MECHANISMS OF SPATIAL AND NON-SPATIAL AUDITORY SELECTIVE ATTENTION

Aspasia Eleni Paltoglou, MA

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

December 2008

Abstract

Selective attention is a crucial function that encompasses all perceptual modalities and which enables us to focus on the behaviorally relevant information and ignore the rest. The main goal of the thesis is to test well-established hypotheses about the mechanisms of visual selective attention in the auditory domain using behavioral and neuroimaging methods.

Two fMRI studies (Experiments 1 and 2) test the hypothesis of feature-specific attentional enhancement. This hypothesis states that when attending to an object or a feature, there should be an enhancement of the response in the sensory region that is sensitive to that object or feature. Experiment 1 investigated feature-specific attentional modulation mainly within the tonotopic fields around primary auditory cortex. Experiment 2 investigated feature-specific attentional modulation mainly around non-primary auditory cortex, when attending to frequency modulation or motion of the same auditory object. Experiment 1 showed evidence for feature-specific enhancement, while Experiment 2 did not. The role of competition among concurrent auditory objects as a necessary factor in driving feature-specific enhancement is discussed.

A second hypothesis from vision research is that spatial perception and attention is much more precise in the centre than in the periphery. Experiment 3 used a masking release paradigm to investigate whether the acuity of auditory spatial attention was similarly increased in the midline. Although location discrimination of sounds segregated by inter-aural time differences was more precise at the midline than at the periphery, spatial attention was not.

Therefore for this task at least there was no effect of eccentricity on auditory spatial attention.

The results of these three studies are discussed in view of selective attention as a flexible process that operates in different ways according to the specifics of the task.

Acknowledgements

I am deeply grateful to my supervisor Prof. Deb Hall, for all her support and patience. A good supervisor makes such a difference, and I had one of the best. Many thanks to Dr. Chris Sumner for his support, especially his useful comments on the thesis. To Dr. Michael Akeroyd for his help with matlab codes for stimulus generation. To Kay Head and John Foster for helping with scanning. Many thanks to Dr. Jon Peirce for offering his support whenever asked for.

Many thanks to my family and friends for all their love and support. Dad, come on, it’s your turn to finish your master’s this year! Thank you for always finding a funny side in everything. Many thanks to my favorite undergraduate lecturer, and now good friend Dr. Kate Michalopoulou, for encouraging me to start my adventure in the UK. It’s been great, you were right.

I’m grateful to my favorite man, Adam Freeman, for keeping me happy and sane, and saving me from all my bad dreams.

I’ll miss everybody at IHR, especially good friends such as Daudi Magezi, Barrie Edmonds, Jeanne Giraud, Daphne Garcia, Ana Alves-Pinto, Bernhard Seeber. I can’t stress enough how happy I’ve been in IHR, and how much I feel I’ve changed for the better during my time here. All the help I needed, it’s always been available. And the atmosphere has always been friendly, embracing and encouraging. And boiling with ideas about all things hearing. So it’s only fair to dedicate this thesis to all my IHR colleagues, past and present.

Declaration

This thesis is the candidate's own original work and has not, whether in the same or different form, been submitted to this or any other University for a degree. All experiments were designed and analyzed by the candidate, and all testing was conducted by the candidate. The only exception was in Experiment 3 (Chapter 5) in which behavioral screening of 14 participants and fMRI scanning of four of them was conducted by two final-year undergraduate project students. Nevertheless, the candidate had responsibility for supervising these students and so was involved throughout.

Conference Presentations

Roberts, K.L., Todd, A., **Paltoglou, A.E**., Summerfield, A.Q., & Hall, D.A. (2005) Auditory spatial attention: Influence of stimulus presentation method. Poster presented at British Society of Audiology meeting, Cardiff, UK.

Paltoglou A.E. and Hall, D.A. (2005) The azimuthal distribution of spatial attention using ITD cues. Poster presented at the Experimental Psychology Society, London Meeting.

Paltoglou A.E. and Hall, D.A. (2006) Attentional modulation of frequency-sensitive regions of human auditory cortex measured with fMRI. Poster presented at the International Conference for the Auditory Cortex, Belton Woods, Grantham.

Paltoglou A.E. and Hall, D.A. (2007) Selective attention to spatial and non-spatial sound features measured with fMRI. Poster presented at British Society of Audiology meeting, London, UK.

Paltoglou A.E. and Hall, D.A. (2008) Selective attention to spatial and non-spatial sound features measured with fMRI. Oral presentation at the Association for Research in Otolaryngology Midwinter Meeting, Phoenix, Arizona, USA.

Publications

Hall D.A. and Paltoglou, A.E. (2009) FMRI of the central auditory system. In M. Filippi (Ed.), New Jersey: Humana Press.

List of Abbreviations

A	Anterior
AAF	Anterior auditory field (cat)
AFC	Alternative forced-choice
AI	Primary auditory cortex
AII	Non-primary auditory cortex
AL	Antero-lateral
AM	Amplitude modulation
ANOVA	Analysis of variance
BA	Brodmann area
BF	Best frequency
BOLD	Blood-oxygen-level dependent
CM	Caudo-medial (monkey)
cm	Centimeters
CP	Caudal parabelt (monkey)
D	Dorsal
dB	Decibels
DLPFC	Dorsolateral prefrontal cortex
DZ	Dorsal zone (cat)
EEG	Electroencephalography
EPI	Echo-planar imaging
ERP	Event-related potential
FA	False alarm(s)
FM	Frequency modulation

FP	False positive
fMRI	Functional magnetic resonance imaging
FWHM	Full width at half maximum
GLM	General linear model
HG	Heschl’s gyrus
HRTF	Head-related transfer function
HS	Heschl’s sulcus
Hz	Hertz
IC	Inferior colliculus
ILD	Interaural level difference
IOR	Inhibition of return
IPL	Inferior parietal lobe
IT	Inferior temporal
ITD	Interaural time difference
L	Lateral
LA	Lateral anterior (human)
LFP	Local field potentials
LGN	Lateral geniculate nucleus
LIP	Lateral intraparietal area
MEG	Magnetoencephalography
MFG	Middle frontal gyrus
MGB	Medial geniculate body
min	Minute(s)
mm	Millimeter(s)

MM	Medio-medial
MMN	Mismatch negativity
MNI	Montreal Neurological Institute
ms	Millisecond(s)
µs	Microsecond(s)
MT	Motion-sensitive visual region (primates)
OP	Operculum
PA	Posterior anterior (human)
PAC	Primary auditory cortex (human)
PAF	Posterior anterior field (cat)
PET	Positron emission tomography
PP	Planum polare
PT	Planum temporale
RF	Receptive field
RM	Rostro-medial
ROI	Region of interest
RP	Rostral parabelt
RTL	Rostro-temporal lateral (monkey)
RTM	Rostro-temporal medial (monkey)
Sb	Subject(s)
s	Second(s)
SEF	Supplementary eye field
SNR	Signal-to-noise ratio
SOA	Stimulus onset asynchrony

SOC	Superior olivary complex
SPL	Superior parietal lobe
STG	Superior temporal gyrus
STRF	Spectro-temporal receptive fields
TE	Echo time
TMS	Transcranial magnetic stimulation
TPJ	Temporo-parietal junction
TPO	Temporo-parieto-occipital junction
TR	Repetition time
V	Ventral
V1	Primary visual cortex
V2	Orientation-sensitive visual region
V4	Color-sensitive visual region
V5	Motion-sensitive visual region (humans)

Table of Contents

Chapter 1: Mechanisms of selective attention

1.1 Introduction

The research presented in this thesis investigates the mechanisms of auditory selective attention using both functional magnetic resonance imaging (fMRI) and behavioral testing. Our sensory system constantly receives information from the surrounding environment. We cannot process all of this information because our cognitive system has limited processing capacity (Desimone & Duncan, 1995). Selective attention refers to the ability to focus on behaviorally relevant information, thus attenuating or ‘ignoring’ information that is less relevant. Sensory processing can proceed in the absence of selective attention. Sensory processing of sound features is reviewed in Chapter 2, while attentional modulation of this process is reviewed in Chapter 1.

Much more research on sensory and attentional processing has been conducted in vision than in audition. Typically, auditory research is influenced by hypotheses that have been tested and successfully proven in the visual system. For example, a model of auditory cortical processing that was popular in the last decade posited two pathways, a dorsal ‘where’ pathway and a ventral ‘what’ pathway (Lomber & Malhotra, 2008). This dual-route framework was first proposed for the visual modality (Mishkin, Ungerleider, & Macko, 1983). Another example concerns the attentional syndromes of unilateral spatial neglect. It has been shown that these patients are not only impaired at attending to visual objects in left hemispace, but also to auditory objects in left hemispace (Eramudugolla, Irvine, & Mattingley, 2007). It is reasonable to assume that the two systems operate by similar mechanisms since they share common characteristics, such as hierarchical organization and spatially segregated (‘modular’) regions for processing different features.

1.1.2 Overview of the thesis

Two sets of hypotheses were tested. In Experiments 1 and 2, the main hypothesis concerned whether attentional modulation is feature specific, i.e. whether the attentional enhancement and suppression of the response is restricted to cortical regions that are sensitive to the attended feature. This hypothesis was tested using fMRI which is the ideal tool for investigating the spatial organization of the cortical response because of its very good spatial resolution. Positive evidence for both feature-specific attentional enhancement and suppression would give support to the notion of attention as a supramodal mechanism that operates in a similar way across the different modalities. In Experiment 3, the main hypothesis concerned whether the efficiency of auditory spatial attention varies as a function of eccentricity. In vision, when the task involves a static visual array, visual perception is optimal at the fovea and declines as the target moves to the periphery of the visual field (Levi, McGraw, & Klein, 2000). Auditory spatial acuity also shows the same effect of eccentricity (Mills, 1958). It was therefore expected that the efficiency of auditory spatial attention should also diminish as the target moves to the periphery (i.e. away from centre).

To illustrate the rationale for Experiments 1-3, Chapter 1 reviews some of the behavioral, neuroimaging and computational modeling literature for visual and auditory processing and for selective attention. This chapter has two parts. The first part includes a brief review of the attentional theories as well as behavioral evidence for filtering of irrelevant information. This is followed by a discussion of behavioral paradigms, such as the probe-signal paradigm, used to examine the gradient of attention. The second part discusses the literature on the neural correlates of visual and auditory selective attention as examined by animal physiological and human neuroimaging studies. Mechanisms include an enhancement of the neuronal response and a sharpening or shifting of the receptive field for the attended stimulus. Additionally, some of the computational models proposed to explain the neural implementation of enhancement are discussed. Note that, throughout Chapter 1, results from the visual and auditory systems are discussed and compared. To provide further background for the rationale of Experiments 1-3, Chapter 2 briefly reviews findings of sensory coding for sound features: frequency, spatial location, spatial motion and frequency modulation (FM). Selective attention towards these key sound features will be manipulated in later experiments.

Chapter 4 describes Experiment 1 which investigated whether attending to a particular sound frequency enhanced the response in primary auditory cortical regions that were sensitive to that frequency. This first experiment was quite complex and so Chapter 3 provides an account of the initial pilot studies conducted to identify the optimal parameters for the stimuli, task, voxel resolution and echo time (TE). Chapter 5 describes Experiment 2 which continued the same theme by investigating whether attending to auditory motion or frequency modulation enhanced the response in non-primary auditory cortical regions that were sensitive to each feature. Chapter 6 presents Experiment 3 which investigated whether the acuity of auditory selective attention declines with increased eccentricity. Finally, Chapter 7 discusses the findings of the three studies and provides a synthesis of the conclusions with respect to a model of auditory selective attention.

1.2 Filtering of irrelevant information

Behavioral research on selective attention started in the 1950s and at first involved experimenting mainly with auditory stimuli. This research suggested that the human cognitive system has limited processing capacity, and could not process all the incoming information. One of the paradigms used was dichotic listening, whereby different messages are played simultaneously to each ear. Cherry (1953) showed that when listeners had to repeat information presented in the attended ear, they could report very little of the information in the unattended ear. They could only report the physical characteristics of the sounds, but not the meaning. These two sets of results inspired Broadbent's early selection model of attention (Broadbent, 1958). This model proposes that only sensory input that is attended can be fully processed, while the rest is filtered out very early. The filter lets through all the attended information, while it lets through only some physical characteristics of the unattended information.

The work of Treisman (1960) led to further refinements of this theory, namely that attentional selection could occur much later than Broadbent had assumed. She showed that unattended information *could* be processed. The degree to which it was processed, depended on how relevant its content was to the attended message. If the messages in the two ears were similar then listeners could retrieve some information from the unattended ear. Treisman's findings led to the formation of the late-selection theory by Deutsch and Deutsch (1963). They claimed that all incoming information was analyzed for meaning, no matter if it was attended or not. After information entered the perceptual system, it was grouped or segregated and assigned a weight according to its importance to the person. This is a late selection model of attention.

Recently, the late selection model of attention has been further revised on the basis of new evidence. Nilli Lavie has demonstrated in a series of studies that the perceptual load of the selective attention task determines whether, and to what extent, the unattended information is processed (for a review see Lavie, 2005). Note that perceptual load is not identical to task difficulty although it is related to it, since presumably a higher load task is also more difficult. As Lavie notes, perceptual load is increased when i) the number of items is increased, or ii) the same number of items poses more demands on attention. Manipulating task difficulty would involve, for example, manipulating the contrast of a visual display (Ress, Backus, & Heeger, 2000).

