Chapter 2

Functional imaging of the central auditory system using PET

Introduction

Hearing is one of the most important senses for communication. In combination with vision and the ability to speak, it contributes in a significant way to the human quality of life. Hearing loss may be serious handicap in social interaction and thus constitutes an important restriction on human functioning. About ten percent of the population suffers from hearing loss. Investigating hearing and hearing loss involves the outer and inner ear, where the sounds are received, up to the brain, where they are interpreted to form a meaningful percept.

The auditory nervous system is the most complex of all sensory pathways. Its structure and function in humans is still poorly understood compared to e.g. the visual system. The last few decades, the knowledge of the central auditory system has increased rapidly thanks to functional imaging tools like Positron Emission Tomography (PET), functional Magnetic Resonance Imaging (fMRI), electroencephalography (EEG) and magnetoencephalography (MEG).

This chapter reviews the current knowledge of the central auditory system. However, this can only be understood thoroughly if we also consider the neuronal networks in which the auditory cortex is embedded, e.g. its subcortical connections as well as connections to other cortical areas. Firstly, we describe the anatomy of the central auditory system. Subsequently, the principles of functional neuroimaging, with an emphasis on the PET-
technique, are given and the state-of-the-art in functional neuroimaging in auditory research is briefly discussed. Finally, we discuss the future perspectives of the PET-technique in functional imaging of the auditory pathways.

Anatomy of the central auditory nervous system

The description of the anatomy below is a simplified representation of current knowledge on the central auditory pathways. Most of this knowledge is based on postmortem human studies and invasive studies with experimental animals like cats and monkeys. The auditory system of the latter is comparable to that of humans, although some structures of connections might slightly differ (Møller, 2000). The auditory nervous system consists of ascending and descending systems. The ascending system has neural fibers that connect the cochlea to the auditory cortex via main nuclei throughout the brainstem. This is the basis for the parallel and hierarchical neural processing that is characteristic of the classical ascending pathway.

Afferent auditory connections

A schematic view of the afferent auditory pathways is shown in Figure 1. The hair cells of the cochlea are innervated by bipolar sensory neurons, which have their cell body in the spiral ganglion. The central axons form the cochlear part of the vestibulocochlear (eight cranial) nerve. Neurons of the eight cranial nerve project to the ipsilateral cochlear nucleus, which is located at the junction of the medulla and pons. The cochlear nucleus can be divided in three parts: an anteroventral, posteroventral and dorsal cochlear nucleus. Each primary nerve fiber connects to all three divisions of the cochlear nucleus and thus presents the information of a single afferent fiber into a multiple channel system (Eggermont, 2001). This
Functional imaging of the central auditory system using PET

Figure 1. A schematic view of the principal human afferent (solid line) and efferent (dotted line) auditory pathways (adapted and modified from Kingsley, 1996).
represents the first stage of the parallel processing in the auditory system. The cochlea is tonotopically organized, which means that each frequency component of a sound stimulates a distinct region of the cochlea. The nerve fibers throughout the auditory system are organized in a systematic way that preserves the tonotopy (Brawer et al., 1974).

Axons from the dorsal cochlear nucleus cross the midline at the dorsal acoustic stria and travel to the contralateral inferior colliculus via the lateral lemniscus. Axons from the posteroverentral cochlear nucleus form the intermediate acoustic stria. Some axons synapse in the ipsilateral and contralateral superior olivary complex, while the main branch joins the lateral lemniscus to ascend to the inferior colliculus. The superior olivary complex is located in the caudal pons and is divided in three parts: the medial superior olive, the lateral superior olive and the nucleus of the trapezoid body. These three nuclei are surrounded by a number of small nuclei, i.e. the peri-olivary nuclei. Some axons of the anteroventral nucleus project to the ipsilateral superior olivary complex. Others travel across the midline via the most ventral acoustic stria, i.e. the trapezoid body, and terminate in the contralateral superior olivary complex. The superior olivary complex is the first structure in the ascending hierarchy to receive binaural input. Its medial and lateral superior olives are involved in the detection of interaural time and intensity differences.

The lateral lemniscus carries fibers from the dorsal and intermediate acoustic striae and from the superior olivary complex. The fibers of the lateral lemniscus terminate in the inferior colliculus, located in the midbrain tectum, and many of them send collateral axons to the nucleus of the lateral lemniscus. The left and right nuclei of the lateral lemniscus are interconnected via the commissure of the lateral lemniscus. The inferior colliculus is believed to integrate spectral and temporal information to construct a spatial auditory map.

Afferent fibers leaving the inferior colliculus innervate mainly the ipsilateral medial geniculate body, the thalamic nucleus of the
auditory system, and the contralateral inferior colliculus. There are two thalamocortical pathways: the first one starting at the ventral part of the medial geniculate body and ending in the primary auditory cortex (PAC), the second one from the medial and dorsal nuclei in the medial geniculate body to the secondary regions. These thalamocortical projections are called the auditory radiation. As described below, there are tight reciprocal connections between the medial geniculate body and the auditory cortical areas. This allows a close regulation of information in both thalamus and cortex. Auditory areas of both hemispheres are connected via the corpus callosum (Kingsley, 1996; Eggermont, 2001; Henkel, 2002; Martin JH, 2003; Griffiths and Giraud, 2004; Langers et al., 2005).

