift gewoon net iets te kort blijkt: een dikke merci!

Maastricht/Groningen, December 2005,
Dave Langers.

■ Q.E.D.

© 1998 Randy Glasbergen, reproduced with permission

**" You've been working awfully hard lately.
If you need a little fresh air and sunshine, you can go to
www.fresh-air-and-sunshine.com "**