Lavie's behavioral studies have shown that there is reduced (or eliminated) interference by distractors in conditions with higher perceptual load (Lavie, 1995; Lavie & Cox, 1997). Conditions with lower perceptual load lead to interference from task-irrelevant distractors, because the available resources are not taken up entirely by the primary task. This is known as the perceptual load theory.

Furthermore, there is neuroimaging evidence looking at the neural correlates of the effects of perceptual load on selective attention (for similar effects of task difficulty see Section 1.8.3, Sylvester, Jack, Corbetta, & Shulman, 2008). Rees and colleagues (1997) showed that a moving distractor would be processed to different degrees depending on the attentional load of an unrelated task. The distractors were dots at the periphery of the visual field which either moved or were static. Participants were instructed to ignore the dots and focus on the midline, where they performed a word-based task. In the low-load condition, they had to respond whenever the word contained capital letters. In the high-load condition, they had to respond whenever the word contained two syllables. Rees and his colleagues measured the degree of visual motion processing in visual motion regions of the brain (V5) in the low- and high-load conditions using fMRI. As Figure 1.1 shows the response in the low-load condition was greater for the moving than for the static distractors. In contrast, in the high-load condition, there was little or no response for either the moving or the static distractors. However, it is interesting to note that in one of the high load ‘no motion’ epochs, there was a slight increase in the response, which is rather inexplicable, and is not commented by the authors.

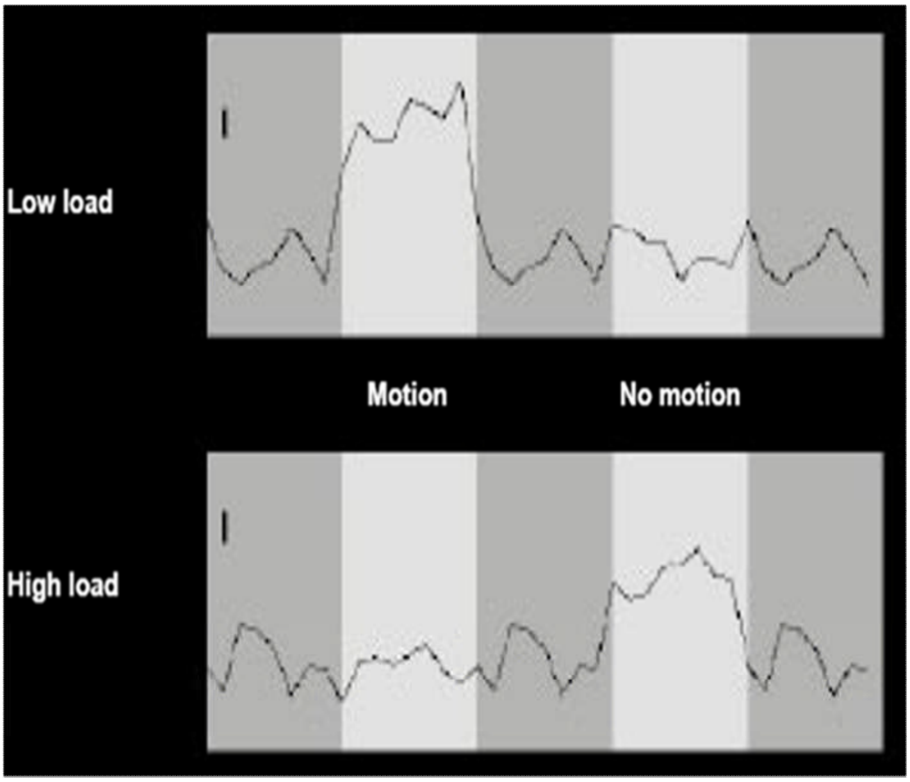


Figure 1.1 Group average response in left V5, for the four experimental conditions (light grey) and the resting baseline (dark grey). The vertical line on the left of the graphs denotes a value of 0.1% fMRI signal change (Rees et al., 1997).

1.3 Objects, spatial locations and features-based attention

An early model considered visual selective attention as a ‘spotlight’, where one could attend to anything within the spatial spotlight (Posner, Snyder, & Davidson, 1980). This model emphasizes the importance of attending to spatial locations. However, more recent evidence suggests that each object can represent one ‘unit’ of attention. The more objects within the visual field, the more resources are required for stimulus processing. Behavioral testing by Duncan (1984) showed that, although there was no cost when attending to two features within the same object, there was a cost when attending to two features from separate objects. This phenomenon is called the ‘same-object advantage’. These results cannot be explained by the spotlight theory, as it predicts that within a particular space performance is without cost and it drops outside this spotlight. Neuroimaging evidence for object-based attention has been reported by O’Craven et al. (1999). They showed that when attending to a feature of an object, an enhanced response was also found to unattended features.

There is contrasting evidence that it is possible to selectively attend to different features within a visual object (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990) since response enhancement has been found exclusively in the regions that are sensitive to that feature. The most pervasive support for the existence of feature-based attention comes from fMRI (Sàenz, Buracas, & Boynton, 2003) and physiological studies (Bichot, Rossi, & Desimone, 2005). When several visual objects shared the same attended feature (e.g. the color red), the response to the unattended object also showed a response enhancement.

There has been little research on this topic in audition, but one fMRI study is also worthy of mention here because it has shown that object- and feature-based selective attention can coexist (Krumbholz, Eickhoff, & Fink, 2007). Object-based attention was demonstrated by showing that activity in motion-sensitive regions was enhanced even when attention was not directed to spatial motion, but to another feature of the same object. Specifically, this condition (‘ignore motion but attend to another feature of the same object’) showed enhancement compared to a condition whereby the two features were part of separate objects and motion was unattended. Feature-based attention was demonstrated by an enhanced response in the motion-sensitive region when attending to motion compared to attending to pitch for the same object. On the other hand, there was no attentional enhancement in the pitch-sensitive region. The authors speculate that either pitch is a feature readily discriminated and thus does not require attention, or the pitch task used was too easy to require attention.

Thus, it appears that participants can attend flexibly to any of these three levels of sensory representation; locations, features and objects, depending on the task at hand.

1.4 Attending to sound frequency: The auditory filter and the attentional filter

A specific place of maximum vibration along the basilar membrane in the cochlea can be thought of as a frequency-based ‘filter’. When a sound pressure wave that represents a single sinusoid (i.e. single frequency) is presented to the ear, it vibrates a specific place along the basilar membrane more than any other place along the membrane. This process is determined by the stiffness of the basilar membrane.

The notion of an auditory filter describes the width of the narrowest frequency channel that can be processed independently and is of central importance for perceptual encoding, as well as for selective attention. In the peripheral auditory system, the first stage of sensory analysis involves ‘breaking down’ the input sound into its constituent frequency components. Taken along its entire length, the basilar membrane can be thought of as behaving like a ‘bank of band-pass filters’ or a set of critical bands, each sensitive to a different frequency. Hair cells at each place of maximum vibration on the membrane transform the mechanical energy into electrical energy that is transmitted along the auditory nerve. The amount of excitation along the membrane decreases as one moves further away from the best filter location and so a neural tuning curve can be plotted to show the degree of frequency selectivity or ‘width’ of each filter.

The width of each auditory filter (or ‘critical band’) that is underlined by the neural filtering, has been investigated psychophysically by many different paradigms, but one of the most popular method is the notched-noise technique (Patterson & Moore, 1986). In the notched-noise technique, listeners are presented with a pure-tone signal plus two band-pass noise bursts. One noise-burst is below the frequency of the pure-tone signal and one is above. The frequency gap between the signal and the two noise bursts is varied, and the signal detection threshold is measured for the different gaps. The closer in frequency the noise bursts are to the signal, the higher the signal threshold will be, because more of the noise will fall within the auditory filter. For frequencies between 750-3000 Hz, the width of the auditory filter is about 12% of the centre frequency (Patterson & Moore, 1986). For example, for a tone of 1000 Hz, the filter is 120 Hz, 60 Hz below 1000 and 60 above.

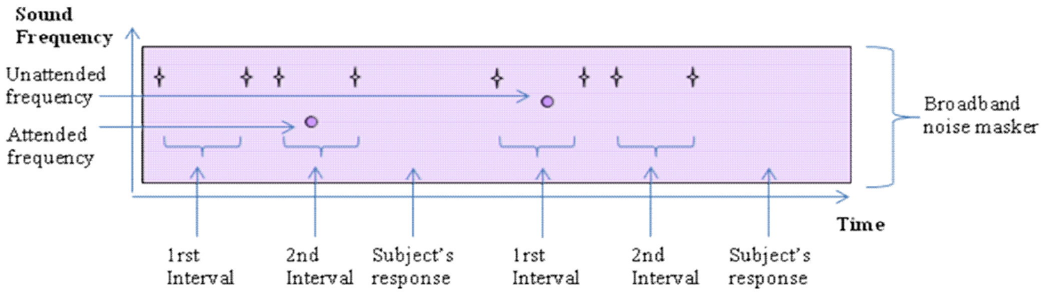


Figure 1.2 Probe-signal method. The stars represent the visual cues at the start and end of each interval. The tones (purple dots) are presented in continuous background noise (rectangle).

The same filter characteristics can be applied to the width of the listening band when participants are specifically required to attend to sound frequency. The auditory filter which is driven by bottom-up input, while the attentional filter is assumed to be modulated by top-down input and so it could be called an ‘active’ filter. In audition, it has long been established behaviorally that when participants expect a pure tone at a specific frequency, their ability to detect that tone in a wideband noise masker is significantly better than when the tone is one of an unexpected frequency (Greenberg & Larkin, 1968). An illustration of the paradigm is shown in Figure 1.2. Each trial contained two intervals, one of which contained a tone. Participants had to indicate which interval contained the tone. In the majority of the trials (about 70%), the frequency was the attended or expected frequency. Participants learned to expect the specific tone by different methods, such as being provided with a lot of practice, or being presented with a pure tone of identical frequency before each trial. The attended tone was either 1000 or 1100 Hz, while the unexpected tones had frequencies higher and lower the expected tone (e.g. for an attended tone at 1100 Hz, the unattended probes were at 600, 700, 800, 900, 1000, 1050, 1150, 1200, 1300, 1400, 1500, and 1600 Hz). The benefit of selective attention for signal detection thresholds can be plotted as a function of frequency (Figure 1.3). The ability to detect tones at frequencies close to the expected frequency was also enhanced, and this benefit dropped off smoothly with the distance away from the expected frequency (Greenberg & Larkin, 1968).

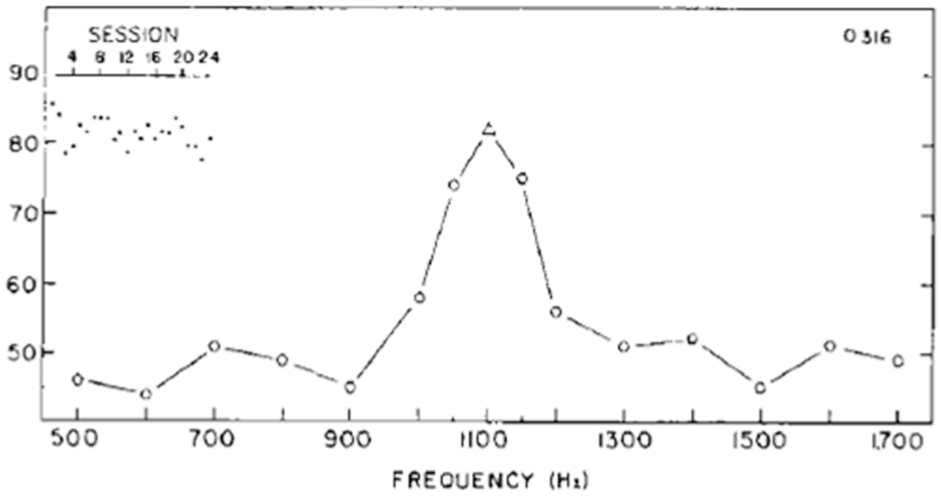


Figure 1.3 Typical results from the probe-signal paradigm (taken from Greenberg & Larkin,1968). Vertical axis denotes % correct (50% is chance). Horizontal axis denotes the different frequencies used.

The original results by Greenberg and Larkin (1968) have been confirmed by other studies, such as the one by Dai et al. (1991). In their study, Dai and colleagues extended the findings by directly comparing the shape of the attentional filter derived using the probe-signal technique, with the shape of the auditory filter derived using the notched-noise technique. The results suggest that both filters were very similar in shape being widest for low frequencies and narrowest for high frequencies (when frequency is plotted on a logarithmic scale) (Figure 1.4). These results support a model whereby listeners attend to frequency through a filter that is centered on the attended frequency, and has a similar shape to the auditory filter. The equivalence between the two filters further suggests that selective attention might be operating at the level of the sensory representation of frequency, (i.e. the cochlea).