Efferent auditory connections

Apart from the ascending pathway, the auditory nervous system consists also of descending pathways, forming reciprocal connections from the auditory cortex to the cochlea (Figure 1). They form feedback loops that provide circuits to modulate information processing, i.e. they may enhance certain signals and/or suppress others.

The primary auditory cortex also has descending projections back to the ipsilateral thalamus. The auditory corticothalamic pathway forms a feedback loop that allows reverberant activity of the thalamus and cortex (He, 2003; Winer, 2005). There are only a few descending pathways from the thalamus. Instead, many of the axons descending from the cortex bypass the thalamus and travel to the midbrain tectum, where they end in the ipsilateral inferior colliculus. Descending axons from the inferior colliculus terminate mainly in the region of the superior olivary complex.

The olivocochlear efferent system arises from cells in the peri-olivary nuclei in the superior olivary complex and is called the olivocochlear bundle. The lateral olivocochlear axons form ipsilateral synapses on the afferent axons of the inner hair cells in the cochlea and occasionally on the inner
hair cell bodies themselves. The medial olivocochlear cells have bilateral projections that terminate directly on outer hair cells in the cochlea. The efferent terminals on the outer hair cells are more numerous than those on the inner hair cells. The direct feedback to the outer hair cells may influence cochlear mechanisms and consequently the sensitivity and frequency selectivity of the cochlea (Kingsley, 1996; Rouiller and Welker, 2000; Suga et al., 2000; Henkel, 2002; Moore and Linthicum F.H., 2004).

Auditory cortical areas

The human auditory cortex is located within the superior temporal lobe. The primary auditory cortex (PAC) is situated in the medial two-third of the transverse temporal gyrus, also called Heschl’s gyrus (Figure 2).
The transverse gyrus is often partially duplicated into a double or occasionally triple convexity (Leonard et al., 1998). If the transverse gyrus is duplicated, the PAC is located in the anterior-most gyrus. The cytoarchitecture of the PAC is described as koniocortex or granular cortex and designated by Brodmann as area 41 (Brodmann K., 1909) (the regions of Brodmann are discussed in a following section). There is no fixed nomenclature for designating the PAC and BA 41 is one of the most used, but also names as A1, Tc1 or core are used. The position, the extent and the absolute size of this area varies across individuals and between the left and right side in one individual (Morosan et al., 2001; Rademacher et al., 2001b). In general, the PAC in the right hemisphere is located more anteriorly when compared to the left hemisphere (Rademacher et al., 2001b). Through the corpus callosum, each auditory cortical area is connected with the reciprocal areas in the other hemisphere. A study comparing the size of the primary auditory areas in males and females found them to be consistently larger bilaterally in women (Rademacher et al., 2001a).

The human primary auditory cortex is surrounded by and reciprocally connected with a secondary auditory region. This region covers the lateral part of the transverse temporal gyrus and extends anteriorly and posteriorly onto the superior temporal plane (Figure 2) (Moore and Linthicum F.H., 2004). This region was designated as area 42 by Brodmann (1909). This area 42 is in turn surrounded by an extensive auditory area (BA 22) that covers the remainder of the superior temporal plane and the lateral surface of the superior temporal gyrus, with exception of its rostral pole. The exact borders of the non-primary auditory areas are not always clearly and uncontroversially defined (Shapleske et al., 1999; Westbury et al., 1999).

It is thought that the primary area is involved in processing basic sound features like frequency and intensity level, and that the non-primary areas play a role in processing spectrotemporally more complex sounds (Wessinger et al., 2001; Hall et al., 2003; Langers et al., 2003). Part of the left BA 22, at the temporooparietal junction, is called the area of Wernicke and
is important in interpretive speech mechanisms. Wernicke’s area is tightly connected with Broca’s area in the frontal lobe (BA 44, 45). The latter is important in speech production. This temporal-frontal pathway is called the arcuate fasciculus and goes from Wernicke’s area via the angular gyrus and supramarginal gyrus to the frontal operculum where the area of Broca is located (Figure 3).

![Figure 3. Location of language related areas in the human brain.](image)

The anatomical division of the superior temporal pole into three auditory regions has also been found in other species of primates. They are called the core, belt and parabelt, resembling BA 41, 42 and 22 respectively in humans. In these animal studies, extensive connections between the three auditory regions have been demonstrated (e.g. Kaas and Hackett, 1998).
Mind and Brain

History of functional neuroimaging

Untill recently, our knowledge of the brain was based predominantly on post-mortem or lesion human brain studies or invasive animal studies. The past few decades this knowledge has increased rapidly by the development of (functional) neuroimaging techniques. With these techniques we are able to identify the anatomy (e.g. with MRI) and functionality (e.g. with positron emission tomography, PET, or functional magnetic resonance imaging, fMRI) of brain areas in vivo. These techniques rely on a long history of brain studies.

In 400 BC, Hippocrates realized that the brain was the material substrate underlying all cognitive and affective processes. Almost two millennia passed before Franz Joseph Gall (1758-1828) proposed a paradigm, which lead researchers to a more functional specialization of the brain. He stated that all mental faculties have their own distinct material substrate in regions of the brain. Gall thought he could determine the relative sizes of cortical areas and hence the differential mental faculties from the size and shape of the skull. The Catholic Church considered his craniology (later termed phrenology) as contrary to religion (that the mind, created by God, should have a physical seat in brain matter, was an anathema). Also the established science condemned him because he could not provide real scientific proof of his theory, but also because phrenology was quickly taken over by quacks and was considered a kind of moneymaking fraud. Nevertheless, Gall made a major contribution to neuroscience in the sense that he used an anatomical method to study the mind-brain relationship.

After Gall, neuroscience developed in a direction of anatomoclinical evidence. Scientists tried to locate mental abilities in the brain, by studying postmortem brains of patients with disorders and brain lesions. Language disorders were the first to be studied systematically. For example, Paul
Broca (1824-1880) and Carl Wernicke (1848-1904) found areas involved in language processing by studying the brains of patients with language loss, these areas were later named after the scientists. Many other mental impairments, like visual recognition disorders, agraphia, memory loss etc. were described and associated with lesions. However, a lot of objections were made to this anatomoclinical approach. The inference from localized symptoms to localized normal functions is anything but straightforward and needs to be complemented by investigations of the healthy brain.

At the end of the 19th century, researchers realized they had to study the intact living brain to get a better picture of the functional organization of the brain. Cerebral thermometry was one of the first techniques used for this purpose. Lombard explored the cerebral temperature in 1867. He placed thermometers on the scalp and measured temperatures in the corresponding areas during mental tasks. He compared the task measurements with a control-state where no mental effort was made. Lombard was one of the first to understand the importance of a control-state or baseline (Marshall and Fink, 2003). Nowadays modern imaging techniques like PET and fMR rely on the principle of a baseline.

The past few decades, new techniques of functional brain imaging emerged: positron emission tomography (PET), functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and magnetoencephalography (MEG). PET and fMRI measure localized changes in cerebral bloodflow, whereas EEG/MEG measure neuronally mediated electrical and magnetic activity. The importance of these imaging techniques lies in providing investigators with a tool to study how the pattern of activity in the human brain in vivo changes as a function of perceptual or psychological variables. In the study of the brain-behavior relationship, these developments are a significant complement to the anatomoclinical approach.
Principles of functional neuroimaging

The main question when using functional neuroimaging tools is “which areas are involved while performing a task?” or in other words “which areas are active during the task?”. The human brain consists of billions of neurons, which transmit information from one brain region to another by electrochemical impulses. Local neuronal electrical activity is a measure of local brain activity. The most direct, but also invasive, way of studying brain activity in the living brain is direct single cell recording, i.e. an electrode is inserted directly into the brain. This technique is restricted to animal studies because of the invasiveness. This technique gives detailed information about the moment of neuronal firing and about the location of the neuron, hence it has a high temporal and spatial resolution, but only a very small area can be recorded. EEG and MEG are more indirect techniques, they measure respectively the electrical or magnetic field resulting from the synchronized activity of many neurons through electrodes placed on the scalp (EEG) or magnetic detection coils (MEG) while subjects perform a task. They have a very high temporal resolution (about 10-100 ms) but a very poor spatial resolution (> 1 cm).

Other functional neuroimaging techniques for studying the living brain are PET and fMRI. The brain of subjects is scanned with PET or fMRI while performing a task. Neuronal activity in regions involved in the task will change and this will change the energy consumption in that region (Figure 4). This change in activity can either be an in- or decrease, denoted by activation or deactivation of that area respectively. A change in the energy consumption in a region translates in a change in blood flow and glucose utilization (for review see Raichle, 1998). In PET-studies, a radioactive tracer is administered to the subject. PET imaging measures regional cerebral blood flow during different conditions directly by measuring the uptake of a radioactive tracer administered to the subject. This gives a direct measure of the perfusion in the brain. In activated brain
regions, the blood flow increases to respond to the higher glucose demand and the blood oxygenation levels change, i.e. the amount of available oxygen in the activation region will increase (higher supply than demand). This accounts for the blood-oxygen-level-dependent (BOLD) signal of fMRI (Gusnard and Raichle, 2001).

Functional imaging studies conducted with PET or fMRI are so called activation studies. They provide a means of measuring local changes in brain activity in the awake subject. In less than an hour or two, a subject
can be lead through a noninvasive battery of imaging procedures that yields a picture of active brain areas across several conditions. Imaging can be done on normal and lesioned brains. In the early activation studies, the investigator established a task, developed the stimuli, instructed the subject and measured brain activity during the task (e.g. Lauter et al., 1985). But soon scientists learned that this method was not satisfying because the brain activity was not specific enough, they could not associate a specific region to a specific task. This is when the subtraction method evolved. The idea behind the subtraction method in functional imaging is to compare the functional signals between two scans, each representing a different experimental condition. The loci of the differences in signal between the two images presumably delineate brain regions differentially involved in the two conditions. The experimental condition with the least demanding task is often called ‘the baseline’ and is sometimes even a complete resting condition (no stimulation, passively resting in the scanner). The choice of an appropriate baseline is fundamental to the understanding of complex systems and has been subject of many discussions (Gusnard and Raichle, 2001).