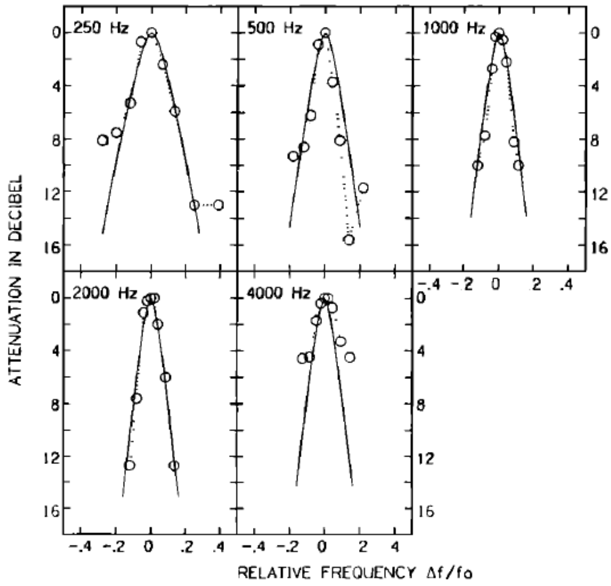


Figure 1.4 Comparison between the auditory filter and the attentional filter (taken from Dai et al., 1991). Circles connected with dotted lines denote the attentional filter, for which % correct scores were converted to a measure of attenuation. Solid lines denote the auditory filter derived from Patterson and Moore (1986).

It is possible to direct attention across more than one pure tone, although there may be a cost involved. To examine this issue, Schlauch and Hafter (1991) manipulated target uncertainty by using one, two or four tone cues prior to each tone target (Figure 1.5). The four cues fell within different auditory filters. The ‘attended’ target had the same frequency as one of the cues (74% of the trials). Again, the further away in frequency the unexpected tones were from the expected tones, the higher the signal-detection threshold. Additionally, the estimated width of the attentional filter increased as the degree of target uncertainty increased. Specifically, the width was 12% of the central frequency for the one-cue condition, 12.4% for the two-cue condition and 13.7% for the four-cue condition. The conclusion from these results was that listeners could divide attention across more than one attentional filter, but not without cost. As the number of filters that were attended increased, the filters became broader, and the detection performance worsened.

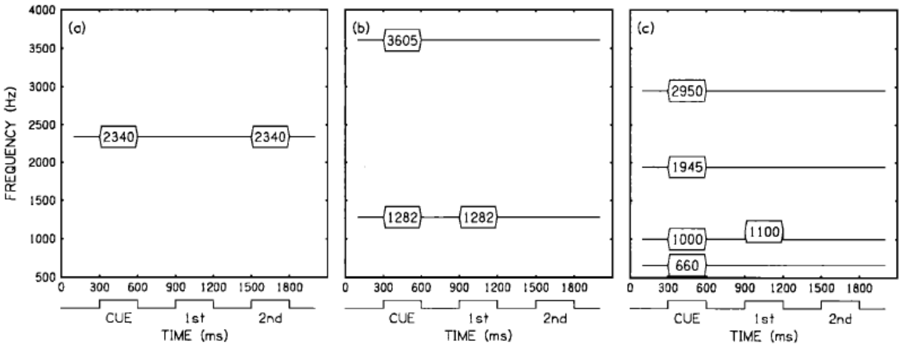


Figure 1.5 Examples of the three conditions reported by Schlauch and Hafter (1991), with (a) one, (b) two and (c) four cues. In (a) and (b) the signal was expected, because it was identical to one of the cues, while in (c) it was unexpected, because it was not.

1.5 Attending to spatial location

Perceptual discrimination of visual targets at the fovea (i.e. midline) is more accurate than at the periphery because the fovea contains a high density of receptor cells thus enabling very sharp acuity (Levi et al., 2000; Morrone, Burr, & Spinelli, 1989). Given this spatial gradient of sensory acuity, the question of interest here is whether visual spatial selective attention also declines from the midline to the periphery. Posner (1978) favored a view that the acuity of visual spatial attention is not influenced by the retinal regions to which attention is directed to. However, there is much accumulated evidence that this is not the case.

Whether performance actually improves or worsens with eccentricity appears to depend on the task. For temporal tasks, such as motion detection, performance is worse towards the midline and improves towards the periphery (McKee & Nakayama, 1984). This result is possibly due to the greater density of motion-sensitive receptor cells (rod cells) in the peripheral visual field. Better performance in the peripheral visual field is also reported for some visual search tasks (Carrasco, Giordano, & McElree, 2006). However, performance is typically best at the midline and declines with eccentricity for static visual tasks that require a high degree of spatial resolution. For example, reaction times are quicker on a spatial cueing task when attending to visual targets at the midline than at the periphery (Golla, Ignashchenkova, Haarmeier, & Thier, 2004; Humphreys, 1981; Yeshurun & Carrasco, 1999). Specifically Golla et al. (2004) examined the effect of eccentricity on visual spatial attention in monkeys and humans, using a spatial cueing paradigm. The results are shown in Figure 1.6. The target was a Landolt ‘C’, which is a circle containing a gap. Participants had to decide whether the gap was on the top or the bottom of the ‘C’. Both the size of the gap and the size of the ‘C’ were varied adaptively according to individual performance. In most trials, the central fixation was followed by a cue (either right or left) and then a target at the same spatial location. The eccentricity of the cues and targets was manipulated across three horizontal positions (3°, 9º and 15°). Performance for both monkeys and humans declined the further away from the midline.

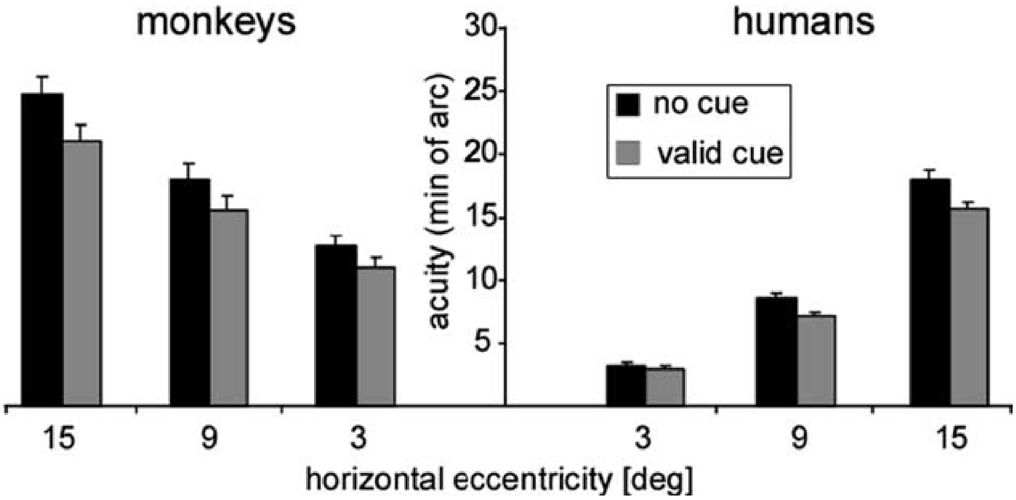


Figure 1.6 Acuity in horizontal eccentricity for both monkeys (left panel) and humans (right panel) across the three conditions (Golla et al., 2004).

As in vision, it has been shown in audition that acuity (as represented by spatial discrimination thresholds) also increases with distance from the midline (Mills, 1958). In Mills' experiment, listeners had to discriminate whether the second of two pure tones was located more to the right or to the left. The sounds were presented at a variety of azimuthal locations varying from the 0º (midline) to 90º (lateral), with 15º steps in between. The minimum audible angle for stimuli presented at 0 º in this study was shown to be about 1º, while for stimuli presented at 75º, it was about 7º. Increased acuity at the midline provides tentative evidence for an auditory equivalent of the visual fovea. An obvious question that logically follows is whether or not auditory spatial attention shows a gradient with eccentricity. This issue will be the focus of Experiment 3.

1.6 Relationship between BOLD signal and neuronal activity

FMRI is the method used to investigate selective attention in Experiments 1 and 2. Although this method has very good spatial resolution, it only indirectly measures neuronal activity. One challenge for neuroimaging is to explain the link between neural activity and the response that is measured during fMRI, the blood-oxygen level dependent (BOLD) signal. It is still not fully understood and one must be careful when making direct inferences between neural firing and hemodynamic responses.

One resolution element measured during fMRI contains millions of neurons. fMRI measures the hemodynamic response that is related to the metabolic demands of an active brain region rather than neural activity itself. Specifically, the BOLD signal reflects a complex combination of variations in cerebral blood flow, cerebral blood volume and oxygen consumption. Simultaneous fMRI and electrophysiological measures suggest that the BOLD signal is correlated most closely with sub-threshold input to, and local processing within, a region because both of these processes place demands on energy metabolism (Logothetis & Wandell, 2004). The BOLD signal is not necessarily correlated with the actual neural firing output of that region. Thus, single-unit recordings of action potentials and fMRI measure different aspects of neural activity. A similar perceptual process that is measured using the different techniques can give different results. fMRI should be viewed as a complementary technique that provides additional insight into the population activity that is not simply captured by the output action potentials.

Differences between single-unit recordings and fMRI are particularly notable when considering the effects of neural inhibition. Cortical activity is characterized by strong excitatory and inhibitory synaptic activity which sum to determine whether or not a given neuron fires. Since fMRI is very sensitive to such local sub-threshold processing, inhibition can either produce enhancement or suppression of the BOLD response. Nevertheless, there is no doubt that there *are* relationships between the BOLD response and neuronal activity. For example, an interesting recent finding is that negative BOLD response in brain regions is associated to some degree with decreases in neural response (Shmuel, Augath, Oeltermann, & Logothetis, 2006). The non-invasive nature and the high spatial resolution makes fMRI an ideal tool to study sensory cognitive processes in humans.

1.7 Neural mechanisms of selective attention

1.7.1 Attentional enhancement of cortical responses: Animal physiology

Selective attention can modify perceptual processing in a number of different ways; by increasing the amplitude of the neural response through an increase in the firing rate of the neuron (Luck, Chelazzi, Hillyard, & Desimone, 1997), or by sharpening and shifting the receptive field (RF) (Womelsdorf, Anton-Erxleben, Pieper, & Treue, 2006; Womelsdorf, Anton-Erxleben, & Treue, 2008). In this, and later sections, enhancement is taken to refer specifically to those circumstances in which selective attention increases the amplitude of the response to the attended stimulus. I make the assumption that an increase in the amplitude of the response (i.e. an increase in mean firing rate) can be detected using fMRI methods and will be reflected as a relative increase in BOLD signal.

The biased competition model of attention has provided an influential framework for studying the enhancement of sensory responses by selective attention (Desimone & Duncan, 1995). According to this model, when two objects are presented in the visual field, they have to compete for representation because the sensory system has limited processing capacity. Representation of the attended stimulus is positively biased by an increase in the firing rate of neurons that are sensitive to the attended stimulus. One way to represent the consequence of such bias is by an enhancement of the response to the attended feature or object.

One of the features of the biased competition model is a greater enhancement of the response for the attended stimulus in feature-specific cortical regions when two objects are presented simultaneously within the cell’s receptive field, compared with when two objects are presented sequentially or when only one object is in the receptive field and the other outside. In a single-unit study, Luck et al. (1997) recorded from neurons in two non-primary visual cortical regions (orientation-sensitive region V2 and color-sensitive region V4) of macaque monkeys. Stimuli had different colors and orientations so that an ‘effective’ stimulus had a particular orientation or color that was optimal for any given neuron in V2 and V4, respectively. There was a greater difference in firing rate between attending to the effective stimulus versus ignoring the effective stimulus when the two stimuli were presented simultaneously in the same receptive field (Figure 1.7A), compared to when they were presented sequentially (Figure 1.7B).

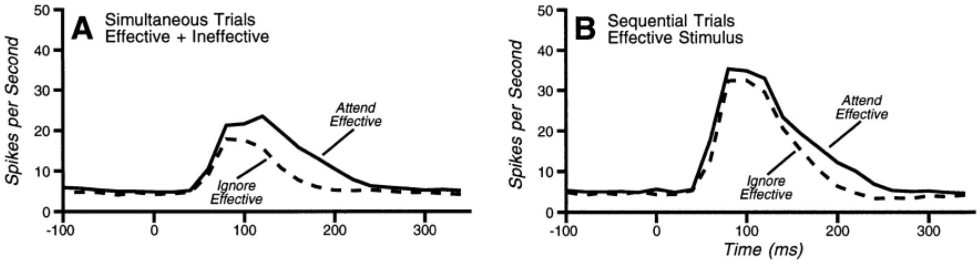


Figure 1.7 ***A****: post-stimulus average response for 29 V4 neurons that showed a significant difference between 'attend effective' (solid line) and 'ignore effective' (dashed line) stimulus conditions, when stimuli were presented simultaneously in the neuron's receptive field. B: post-stimulus average response for 14 V4 neurons that showed a significant difference between 'attend effective' (solid line) and 'ignore effective' (dashed line) stimulus conditions, when stimuli were presented sequentially in the neuron's receptive fields (taken from Luck et al., 1997).*

Additionally, Luck et al. examined the difference between enhancement that was due to a general elevation in spontaneous baseline activity or enhancement that was due to stimulus-specific factors (Figure 1.8) (Hillyard, Vogel, & Luck, 1998; Luck et al., 1997). The baseline activity was measured at 100 ms preceding the stimulus presentation. The enhancement of baseline activity results in fewer excitatory stimulus-driven inputs needed to signal the presence of that stimulus. Figure 1.8 presents 3 average histograms whereby the two stimuli were presented sequentially; one inside the receptive field and one outside (a and b), or both inside the receptive field (c). As Figure 1.8a shows, when comparing the conditions 'attend stimulus inside the receptive field' with 'ignore stimulus inside the receptive field and attend outside', there was a greater baseline response for the former condition. The response to the stimulus was identical for both conditions. On the other hand, in the time period that a stimulus was presented outside the receptive field of the neuron, attending inside the receptive field led to an enhancement of the baseline response compared to attending to that stimulus outside the receptive field (Figure 1.8b). No stimulus specific response was shown for either condition. As shown in Figure 1.8c when attending to a stimulus inside the receptive field, there was a greater stimulus-specific response compared to attending to an identical stimulus presented in another location in the same receptive field, when the two stimuli were presented sequentially. No enhancement of baseline activity was observed. So, when attending to a stimulus inside the receptive field, as opposed to attending outside of the receptive field, there was elevation of the baseline (spontaneous) activity of that neuron but no stimulus-driven attentional enhancement (a). Only when two stimuli were present inside the receptive field of the neuron was there stimulus-specific enhancement (c). The results indicate that both baseline and stimulus-driven enhancement coexist. The two effects could have very different consequences for fMRI measures of selective attention. Increases in baseline activity may decrease the relative difference in BOLD signal between baseline and stimulus conditions while stimulus-specific enhancement may increase the relative difference between baseline and stimulus conditions. As will be discussed in Section 1.8.2, some fMRI studies on humans have managed to differentiate between the two effects.