A drawback of functional neuroimaging techniques like PET and fMRI is that they do not measure the neural activity directly, but only indirectly via vascular changes. This implicates that it is not possible to distinguish between excitatory and inhibitory neural activity since both processes demand higher energy consumption and consequently an increase in blood flow. However, because inhibitory synapses are less numerous and less metabolically demanding than excitatory synapses, the measured responses will probably result mainly from excitation rather than from inhibition (Waldvogel et al., 2000). A deactivated region is sometimes confused with an inhibitory region. However, a deactivated region is an area in which the rCBF is lower during the stimulus/task than during the baseline. Like Demonet and colleagues (2005) stated, deactivation loci correspond to regions that are inhibited rather than inhibiting.

Positron emission tomography (PET)
Positron emission tomography is based on coincidence detection of gamma radiation resulting from annihilation of positrons emitted by a radioactive tracer that is injected into the subject. Positron emitting radionuclides have a surplus of positively charged particles in their nucleus. Because this is an unstable situation, these radionuclides emit a positron, a particle with the same weight as an electron but with a positive charge, to achieve stability in the nucleus. After travelling a small distance (the average range for the most used radionuclides is less than 2 mm), the positron encounters an electron in the tissue and annihilates (Figure 5). Annihilation means that the mass of the two particles is converted into energy by emission of two photons. For electron-positron annihilation both photons have an energy of 511 keV. The two photons are emitted under an angle of 180 degrees. The axis of the pair of emitted gamma photons is randomly oriented. The two photons travel through the tissue and air. If

![Figure 5](image_url)
two detectors, situated opposite one another, detect a photon simultaneously (i.e. within 5-15 ns), coincidence is counted, also known as “coincidence detection”. Since the positions of the two detectors, which detected the two gamma photons, are known, the line on which the annihilation took place is determined. This line is called ‘the line of response’. A PET camera consists of thousands of scintillation detectors, positioned in multiple rings, which register the radiation at different angles. This allows mapping of the distribution of radioactivity. For now however, the major limiting resolution factor of most PET-scanners is the detector size. The spatial resolution that can be attained with the currently used clinical PET cameras is 4-5 mm.

Limitations of PET

One should be aware of the limitations of a functional neuroimaging technique in order to interpret results appropriately. One of the most important limitations of PET is its use of radioactive tracers. There are strict guidelines regarding the administration of ionizing radiation. Typically, subjects are scanned 10-12 times at 500 MBq [H\textsubscript{2}\textsuperscript{15}O] per scan. Twelve scans represent a total radiation dose of 5.6 millisieverts [mSv]. According to the International Committee on Radiological Protection, administration of 1-10 mSv has a minor to intermediate level of risk. Concerns about the radiation dose limit repeated tests in one subject and exclude PET as a research tool in children. A second limitation is the poor temporal resolution of PET. One of the conditions of PET-research is that during the scan, a steady-state (i.e. constant perfusion) in the brain is accomplished. This means that the stimulation (or task) must begin before the actual measurement and has to continue throughout the scanning period. As stated above, the optimal scanning time is 120 seconds (Kanno et al., 1991). All data collected during these two minutes are summed and hence no inferences can be made on temporal resolution of less than two minutes. Thirdly, the spatial resolution of PET-cameras is limited (4-5 mm). However, some experimental so called
‘Brain PET scanners’ have a better resolution which can be compared to BOLD fMRI scanning. PET-studies investigating the auditory ascending pathway up to the thalamus encounter a lack of sensitivity and power. Activations of auditory brainstem nuclei are more often absent (e.g. Mirz et al., 1999) than present (e.g. Lockwood et al., 1999) when imaging sound processing.

Advantages of PET

On the other hand, although fMRI is overtaking PET as an imaging tool of preference because it does not involve the use of ionizing radiation, PET has several advantages. First and very important in auditory research: PET is a quiet imaging technique. Investigators do not have to worry about possible influences of the very loud scanner noise of fMRI devices. Secondly, patients with (auditory) implants that are not magnet compatible can, in contrast to fMRI studies, participate in PET-studies. Thirdly, close contact with subjects before, during and after scanning is possible in PET-protocols. Fourthly, because PET-scanners are head-only devices, unlike most fMRI equipment, they are less likely to provoke claustrophobic reactions. Furthermore, PET is less sensitive than fMRI for artifacts resulting from subject motion during scanning.
Data analysis

In a functional brain study with PET, a subject is scanned several times under different conditions. To improve the signal to noise ratio, multiple scans of one condition are obtained. But because of the radiation dose, multiple scanning in one person is limited. The signal of one subject is often too weak and therefore it is necessary to scan multiple subjects. This allows making assumptions about the population. Consequently, the data analysis consists of several preprocessing steps, first combining scans within a subject and then comparing the scans of multiple subjects. The aim of the data analysis is to compare different images of different conditions to identify brain areas that respond differently to each condition.

During the whole scanning period, which typically lasts 2 hours, it is possible that a subject’s head changes position. A first step in the data-analysis is the realignment of all scans to one of the scans in the series to obtain the same orientation of all scans of one subject (Figure 6). Next, all participants will have slightly different shaped brains and to compare these different brains of different subjects, it is necessary to transform all the images to match a template brain. All the scans of all the subjects are transformed into a standard coordinate system, which allows comparison of the data to other data sets that are transformed to the same coordinate system. This step is called spatial normalization (Ashburner and Friston, 1999). A final step in the preprocessing is smoothing with an isotropic Gaussian Kernel (Worsley et al., 1996). Smoothing reduces intersubject variability in neuroanatomy and smoothness of the data is required when applying the random Gaussian field theory, needed to correct for the problem of multiple comparisons (see below).

The data are now ready for the actual statistical analysis. Depending on the data, several variables (scans, subjects, groups, conditions) and covariates (e.g. age, scan order, global blood flow) are fitted in a model to represent the data.
Functional brain mapping studies are usually analyzed with some form of statistical parametric mapping (SPM). The term statistical parametric mapping refers to the construction of spatially extended statistical processes to test hypotheses about regionally specific effects (Friston et al., 1991; Friston, 1996). Statistical parametric mapping constructs a map of the brain, which consists of voxels that are, under the null hypothesis, distributed according to a known probability density function, usually the Student’s T or F distributions. Each voxel typically represents the activity of a particular co-ordinate in three-dimensional space. The exact size of a voxel can be chosen by the researcher and depends on the imaging technique being used. Parametric statistical models use the general linear model to describe the variability in the data in terms of experimental and confounding effects, and

Figure 6. Full-color in appendix. Overview of the image preprocessing and analyzing steps (from Johnsrude et al., 2002).
residual variability. The analysis of variance assumes that the underlying distributions are normally distributed and that the variances of the distributions being compared are similar. Hypotheses expressed in terms of the model parameters are assessed at each voxel with univariate statistics.

To explore the effect of a condition on brain activity, contrast weights are attached to each condition. With SPM we test a null hypothesis of no difference in blood flow between the contrasted conditions (Friston et al., 1991). Often one-sided t-tests are applied, implicating that increases (activations) and decreases (deactivations) in blood flow are tested separately. SPM performs the analysis voxel-by-voxel and a correction for these multiple comparisons has to be made. In this thesis we used the False Discovery Rate (FDR) correction to control for these multiple comparisons (Genovese et al., 2002). A common significance threshold is $p<0.05$ corrected for multiple comparisons.

As opposed to the voxel-by-voxel analyses of the standard SPM analyses, also region of interest (ROI) analyses can be conducted by pooling the SPM results. One way of performing ROI analyses is using MarsBaR (MARSeille Boîte À Région d’intérêt, Tzourio-Mazoyer et al., 2002). A ROI analysis pools the data of all voxels in a defined region and hence the statistical sensitivity is improved and inferences about the whole region can be made.
Data identification

The final step of the analysis is the identification of activated regions. There are several ways of labeling an area, namely with stereotaxic coordinate labels, macroanatomic labels or microanatomic labels.

Identification with macroanatomical labels relates brain activity to cortical gyri/sulci or deep brain nuclei. Macroanatomical labels are of great use when there is an established relationship between structure and function. For example, although the exact location of the primary auditory cortex varies across subjects, it has a clear relationship with Heschl’s gyrus (Rademacher et al., 2001b). Unfortunately this is not always the case, because many sulci in the brain are highly variable between subjects (e.g. Juch et al., 2005).

Identification with stereotaxic labels is based on a standard coordinate system, like the widely used atlas of Talairach and Tournoux (Talairach J and Tournoux P, 1988). This atlas displays labeled diagrams of slices of a postmortem brain of a 60-year old woman. A coordinate in a brain that has been transformed to this system matches the same coordinate location in the atlas slices. In the Talairach coordinate system, a brain is aligned to the anterior and posterior commissures (AC and PC), two relative invariant structures in the brain. The brain is rotated so that the AC-PC line is on a horizontal plane and the interhemispheric fissure is on a vertical plane. A coordinate system relative to a x-, y- and z-axis is defined. The origin of these axes is the anterior commissure (Brett et al., 2002). A major advantage of a common coordinate system is the ability to compare different studies using the same coordinates. The limitation of the Talairach coordinate system is the fact that only one female template brain was used to construct the atlas.