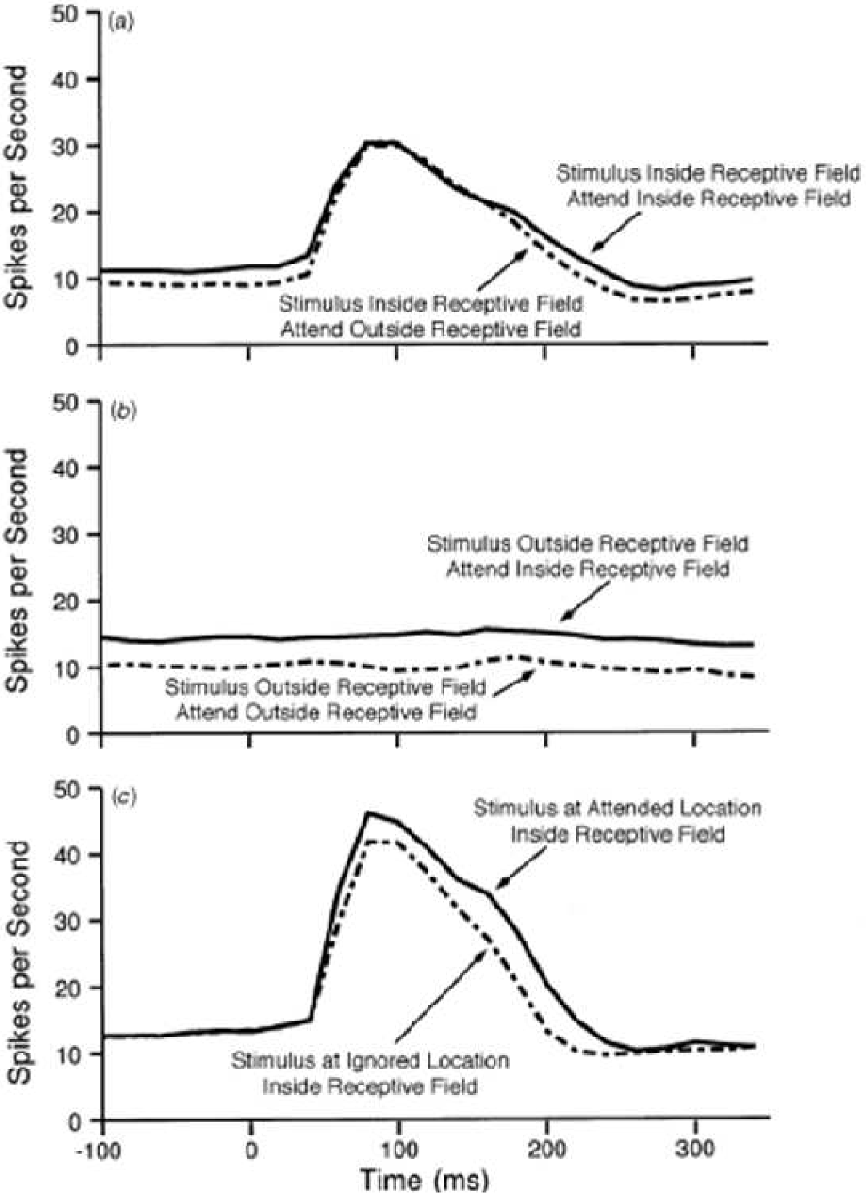


Figure 1.8 ***(a)*** *post-stimulus averaged responses for 74 V4 neurons for 'attend stimulus inside the receptive field' (solid line) versus 'attend outside the receptive field' (dashed line). In this condition, two identical stimuli were presented sequentially; one inside and one outside the receptive field.* ***(b)*** *post-stimulus averaged responses for 40 V4 neurons for 'attend inside the receptive field' (solid line) versus 'attend outside the receptive field' (dashed line), when one stimulus was presented outside the receptive field.* ***(c)*** *post-stimulus averaged responses for 37 V4 neurons for 'attend one stimulus inside the receptive field' (solid line) versus 'attend other stimulus inside the receptive field' (dashed line), when two identical stimuli were presented sequentially inside the receptive field. Note that '0' on the x axis is when the stimulus appears (Hillyard et al., 1998; Luck et al., 1997).*

In support of the biased competition model, the significant effects of selective attention were found only when two objects were presented within a neuron's receptive field. When only one object was in the receptive field, and another outside, there were no significant effects of (stimulus-driven) attention. This result agrees with previous studies (Haenny, Maunsell, & Schiller, 1988; McAdams & Maunsell, 1999; Moran & Desimone, 1985; Motter, 1993) suggesting that some sort of competition for resources is necessary. This effect is very significant for Experiment 1, because effectively, one auditory object (e.g. high-frequency sound) is presented in receptive field, and another auditory object (e.g. low-frequency sound) outside the receptive field. If attention functions in the same way in the auditory system as it does in the visual system, there should not be any effect of attention whatsoever. In fact, Luck et al., (1997) showed that there was no baseline or stimulus-specific enhancement in the primary visual cortical region V1 (i.e. more baseline activity when attending to stimulus inside the receptive field, than attending outside), although there was in V2 and V4 for the same (attended inside/unattended outside) conditions.

The studies discussed above, investigate the combined effects of feature and spatial attention, since the stimuli are usually in the stimulus' receptive field. However, other studies have managed to dissociate between feature-attention and spatial attention, by employing visual search tasks. Specifically, Bichot et al. (2005) presented monkeys with a cue, which indicated the color of the target. They recorded from V4 neurons that contained the target (e.g. a red stimulus), but to which the monkey had not yet made a saccade to. The recording was done when the monkey started making saccades to find the target. They showed that V4 neurons that contained the target (red object), and this color was the neuron's preferred color (but not when it was not preferred), there was a response enhancement, as well as synchronization in the gamma frequency range (30-60 Hz, see Section 1.7.4). In a second experiment, they investigated feature conjunction (color and shape). This time, monkeys had to search for a stimulus that had specific shape and color. The important result here was that the receptive fields of the neurons that contained distracters that were identical to the target in one stimulus dimension (e.g. only in color) showed enhancement and synchronization of firing, despite the fact that they did not contain the target. In a third experiment, they showed that when attending to the stimulus' receptive field, there is more enhancement than when not attending, i.e. location-specific enhancement. These experiments provide evidence for both feature-specific enhancement, implemented in a parallel manner across the visual space, and location-specific enhancement, implemented in a serial manner for the different attended locations.

Not all studies support the biased competition model of selective attention. The model predicts that some sort of competition for resources is necessary for attentional enhancement. However, feature-specific enhancement has been recorded in response to a single stimulus presented in the neuron's receptive field, in motion-sensitive region MT (e.g. Treue & Maunsell, 1996) and in color-sensitive region V4 (e.g. Spitzer, Desimone, & Moran, 1988). Spitzer et al. (1988) also showed that the more difficult the discrimination task, the larger the enhancement. This presumably indicates that, as long as the task is sufficiently difficult, the effects of attention can be shown even if only one object is presented within the neuron's receptive field.

Evidence for stimulus-specific enhancement of responses to attended sounds has also been provided in the auditory cortex (Fritz, Elhilali, David, & Shamma, 2007; Fritz, Shamma, Elhilali, & Klein, 2003). The experimental paradigm differed from those reviewed in the visual system. For example, attention was not directed to the most effective stimulus for that neuron and there was no competing stimulus (c.f. Luck et al. 1997). In a series of experiments, awake-behaving ferrets were trained to perform a number of frequency-based tasks. Unlike primates, it is difficult to train ferrets to switch attention from one target stimulus to another. A period of training on a tone-detection task focused attention on a target that was close to the neuron's best frequency (BF). The effect of attention was measured by mapping the spectro-temporal receptive fields (STRF) using a broadband rippled noise. The receptive fields showed strong facilitation around the target frequency and this persisted for 30-40 ms (Figure 1.9). This effect can be thought of as an enhancement of the response to the frequency that is attended. However, given that there is also a change in the shape of the STRF, aspects of the attentional modulation fit with the view that attention can also shift or expand the receptive field.

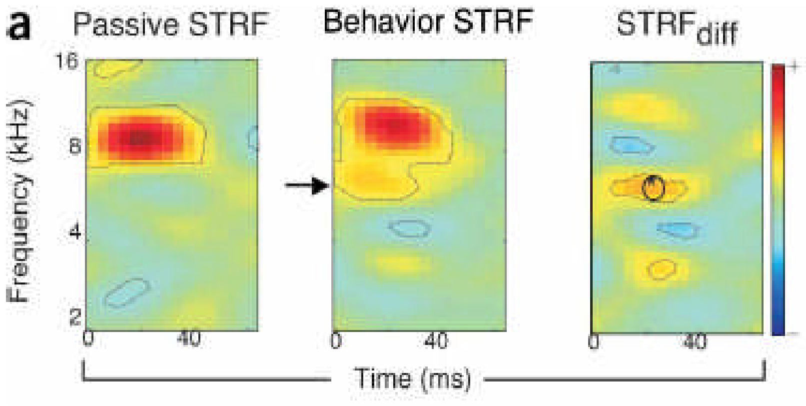


Figure 1.9 Effect of attention on the receptive fields of neurons in the primary auditory cortex of ferrets. The ferrets were performing a tone detection task (adapted from Fritz et al., 2003). In this particular task, they had to detect a tone of 6 KHz. The STRF difference plot on the right shows the result of a subtraction between passive and behavior STRF, shown on the left and centre.

1.7.2 Attentional enhancement of cortical responses: Computational models

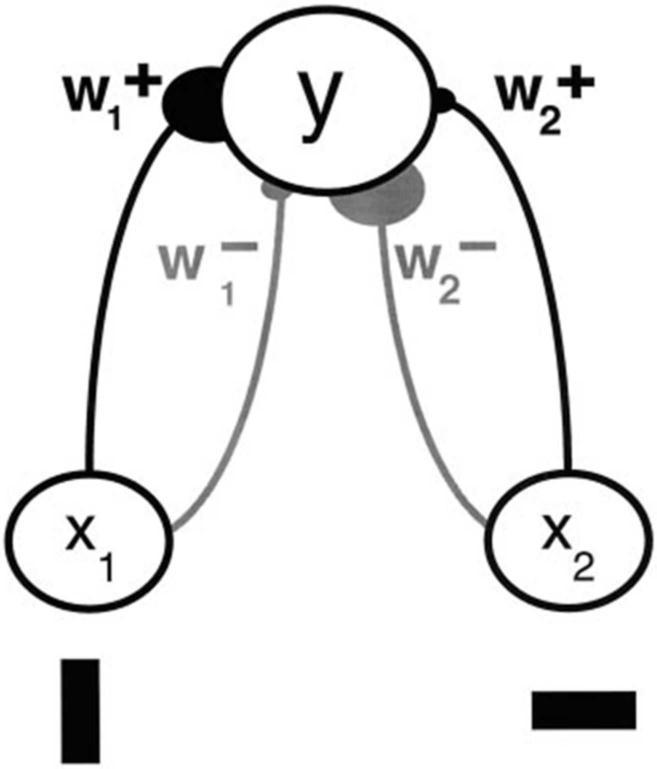
A number of different computational models of selective attention have been proposed to explain the physiological and behavioral data. These computational models are posed to explain the biased competition model of selective attention and they usually describe the situation in which two visual objects are presented within the neuron's receptive field or outside it.

Reynolds et al. (1999) demonstrated physiologically that attention is mediated by a biased competition mechanism, as Desimone and Duncan (1995) had proposed. They aimed to investigate the sensory response and the effects of attending to different stimuli for neurons in regions V2 and V4. Region V4 represents a color-sensitive region of the visual system receiving inputs from region V2. According to the biased competition model, when many stimuli are presented within one receptive field, they compete for resources and exert inhibitory influence on each other. Attending to one stimulus biases competition for this stimulus. In the first experiment, where they investigated the sensory response, they presented monkeys with a reference stimulus, with a probe stimulus, or with both of them, within the receptive field of the neuron while the monkeys were passively viewing. The authors compared the response of the neurons when only the reference was present with the condition where both reference and probe were present. The results showed that, when both reference and probe stimuli were present the response of the neuron changed according to whether the probe was a less or more effective (i.e. preferred) stimulus in terms of colour and orientation, compared to the reference. Specifically, if the probe was more effective than the reference, there was an enhancement of the neuronal response when both stimuli were present compared to when only reference was present. If the probe was less effective than the reference, there was a suppression of the response when both stimuli were present than when only the reference was present. Finally, if the reference and probe were equally effective, then there was no enhancement or suppression due to the probe.

In the second experiment, Reynolds et al. (1999) investigated the effect of attention, by directing attention either to two stimuli inside the receptive field of the neuron, or to two stimuli outside it. There were three ‘attend outside receptive field’ conditions, where the monkeys had to attend to two stimuli outside the receptive field, while, in the receptive field i) only the reference, ii) only the probe, iii) or both stimuli appeared. There were two ‘attend inside receptive field’ conditions, where the monkeys had to attend to the reference or the probe, when both stimuli were present inside the receptive field. When directing attention to one stimulus or the other within the receptive field, the neural firing rate depended on whether the stimulus that was attended to was effective or not for the neuron. For example, if the probe suppressed the response of the neuron, then attending to the reference reduced this suppression. If adding a probe enhanced the response of the neuron, then attending to the reference reduced this enhancement, while attending to the probe enhanced the response of the neuron even more. If the reference and probe were equally effective, then attending to either leads to a moderate increase in the neuronal response. The results of the selective attention task in Experiment 2 resemble those of sensory representation in Experiment 1. These results illustrate a close interaction between selectivity, sensory interaction and attention, by which biased competition is implemented. The authors note that this attentional modulation might be due to top-down bias from higher regions, biasing the neuron to respond preferentially to the attended location. These results support the biased competition model of selective attention.

Reynolds et al. (1999) developed a model using a simple neural circuit shown in Figure 1.9 to describe the results. In this model, selective attention increases the excitatory synaptic weights and decreases the inhibitory synaptic weights that input to those neurons that are tuned to the target stimulus or target feature. The computational model simply determines the firing rate of the neuron by the sum of excitatory and inhibitory inputs from projecting inter-neurons and so the resulting post-synaptic potential therefore encodes the attended target. As the figure shows, x_1 is a neuronal population sensitive to the attended stimulus, while x_2 is sensitive to the unattended stimulus. Attention enhances the strength of the excitatory input (w_1+) coming from cells that are sensitive to the attended stimulus. This simple circuit can explain neural outputs even in the presence of competing synaptic input from neurons that are not best tuned to the target. The authors speculate that this is achieved by increasing the efficiency of the synapses projecting from neurons that are best tuned to the ‘to be’ attended stimulus.

The model can predict the response when there is no competition within the receptive field. This is typically the case in V1 where neurons have very narrow receptive fields. The model predicts that if the response to the stimulus is not at its maximum, then it can be further enhanced by attention. The model can also account for a general increase in spontaneous baseline firing rate, as found by Luck et al. (1997). This increase occurs because attention increases the efficacy of the synapses from the afferent neurons that are sensitive to the attended location or stimulus. The greatest increase occurs if attention is directed to the centre of the receptive field.



$$E = x_1 w_1^+ + x_2 w_2^+ \qquad (1)$$

$$I = x_1 w_1^- + x_2 w_2^- \qquad (2)$$

Figure 1.10 Illustration of the model by Reynolds et al (1999). The big circle denotes the measured neuron and y denotes the firing rate of the neuron. The two smaller circles indicate populations of "input" neurons that are sensitive to either the reference (left) and probe (right) stimuli. These two neural populations, with mean firing rate x_1 and x_2 respectively, project to the measured neuron on the top. The black lines represent the excitatory projections from each of the input neural populations to the measured neuron. The gray lines represent the inhibitory projections to the measured neuron. w^+ represents the weight of the excitatory projection from the ith input population, and w^- represents the weight of the inhibitory projection from the input population (Reynolds et al., 1999).

A single-cell multi-compartmental model was proposed to describe how feed-forward input from the lateral geniculate nucleus (LGN) can influence the selectivity of a color-sensitive V4 neuron (Archie & Mel, 2000). In this model, the preferred stimulus was proposed to increase the number and the spatial proximity of synchronized excitatory synaptic inputs and therefore exert a greater influence on the overall sum of all pre-synaptic inputs (Figure 1.11).

Although this model primarily describes the sensory representation of visual input, it is possible that selective attention would also increase the feed-forward excitatory input to the attended stimulus. In this example (Figure 1.11), the neuron has four basal dendritic branches. The preferred stimulus elicits a greater amount of synchronous excitatory synaptic input along one of the four dendritic branches, while the non-preferred stimulus elicits a more dispersed pattern of input along the branches.

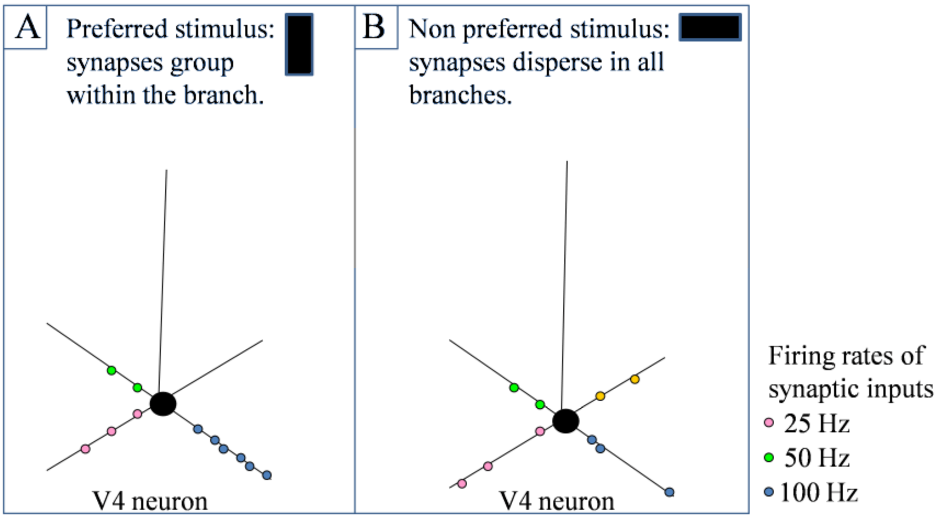


Figure 1.11 A single-cell multi-compartmental model by Archie & Mel (2000). According to this model, competition between stimuli is resolved by spatially separating the inputs that a neuron receives.

Other computational models test the prediction that attention is mediated by a top-down excitatory influence that modulates activity of neurons that respond to the attended stimulus (Deco, Pollatos, & Zihl, 2002) (Figure 1.12). When multiple stimuli are present, neurons that prefer one stimulus send inhibitory inputs to other neurons via the inhibitory pool that prefer another stimulus. Such reciprocal inhibitory activity captures the competitive nature of the biased competition model. When selectively attending to one particular stimulus, any neuron that prefers the attended stimulus receives additional excitatory input from higher-level centers which enhances the neural firing rate, and thus counteracts the inhibition from neurons that are sensitive to the non-preferred stimulus.

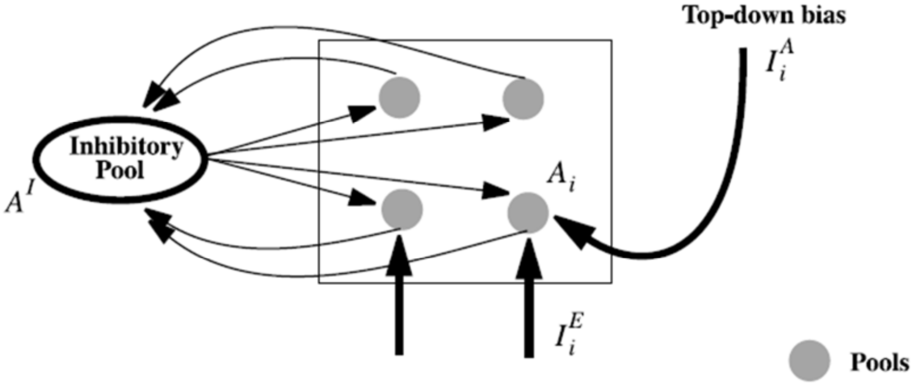


Figure 1.12 Basic module which forms part of the computational model by Deco et al.(2002). The module consists of a competitive network containing both the visual region, as well as external top-down bias. The excitatory pools are connected with a common inhibitory pool and exert mutual influence. This is how competition mechanism is implemented. Top-down excitatory input can bias the competition in favor of a particular pool.

It has been observed that neural firing can decrease relative to baseline when a non-preferred stimulus is attended (suppression), not just increase when the preferred stimulus is attended (enhancement). Mishra et al. (2006) proposed a mechanism to account for the negative and positive effects of selective attention on neural firing (Figure 1.13A). In their model of V4 neurons, the neuron has two sets of excitatory inputs (one for the preferred stimulus and one for the non-preferred stimulus) and one set of inhibitory inputs from inhibitory inter-neurons. The three sets of inputs are highly correlated over time, but the two sets of excitatory inputs are always out of phase. The crucial factor that determines neural firing is the phase of the inhibitory input with respect to the excitatory inputs. Firing output increases when the excitatory input for the *preferred* stimulus is out of phase with the inhibitory input (‘NP+Attend P’, Figure 1.13B). On the other hand, firing output decreases when the excitatory input for the *non-preferred* stimulus is out of phase with the inhibitory input (‘P+Attend NP’, Figure 1.13B). It is important to note that the model is optimal when the phases of the two excitatory inputs are clearly different, in other words the preferred and the non-preferred stimuli are very different from each other. Additionally, in the conditions whereby only one stimulus was presented at a time, attending to the preferred stimulus (‘Attend P’) showed a much greater firing rate than when only the non-preferred stimulus is attended (‘Attend NP’).

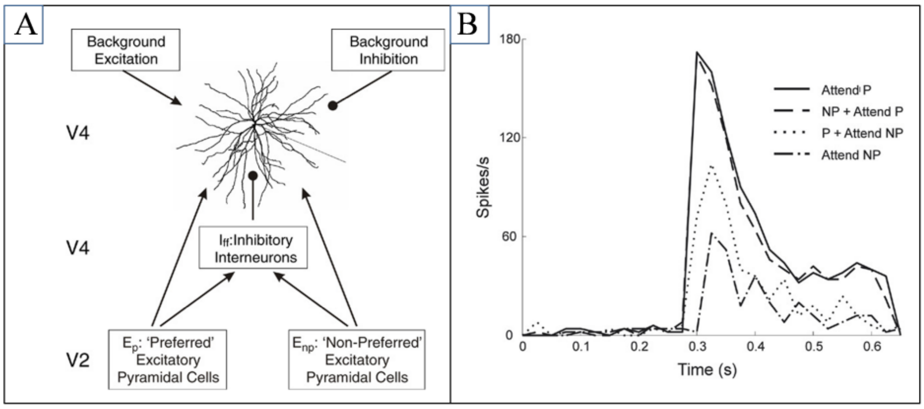


Figure 1.13 ***(A)*** *Illustration of the model by Mishra and colleagues (2006). According to this model, the V4 neuron receives excitatory and inhibitory inputs from V2 neurons. The model contains two excitatory synaptic inputs from two separate V2 neuronal populations, a ‘preferred’ and a ‘non-preferred’. As the authors note, the ‘preferred’ excitatory input (E_p) has more synapses than the ‘non-preferred’. Additionally the model contains one feed-forward inhibitory interneuron pool (I_{ff}) was used, the firing rate of which was a linear function of the firing rate of the active V2 excitatory neuronal populations. Furthermore, the V4 neuron received continuous uncorrelated background excitatory and inhibitory inputs. Arrows denote excitatory projections, circles denote inhibitory projections.* ***(B)*** *Firing rate of the V4 neuron for the different conditions. ‘P: preferred. NP: non-preferred.Attend P’ ‘Attend NP: one stimulus is presented alone and is attended. ‘NP + Attend P’, ‘P+Attend NP : both stimuli are presented but only one is attended.*

Corchs and Deco (2002) proposed a computational model to unify three sets of visual data; an fMRI dataset reported by Kastner et al. (1999) and two sets of single-unit recordings reported by Luck et al. (1997) and Reynolds et al. (1999). This computational model gives an idea of how the biased competition hypothesis might be implemented in the brain. This model is composed of many populations of neurons that are part of different brain regions and are interconnected. The model is composed of six modules, as shown in Figure 1.14. According to the biased competition model, competitive neural interactions occur at both microscopic and macroscopic levels. The model by Corchs and Deco consists of six parts: V1, V2-V4, IT, PP, v46, d46, organized in two streams: dorsal and ventral. Information from the lateral geniculate nucleus (LGN) enters V1, and then is directed into two separate streams: the ‘what’ pathway that consists of V2-V4 and IT, and the ‘where’ pathway that includes posterior parietal regions. According to this model, when multiple stimuli are presented at the same time their representations compete within the object recognition pathway and also possibly within the spatial location pathway if they are presented at the same point in space. V1 is responsible for extracting simple features of the objects. The IT region is responsible for object recognition. The regions V2-V4 are the intermediate stage between IT and V1, and they channel the responses from V1 to IT to make sure that there is little translation invariance. Finally, region v46 in the prefrontal cortex is responsible for keeping the object in the short-term memory. In the dorsal stream, PP is responsible for updating the position of the object that is attended and exerting influence on the earlier visual regions when selectively attending to space. Finally, d46 may be involved in keeping the location of the object in the short-term memory and in creating attentional bias for the location that is attended. Competition between representations creates inhibition between neurons. According to this model, the neuronal population of a region is connected with a common inhibitory neuronal input. Consequently, the more neuronal populations are active within a region, the more active the inhibitory inputs will be. Thus, only the very excited neuronal populations will survive the competition. Additionally, top-down modulation will provide excitatory feedback to a specific population, so that they are the winners of the competition. While this model predicts the neural consequences of attending to non-spatial features of competing objects, the model does not predict the role of attention in facilitating motion and location representations in V5/MT.