Microanatomical labels refer to the cytoarchitecture of the layers of the brain. Regional variations in morphology and distribution of layers often correlate with regional specialization of function (Brett et al., 2002). A century ago, Korbinian Brodmann studied a postmortem brain and divided
the brain in different regions based on the laminar structures (basically, the appearance of the cortex under the light microscope) and numbered the different regions (Figure 7). These so-called Brodmann areas are still widely used (Brodmann K., 1909). It is common in functional imaging to use the Talairach coordinate system to assign Brodmann areas to a site.
of activation (Brodmann K., 1909; Talairach J and Tournoux P, 1988). However, the cytoarchitectonic labels shown in the Talairach atlas are not based on a microstructural analysis. The authors inferred where the areas of Brodmann could be located in their female subject, based on the sulcal landmarks and gross morphology (Brodmann K., 1909; Talairach J and Tournoux P, 1988; Eickhoff et al., 2005). There is a general consensus that a model based on microstructure can be regarded as a functional model of the cortex (Eickhoff et al., 2005). In case of the auditory cortex, there is a clear link between the microscopic appearance of the region and its function.

However, large intersubject variability exists in location and size of cytoarchitectonically defined areas. To overcome this problem, Zilles and colleagues defined probabilistic cytoarchitectonic maps of various cortical areas, including the primary auditory cortex (e.g. Amunts et al., 1999; Amunts et al., 2000; Morosan et al., 2001; Rademacher et al., 2001b). They generated these probability maps by quantifying the degree of intersubject overlap in each stereotaxic position (Rademacher et al., 2001b). The maps are based on the cytoarchitectonic analysis of five male and five female human post-mortem brains. Eickhoff and colleagues (2005) implemented these probability maps in a toolbox which enables investigators to combine functional and anatomical data.
Neuroimaging in auditory research

Most neuroimaging studies on auditory function are confined to the auditory cortex and only a minority studied the subcortical structures (e.g. Giraud et al., 2000; Melcher et al., 2000; Langers et al., 2005). Studies on the auditory cortex can be divided into studies those that focus on the representation of an auditory function in the auditory areas and studies that investigate a neural circuitry on a larger scale.

Functional organization of the auditory cortex

Understanding how the human primary auditory cortex is functionally organized, is an essential step in identifying the neural mechanisms underlying the perception of sound. The relationship between several aspects of auditory stimuli and auditory cortex activation has been investigated with functional neuroimaging tools. Previous functional imaging studies have shown various patterns of activation of the superior temporal areas using different types of non-verbal auditory stimuli like white noise (Mirz et al., 1999), pure tones (Lauter et al., 1985; Lockwood et al., 1999; Mirz et al., 1999), noise bursts (Zatorre et al., 1994), click stimuli (Laureys et al., 2000) and pulsed tones (Tanaka et al., 2000). They show that the cortical response to an acoustic stimulus depends on many factors, like frequency, intensity, presentation rate and complexity of the stimulus (Strainer et al., 1997).

The functional representation of the frequency of an auditory stimulus in the primary auditory cortex was one of the first auditory topics investigated with functional imaging techniques. These studies were based on electrophysiological and cytoarchitectonical findings in non-human primates, which revealed multiple tonotopic maps in primary and surrounding secondary auditory areas (Merzenich and Brugge, 1973; Morel and Kaas, 1992; Morel et al., 1993; Rauschecker et al., 1997).
However, the results in humans were less clear. Discrepancies were found in precise orientation and the number of tonotopic maps (Schonwiesner et al., 2002; Wessinger et al., 2001). Nevertheless, it was consistently shown that high-frequency tones activated the medial and posterior part of the primary auditory cortex and low-frequency tones the anterior-lateral lateral part (Lauter et al., 1985; Pantev et al., 1989; Lockwood et al., 1999; Schonwiesner et al., 2002; Formisano et al., 2003). Hence, the tonotopic organization in the cochlea remains throughout the auditory pathway along the different brainstem nuclei to the primary auditory cortex. But the relationship between cytoarchitectonically defined fields of the auditory cortex and functionally characterized areas in humans remains less established than in non-human primates and questions about the cortical tonotopic maps remain. This might be due to technical limitations of functional neuroimaging tools compared to spatially highly specific electrophysiological measurements in non-human primates.

Not only frequency, but also the intensity of the stimulus has an effect on brain activity. The spatial extent of the neural responses increases with increasing stimulus intensity (Jancke et al., 1998; Lockwood et al., 1999), which is called ampilotopic organization of the auditory cortex. Another important factor is the presentation rate, with the maximal cortical activation to pulsed pure tones at a presentation rate of 5 Hz (Tanaka et al., 2000). Furthermore, in several imaging studies, the effect of stimulation side has been investigated. Studies with pure-tones, consonant-vowel syllables or click stimuli all revealed a lateralization contralateral to the stimulated ear in normal hearing subjects (Lauter et al., 1985; Laureys et al., 2000; Jancke et al., 2002).

Spectral and temporal variations in acoustic stimuli influence the response of auditory areas. The primary auditory cortex in both hemispheres responds to temporal variation, while the anterior superior temporal areas bilaterally respond to the spectral variation. The responses to the temporal features were weighted towards the left, while responses to the
spectral features were weighted towards the right. These findings confirm the specialization of the left-hemisphere auditory cortex for rapid temporal processing (important for language) and right hemisphere specialization for spectral cues (important for music) (Zatorre and Belin, 2001; Zatorre et al., 2002a).