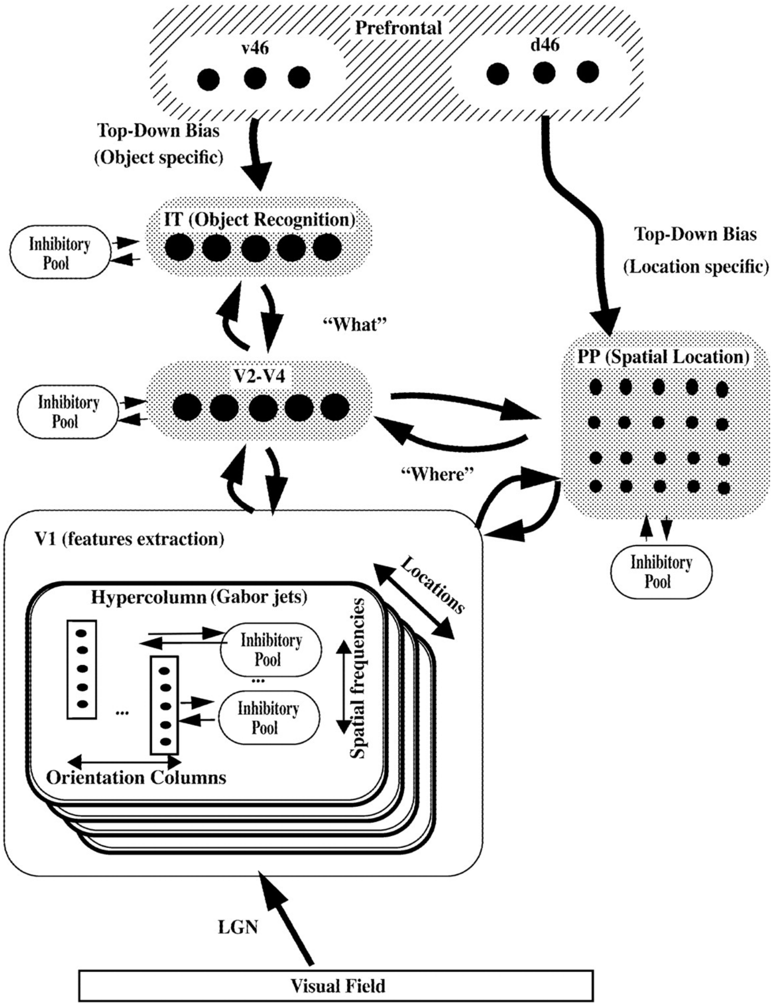


Figure 1.14 The neurodynamical model by Corchs and Deco (2002) to describe the mechanisms of selective attention in the ventral and dorsal visual pathways.

1.7.3 Attentional sharpening or shifting of cortical response properties: Animal physiology

Selective attention can have multiple effects on the receptive field of a neuron. As well as the increased neural firing described above as enhancement, neurons can also sharpen (i.e. narrow) in their tuning properties or can even shift their tuning properties. Sharpening and shifting of response properties tend to be reported for neurons in non-primary regions where they are not only more broadly tuned, and thus can be presented with more than one objects within their receptive field, but also demonstrate more complex tuning characteristics compared to neurons in primary regions.

Evidence for this effect have been reported by Womelsdorf et al.(2006; 2008) where they investigated spatial attention in motion-sensitive neurons in visual cortex. In these studies, macaque monkeys were trained to attend and respond to objects while neural activity in the motion-sensitive region of the cortex (MT) was recorded. Two of the objects, stimulus 1 (S1) and stimulus 2 (S2) were presented in the neuron’s receptive field, while the third, stimulus 3 (S3) was outside the receptive field. All objects contained random dot patterns that moved and created the percept of motion within the object. Selective attention was found to sharpen the receptive field around the attended location as well as shift the receptive field towards the attended location (Figure 1.15a and c). In Figure 1.15a, attending to S1 resulted in shift of the receptive field towards the S1 location. In Figure 1.15c, attending to S2 resulted in shift of the receptive field towards the S2 location. In Figure 1.15b, attending to S3 outside the receptive field, modifies the receptive field so that it is more equally distributed across the S1 and S2 objects. Moreover, there was a positive correlation between the sharpening and shifting effects. In other words, the more the neurons shifted their response properties, the more they sharpened.

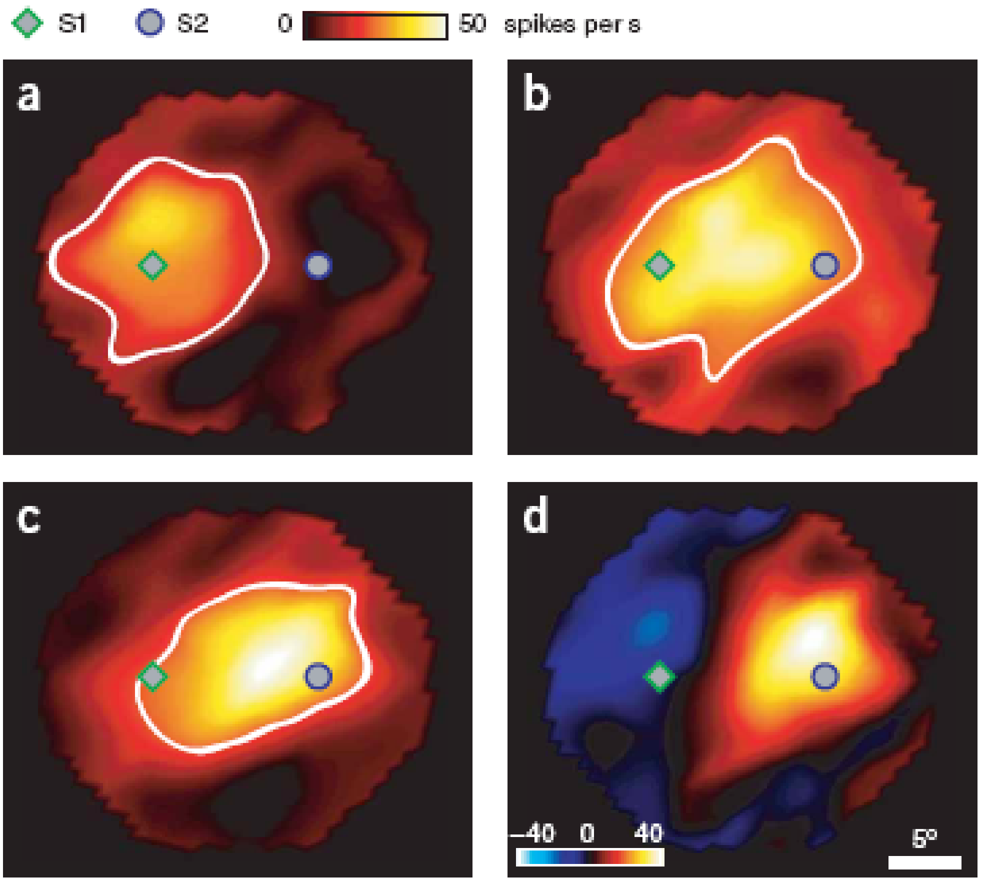


Figure 1.15 Shape of the receptive field of one neuron. (a) and (c): Shape of receptive field (RF) when attending inside the RF, to stimulus S1 (a) or S2 (c), or (b) when attention was directed outside the RF, to stimulus S3 . Color indicates the neuronal response increase in by the presentation of a probe stimulus at that location, over the response when only S1 and S2 were present. (d) Difference map, reflecting subtraction of the RF when attention was directed to stimulus S1 from the RF when attention was directed to stimulus S2.This map shows that shifting attention from S1 to S2 enhances response around S2 and reduces it near S1 (Womelsdorf et al., 2006).

Given the rather coarse spatial resolution of fMRI studies, the technique is rather insensitive to changes in the width of the receptive field or to shifts in the receptive field. Therefore human neuroimaging studies are typically concerned with quantifying whether the amplitude of the response increases or decreases as a function of selective attention.

1.7.4 Synchronization between and within regions as an attentional selection mechanism

There is evidence that selective attention does not only operate by increasing the firing rate of the neurons that are sensitive to the attended object. An additional mechanism has been proposed, whereby selective attention increases the synchronization of activity between regions and within regions in terms of gamma-wave oscillation. Gamma oscillations occur between 30-100 Hz (e.g. Fries, Reynolds, Rorie, & Desimone, 2001; Fries, Womelsdorf, Oostenveld, & Desimone, 2008).

Fries et al. (2001) recorded the local field potentials (LFPs) and single neuron spike rate from the color-sensitive V4 region of macaque monkeys, while they were either attending to a stimulus (target) inside the neuron’s receptive field or to a stimulus (distractor) outside the receptive field. There was an enhancement of gamma-wave synchronization for the ‘attend inside the receptive field’ condition relative to the ‘attend outside the receptive field’ condition. A more recent study by Fries et al. (2008) showed that gamma-wave synchronization was enhanced even when the monkey was expecting to see the stimulus in the receptive field, but the stimulus had not yet appeared. These studies show that gamma-wave synchronization is associated with attentional selection. Additionally, they prove that attentional modulation can occur, even if only one stimulus is present within a neuron’s receptive field, which is in contrast with the results of Luck et al. (1997).

Gamma-wave synchronization between neuronal populations of a visual region and a higher order, parietal region, is also linked to attentional selection (Saalmann, Pigarev, & Vidyasagar, 2007). In this study, macaques were trained to perform a visual-matching task whereby they had to match both the location and orientation of two stimuli presented one after the other with a 800 ms interstimulus interval (Figure 1.16).

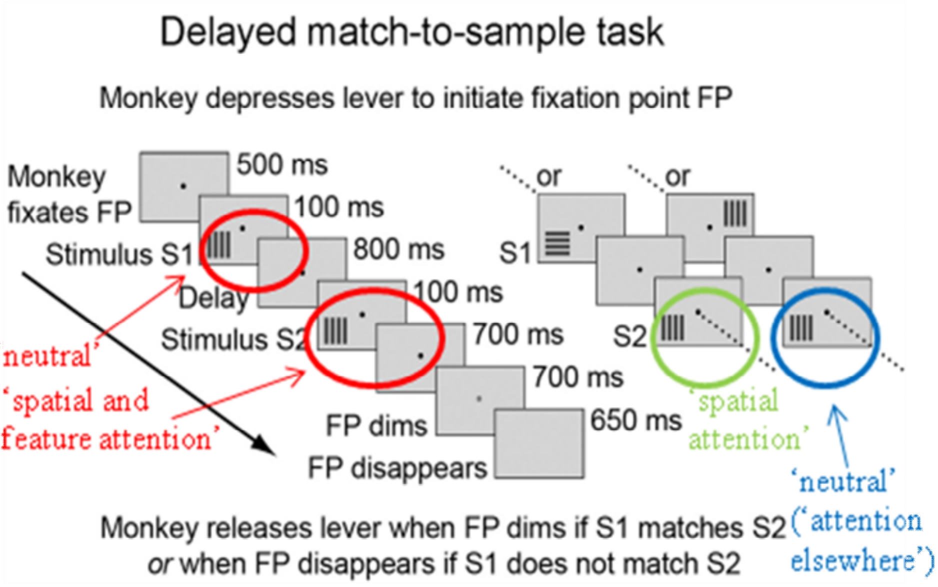


Figure 1.16 Illustration of the different experimental conditions and the matching tasks that monkeys had to perform in the study by Saalman et al. (2007). In this display, the stimulus on the bottom left of any grey quadrant is in the neuron's receptive field, while when presented anywhere else it is outside the neuron's RF. Additionally, when the lines of the stimulus are vertical it is considered a preferred stimulus. The circles indicate that point during the task that corresponds to each condition. For example, the red 'neutral' is the one immediately after the onset of S1. FP= fixation point. S1=stimulus 1. Adapted from Saalmann et al., (2007).

Responses were recorded from both lateral intraparatietal area (LIP) of the posterior parietal cortex and motion sensitive area MT (middle temporal). The coherence of the oscillatory spiking activity between MT and LIP was examined for the two neutral and attend conditions (Figure 1.16). Coherence between MT and LIP was greater for the two 'attend' conditions (red and green, Figure 1.17) compared to the two 'neutral' conditions (blue and black) in the frequency band between 25 and 35 Hz, (ie. a portion of the gamma frequency band).

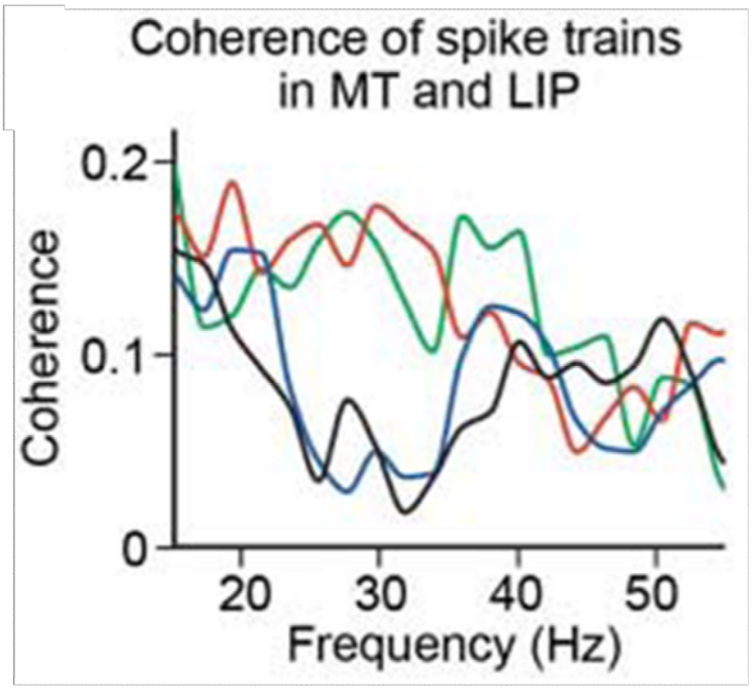


Figure 1.17 Coherence of neuronal response between MT and LIP for the two 'attend' and the two 'neutral' control conditions. Note that the 'neutral' condition coded in black here, was coded in red in the previous figure. The other conditions maintain their previous color coding. Red='spatial and feature attention'. Green=spatial attention, Black=neutral, Blue=neutral (attention elsewhere). Adapted from Saalmann et al. (2007).

The timing of activity between MT and LIP was also analyzed in terms of the number of spikes. The analysis for S1 and S2 showed that, when they were attending in a particular point in space, the peak in spike activity within LIP and MT occurred at a similar latency, but the peak in LIP occurred slightly earlier than in MT. The latency difference suggests that LIP exerts a top-down attentional influence onto MT.

1.7.5 Attentional suppression of cortical responses to irrelevant information: Animal physiology

Many studies of auditory attention highlight the evidence for attentional enhancement. However, selective attention might operate by dual mechanisms not only by enhancing relevant information but also by suppressing irrelevant information. In this, and later sections, suppression is taken to refer specifically to those circumstances in which selective attention decreases the amplitude of the response of neurons tuned to one stimulus, when that stimulus is no longer the focus of attention. I make the assumption that a decrease in the amplitude of the response (i.e. a decrease in mean firing rate) can be detected using fMRI methods and will be reflected as a relative decrease in BOLD signal. Note that this is the type of suppression examined in Chapter 4.

In the visual system, suppression has been shown when attending to a non-preferred stimulus (Treue & Trujillo, 1999). In this experiment, a non-preferred stimulus (pattern A) and a preferred stimulus (pattern B) were presented in the receptive field of an MT neuron. Tuning curves were plotted as a function of the pattern B’s motion orientation (Figure 1.18). When attending to a non-preferred stimulus (pattern A), there was a decreased response compared to attending to a preferred stimulus (pattern B), or ignoring both (the ‘sensory’ response) (Treue & Trujillo, 1999).

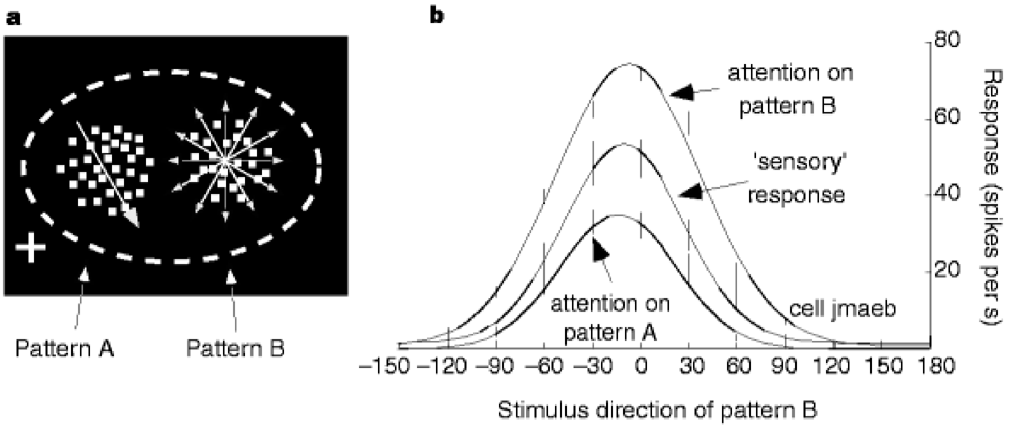


Figure 1.18 Effects of attentional enhancement (upper curve: 'attention to pattern B'), attentional suppression (lower curve: 'attention to pattern A') and sensory response (central curve). The curves are plotted as a function of the different direction of pattern B. pattern A always moves to a non-preffered direction of the neuron (Treue & Trujillo, 1999).

There is similar physiological evidence for the suppression of activity in neurons in the primary auditory cortex that are not tuned to the sound stimulus (Fritz, Elhilali, & Shamma, 2005). In a two-tone discrimination task, awake-behaving ferrets were trained to perform an oddball task in which they responded to an infrequent target frequency and withheld a response to a frequent reference frequency. Replicating the results from the tone-detection task, the receptive fields showed an enhanced response to the broadband rippled noise in a frequency region that was centered on the attended frequency (red arrow and corresponding red region of the STRF in right panel of Figure 1.19). In contrast, there was a suppressed response for the non-attended, reference frequency (blue arrow and corresponding blue region of the STRF in left panel of Figure 1.19). These opposing effects serve to magnify the contrast between the attended and the non-attended frequencies, and thus facilitate the selection of the target.

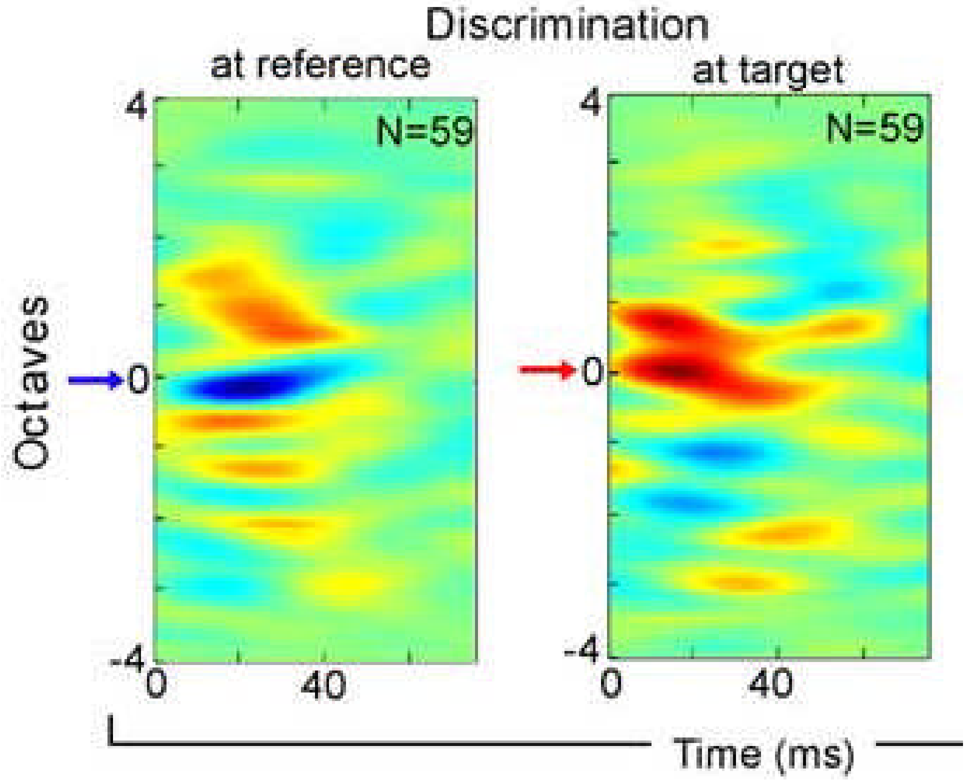


Figure 1.19   Summary of the effect of attention on the receptive fields of neurons in the primary auditory cortex of awake-behaving ferrets during a two-tone discrimination task. Blue colour indicates suppression of firing rates relative to baseline (green), while red indicates enhancement. Adapted from Fritz et al. (2005).

In other circumstances, suppression may refer to decreases in the amplitude of the response in neurons adjacent to those tuned to the attended stimulus. This latter case is generally known as lateral inhibition. A role for lateral inhibition has been incorporated into the biased competition model of selective attention (Desimone & Duncan, 1995). In this model, selective attention not only enhances the response of neurons tuned to the attended stimulus, but also reduces the suppressive interactions from neurons sensitive to nearby distractors. A role for lateral inhibition has also been incorporated into a number of other computational models. For example, in the model proposed by Deco et al. (2002), the inhibitory input to those neurons tuned to the attended stimulus from neurons tuned to the non-attended stimulus is effectively ‘cancelled out’ by top-down excitatory input. Although necessary for a complete understanding of the neural mechanisms of selective attention, the process of lateral inhibition cannot be isolated using fMRI measures of the BOLD signal and so it is not a focus in this thesis.

1.8 Human neuroimaging studies of selective attention

Many of the models support a top-down mechanism of selective attention, whereby when attending to locations, features or objects, frontal and parietal regions of the brain provide the top-down modulatory influence, the ‘sources’ of selective attention (Kastner & Ungerleider, 2000; Posner & Fan, 2004). Through anatomical connections, these act on sensory regions of the brain, the ‘sites’ of selective attention. In the next section, I describe evidence for this widespread attention network. Although many neuroimaging studies identify sources of attention, the evidence is often correlative rather than directional and so it is difficult to prove causality (Vuilleumier & Driver, 2007).

1.8.1 Sources of attentional modulation

Visual spatial attention The first part of this section is devoted to visual spatial attentional studies, because retinotopy is the primary organizational principle of the visual system and most studies investigate spatial attention.

Kastner and Ungerleider (2000) performed a meta-analysis of several blocked-design studies on visual selective attention. Bilateral inferior and superior parietal lobules, middle frontal gyri and frontal eye fields were consistently activated. Supplementary eye fields were also activated but less consistently across studies. There is evidence that these frontal and parietal regions do not appear to respond to an actual visual stimulus, but send feedback to visual regions. For example, Kastner et al. (1999), showed that there was no further enhancement of the response in the frontal and parietal regions when the stimulus was present, than when it was not present. Frontal eye fields are involved in more than voluntary saccadic eye movements. Evidence from both human (Corbetta et al., 1998) and physiological studies on monkeys (Wardak, Ibos, Duhamel, & Olivier, 2006) showed that frontal eye fields are involved in selective attention to visual locations, such as in visual search tasks as well as in discrete shifts of attention (Beauchamp, Petit, Ellmore, Ingeholm, & Haxby, 2001; Donner et al., 2000).

Selective attention is dynamic. There are many processes involved, such as directing attention, detecting a target, responding, as well as ignoring irrelevant targets. Thus, it is difficult to know the precise functional role of each brain region when averaging activity across a long block of time, as in blocked-design studies. This can be solved by using event-related fMRI studies that present participants with a cue and then present the target either in the same location (valid trial) or in different locations (invalid trial) (Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Hopfinger, Buonocore, & Mangun, 2000). Tootell & Hadjikhani (2000) summarized the results of these two studies (Figure.1.20). Response was more pronounced in the right hemisphere than the left. Superior and inferior parietal lobes and dorsolateral prefrontal cortex were typically activated during the spatial cue and these regions represent the control of spatial attention paradigm. Superior parietal lobe and supplementary motor cortex (including frontal eye fields) were typically activated during valid trials and represent the processing of the attended stimulus. Finally, superior parietal lobe, supplementary motor cortex and temporo-parietal junction (TPJ) are activated during invalid trials and represent the disengaging of attention from one location and redirecting attention to another unexpected target. The superior parietal lobule was active during all three events and so plays a crucial role in many facets of spatial attention.