Functional neuroimaging work has revealed a hierarchy of multiple cortical steps in auditory processing that is, like in monkeys (Rauschecker et al., 1997), organized serially and in parallel (Wessinger et al., 2001). The core of the processing hierarchy is located in the medial part of Heschl’s gyrus and later stages of processing spreading laterally and medially form there. Like the visual system, later stages build on the outputs of the early stages. Simple acoustic stimuli like pure tones activate the center of the primary auditory cortex (Figure 8), whereas spectrally more complex tones also activate the surrounding higher order areas (Wessinger et al., 2001).

Figure 8: Full-color in appendix. Activation of auditory areas by noise (A) and music (B). Noise activates the primary auditory cortex, located in the transverse gyrus, whereas the activation of music is located in the primary and surrounding secondary auditory areas. R= right hemisphere, L= left hemisphere.
The knowledge on hierarchical processing was further extended by studies of Kaas and Hackett (1998), Hackett and colleagues (1999), Romanski and colleagues (1999a, 1999b), Rauschecker and Tian (2000) and Romanski and Goldman-Rakic (2002) revealing connections between the secondary auditory areas and areas in parietal and prefrontal lobes in monkeys. They supported a popular working model of two hierarchical sequences of cortical areas, so-called “streams”. The ventral stream specializes in object identification (what is it?) and a dorsal stream specializes in object localization (where is it?). This model was already identified for the human visual system (Ungerleider and Haxby, 1994). With functional neuroimaging it was investigated whether these “where” and “what” streams are also applicable to the human auditory system. Several PET and fMRI studies supported this model (Bushara et al., 1999; Weeks et al., 1999; Belin and Zatorre, 2000; Zatorre et al., 2002b; Arnott et al., 2004). Although some aspects of this model have been argued (Middlebrooks, 2002; Gifford and Cohen, 2005), the auditory dual-stream model has gained support of lesion and functional imaging studies (for review see Rauschecker and Tian, 2000).

Hierarchical processing of the auditory system leads us to complex sound processing like music and language. Speech and music represent the most complex cognitive uses of sound by the human species. They are both built up by rule-base permutations of a limited set of basic elements (phonemes or tones) to result in meaningful structures (words or melodies). Speech and music are interesting stimuli in functional imaging studies, because all normal functioning humans seem to be capable of processing sophisticated music and speech without explicit training and hence they are likely to be related to the functional organization of the human auditory system (Zatorre et al., 2002a). But there is an important difference between the two stimuli: speech consists of rapid temporal changes, whereas music tends to have slower changes but the small variations in frequency are important. Zatorre et al. argue that the auditory cortices in the two hemispheres are relatively specialized, such that temporal resolution (important for speech) is better in
the left auditory cortical areas and spectral resolution (music) is better in the right auditory cortex (Zatorre et al., 2002a).

Speech processing has been extensively studied, because of its scientific and clinical importance and relevance. Speech comprehension comprises several processing stages. The first stage of acoustic analysis of the speech sound in the primary auditory cortex has been described in the preceding section. After the basic acoustic analysis, a speech sound is phonetically and sublexical phonologically processed, i.e. identification and processing of the constituent phonemes. Then, an access to the lexicon and semantics has to be found to retrieve the meaning of the presented word. And finally, words have to put together to form a sentence.

The exact location of the phonological and phonetic speech analysis in the superior temporal region is somewhat controversial. Some studies found left-lateralized posterior superior temporal gyrus activation when listening to a story in a foreign language. Processing foreign speech sounds involves phonological processing of the phonemes, but there is no entry to the semantics of the words (Mazoyer et al., 1993). But others reported bilateral activation of the posterior part of the superior temporal sulcus when comparing pseudowords to tones (Binder et al., 2000). Once the perceived word has been phonetically and phonologically analyzed, the word meaning has to be retrieved. The lexico-semantic processing is not restricted to the superior temporal lobe, but it relies on a whole network of cortical region. This so-called language network involves left anterior superior and middle temporal gyrus, posterior inferior temporal gyrus, angular gyrus, inferior frontal gyrus and dorsal prefrontal areas and in the right hemisphere, it involves posterior inferior temporal gyrus and the angular gyrus (Binder et al., 2000; Demonet et al., 2005). Furthermore, neuroimaging studies on neural substrates of syntax processing have shown the involvement of the left inferior frontal region (Caplan et al., 1999). However, it remains to be shown whether such activations reflect highly specific grammatical/syntactic processing or working memory involvement in general. For a
review on functional neuroimaging and language processing, see Demonet and colleagues (2005).

A relative specialization in the right hemisphere for music processing is supported by functional imaging data from a wide variety of perceptual tasks: pitch judgements in melodies or tones, imagery for tunes, timbre judgements in dichotic stimuli, etc. (Zatorre et al., 1994; Halpern and Zatorre, 1999; Zatorre, 2001; Zatorre, 2003). It is beyond the scope of this introduction to describe in detail the many different (sub)cortical regions found in this multitude of tonal tasks. In short, the overall pattern that emerges is similar to the hierarchical organization of speech stimuli, but now primarily in the right hemisphere: primary auditory areas are involved in low-level processes related to extraction and ordering of pitch information, whereas the secondary and associative areas are involved in the more complex analysis of processing tone patterns (Zatorre et al., 2002a). Finally, because music processing involves organization of sounds in time, working memory is involved and interactions with right frontal lobes are observed while processing music (Zatorre et al., 1994).

Although is it clear that music and language have in part separate neural pathways, there is evidence that some functions, such as syntax, may require common neural resources for speech and music (Patel, 2003). In other words, the ability to organize a set of words into meaningful sentence and the ability to organize a set of notes into a structured melody might engage the same brain structures.

The cortical response to an acoustic stimulus depends on many variables as shown above, like frequency, intensity, melody. But also the subject’s attention can influence neural responses to sounds. When greater demands are made on auditory selective attention during dichotic tasks, activation within the primary auditory cortex (together with prefrontal attention areas) was increased (Pugh et al., 1996; Tzourio et al., 1997; Zatorre et al., 1999). Attentional modulation presumably performs an important gating function for perceptual information.
Non-auditory functions of the auditory cortex

It is generally believed that a primary sensory cortex is restricted to responding to that particular sensory input and that information of different modalities integrate on a higher level in the cortical hierarchy, namely in the secondary and associative areas. Although in many respects this view is probably still correct, recent studies have yielded data suggesting that primary cortices may have a function in processing stimuli other than their primary modality.

Neuroimaging data of congenitally blind and deaf subjects demonstrated input variability in a primary sensory area after loss of that sensory modality. Blind subjects showed activation of the primary visual cortex during tactile and auditory tasks (Sadato et al., 1996; Cohen et al., 1997; Weeks et al., 2000). Later, also the input plasticity of the primary auditory cortex was investigated in early deaf subjects. It was found that secondary, but also primary auditory areas were involved in the processing of purely visual stimuli (Finney et al., 2001). This suggests that removal of one sensory modality leads to neural reorganization of the remaining modalities.

In the light of plasticity of the auditory cortex, patients with cochlear implantation are a unique subject group. With a cochlear implant deaf subjects can (re)gain (part of) their hearing ability. Deafness is a model of sensorineural deprivation and cochlear implantation provides an interesting opportunity to examine how cortical responses change as a function of time when sensory stimulation is again provided via the implant device. The past ten years, functional neuroimaging has been used to answer plasticity questions related to this patient group. It is beyond the scope of this introduction to review all these experiments (for review see Truy (1999) and Syka (2002). Still, one study of Giraud and colleagues (2001) clearly marks the important integrating role of the auditory areas. They studied postlingually deaf subjects with PET at different time points after
implantation to measure brain activation during stimulation with words or various noise control conditions. It was shown that the primary auditory cortex activated increasingly with postimplantation time. It reflects that cortical activity is associated with decoding speech sounds, which improves over time. Another remarkable result in the study of Giraud et al. was a concomitant increase in visual cortical activity. The authors propose that the visual cortex participates in speech decoding because deaf subjects have learned to use lipreading along with auditory information to understand speech. Since a cochlear implant provides a degraded input to the auditory cortex, compared to a natural situation of speech comprehension, implant users rely on lipreading in addition to the auditory input.

This interaction between visual and auditory information in speech processing is not new. Already in the 70’s, McGurk showed that visual speech information can influence the auditory percept, also known as the McGurk-effect (McGurk and MacDonald, 1976). Two decades later, a functional neuroimaging study showed that viewing silent videos of people mouthing words is sufficient to activate the primary auditory cortex in normal hearing individuals (Calvert et al., 1997). This shows that purely visual stimuli can activate the primary auditory cortex in persons with a normal hearing ability, although these results could not be repeated by others (Bernstein et al., 2002; Paulesu et al., 2003).

Activation of the primary auditory cortex without acoustic stimulation was also found during auditory verbal hallucinations in schizophrenic subjects (Dierks et al., 1999). In experiments with non-schizophrenic subjects, no primary auditory cortex activation was found during hallucinations (e.g. Griffiths, 2000). It is thought that activation of the PAC during psychotic hallucinations might help to explain why the hallucinations are perceived as real (attributed to external sources) (Dierks et al., 1999).

These examples of normal and impaired subjects demonstrate that the function of what we call the auditory cortex is not defined on the cortical
level by some independent intrinsic factors, but rather by the pattern of input connections and type of sensory input. The traditional view of primary sensory cortices being devoted to solely processing that modality is too straightforward and is more subtle.

In conclusion, functional neuroimaging studies increased the knowledge on the function of the central auditory system. Researchers proposed answers on fundamental and clinical questions about auditory processing. However, the central auditory system is still poorly understood in comparison to other sensory systems. Therefore, future research is needed to further map the different aspects of auditory processing and their relationship with hearing impairment and its treatment.
References


Chapter 2