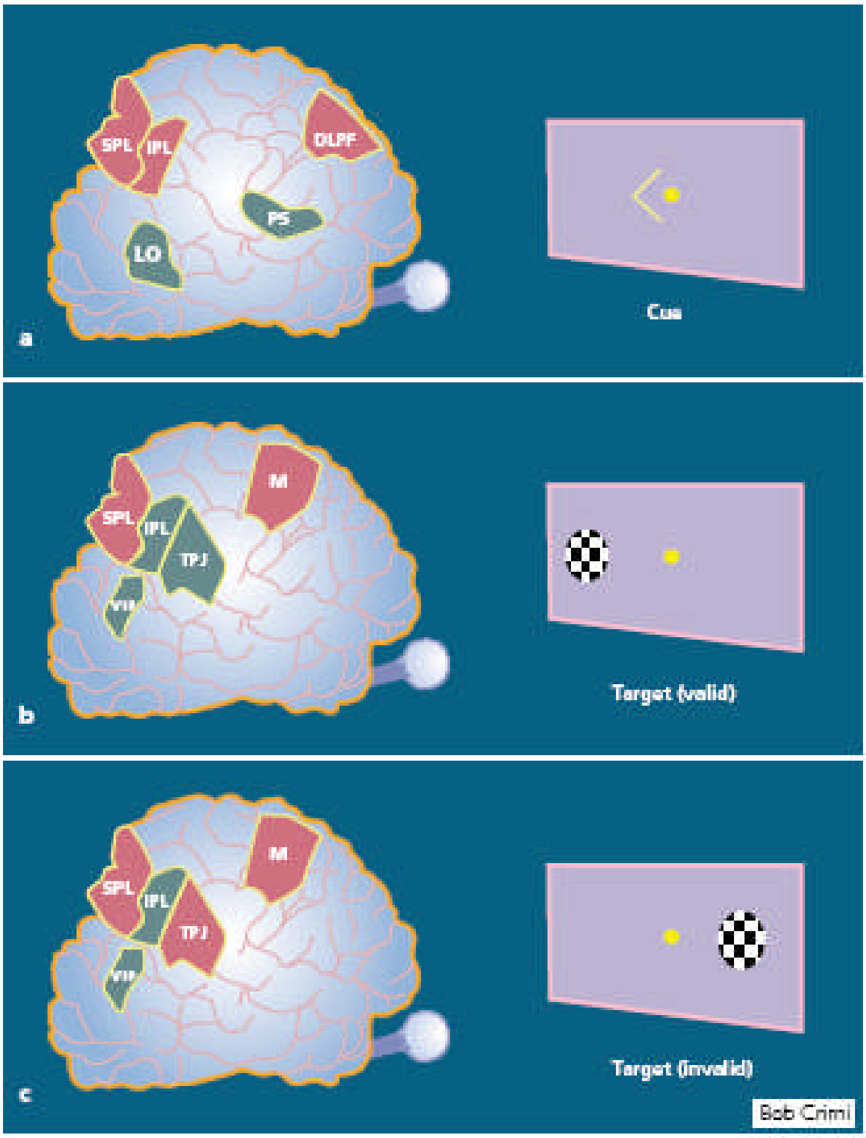


Figure.1.20 Regions activated when ***(a)*** *presented with a spatial cue* ***(b)*** *presented with a valid target (same location as cue) and* ***(c)*** *presented with an invalid target. The figures present a summary of the results from two studies by Hopfinger et al. (2000) and Corbetta et al.(2000) who used a spatial cueing paradigm. In red are the regions that showed high BOLD response, while in green are the regions that showed lower BOLD response. Note that other regions than the ones shown here were activated. DLPF=dorsolateral prefrontal cortex, IPL=inferior parietal lobule, LO=lateral occipital region of visual cortex, M =supplementary motor region, PS=peri-sylvian), SPL=superior parietal lobule, TPJ=temporal-parietal junction, VIP=ventral intraparietal region. (Tootell & Hadjikhani, 2000).*

Corbetta and Shulman (2002) refer to a segregated system for feature-based attention that distinguishes between ventral (inferior) and dorsal (superior) attentional source networks. The first appears to be involved in bottom-up modulation and involves TPJ, prefrontal cortex and inferior frontal gyrus. This network is proposed to play a role in switching attention, or when attention is grabbed by a salient stimulus. The second appears to be involved in top-down modulation and includes inferior parietal sulcus and frontal eye fields. This network is proposed to play a role in voluntarily attending to a stimulus.

Similar patterns of attention-related activity are reported when selective attention is directed to other visual features, such as color or motion (e.g. Shulman et al., 1999). Some of these regions also form part of a supramodal attentional system. Specifically, TPJ is involved in task switching in visual, auditory and tactile modalities (Corbetta & Shulman, 2002; Downar, Crawley, Mikulis, & Davis, 2001) (Figure 1.21).

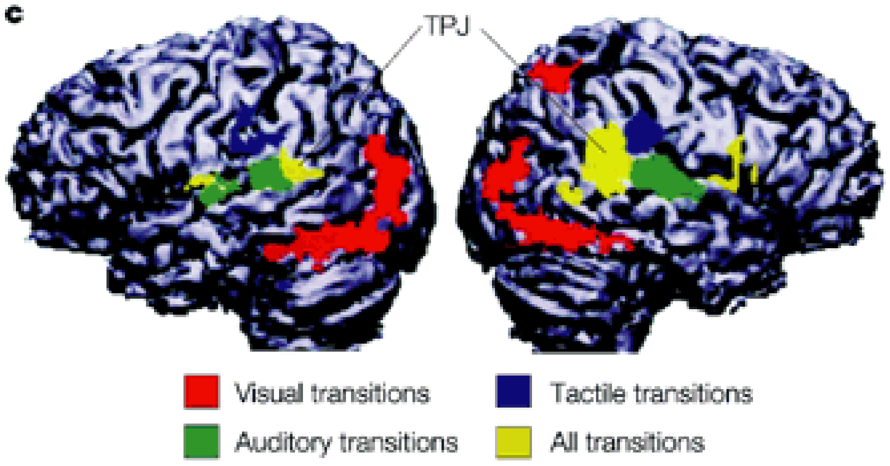


Figure 1.21 FMRI evidence that the TPJ responds to transitions in various stimulus modalities, such as auditory (transition from frog sound to running water sound), and visual (blue to red) (Corbetta & Shulman, 2002; Downar et al., 2001)

Selective attention and working memory are very much related mechanisms (Knudsen, 2007). For example, the participant has to hold the targets in working memory to be able to perform the task. The prefrontal cortex appears to be the most important region for this process because neurons show a persistent response, firing not only during the presence of a target, but until a response is made (Fuster & Alexander, 1971).

Knudsen has described a model of selective attention that combines working memory, competitive selection, attentional enhancement and suppression ('sensitivity control'), and filtering salient information (Knudsen, 2007). According to this model, working memory maintains the relevant targets and selective attention biases the response according to the behavioral goal (Figure 1.22).

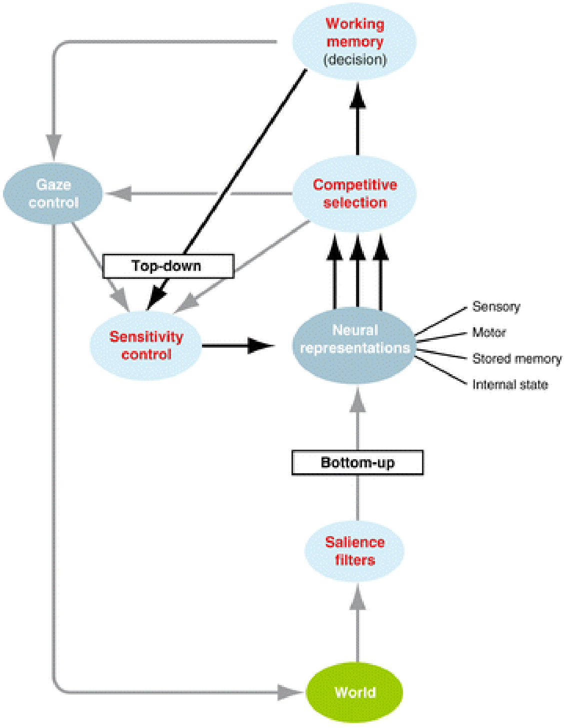


Figure 1.22 Model of selective attention (Knudsen, 2007). In red are the components that are relevant to selective attention. Salience filters refer to bottom-up (involuntary) processes, whereby infrequent stimuli are selected. Competitive selection selects the most salient stimulus which enters working memory, which can exert top-down influence which modulate sensory response. Sensitivity control refers to enhancement of the response for the attended feature, which is what is investigated in Experiments 1 and 2. The black arrows indicate top-down influence, while the grey arrows indicate bottom-up influence.

Auditory spatial attention Auditory spatial attention has been shown to engage the bilateral premotor cortex, left superior parietal lobe and right middle frontal gyrus (Degerman, Rinne, Salmi, Salonen, & Alho, 2006). Degerman et al. (2006) also showed that the right inferior parietal lobule was involved in both attending to the pitch as well as the ear of presentation. This result is consistent with the supramodal model of attention proposed by Corbetta and Shulman (2002) since the inferior parietal cortex forms part of the dorsal attentional source network. There is some evidence that source regions can be differentially involved in auditory selective attention depending on the nature of the task and the attended stimulus or feature (Degerman et al., 2006; Krumbholz et al., 2007).

1.8.2 Sites of attentional modulation. Evidence for feature-specific attentional modulation

The sites of attention refer to the modality-specific sensory regions that process the stimulus and these are modulated by source regions. Feature-specific attentional enhancement refers to the enhancement that occurs in the regions that are sensitive to the attended feature/stimulus.

Visual spatial attention In a seminal study, Tootell et al. (1998) showed an increase in visual activity in regions that were sensitive to the attended spatial location. The details of this study are described in Section 4.1 where they are most relevant to the design of Experiment 1. Feature-specific enhancement of the response has been shown with other visual features such as color and motion (Figure 1.23) (Schoenfeld et al., 2007). The general principles of feature-specific enhancement are also supported by evidence from animal electrophysiology discussed earlier in Section 1.7.

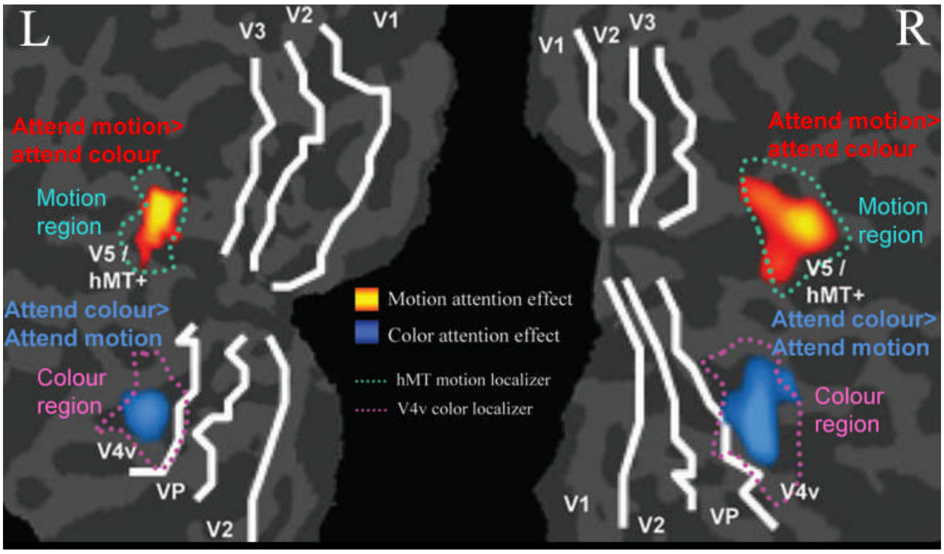


Figure 1.23 fMRI results of attention effects on flattened cortical surfaces of the 2 hemispheres for one participant. The red region denotes the response for the contrast 'attend motion>attend color', while response for the reverse contrast is shown in blue. Additionally, the motion-sensitive regions V5 (green dotted line) and the color sensitive region V4 (pink dotted line) were identified by contrasting all motion conditions with stationary/grey baseline, and all coloured conditions with the same baseline respectively. The borders between the visual areas were defined by retinotopic mapping. Adapted from Schoenfeld et al., (2007).

Auditory attention The question of interest in this thesis is whether auditory attention also operates by enhancing the response in a feature-specific way, either for sound frequency or for other features that are represented in specific regions of the auditory cortex, such as pitch, FM, and spatial motion.

There is little evidence for feature-specific attentional enhancement in the auditory cortex in humans. Petkov et al. (2004) were unable to find any evidence for feature-specific enhancement, but instead found that enhancement was widespread across the non-primary auditory cortex. The auditory stimuli were three pure tones of 350, 1400 or 4500 Hz, or three narrow band noise bursts centered at these frequencies. Stimuli were presented either in the right or in the left ear. There were also visual stimuli, which consisted of geometric shapes. Participants had to press the button whenever they heard a repetition of a tone or press the button when they saw a repetition of a particular shape. The authors first mapped the sensory response to sounds by subtracting the ‘attend visual’ condition (without sounds) from the ‘attend visual’ condition (with sounds). To investigate the effect of selective attention they subtracted the ‘attend visual’ condition (with sound) from the ‘attend auditory’ condition. Although the sensory response to sounds occurred in Heschl’s gyrus (HG, primary auditory cortex), attentional modulation occurred mainly in the superior temporal gyrus (STG, non-primary auditory cortex).

Two recent neuroimaging studies of the central auditory system do claim to support feature-specific auditory selective attention (Ahveninen et al., 2006; Krumbholz et al., 2007). This effect was specific to attending to spatial features. Ahveninen et al. (2006) found that in posterior auditory cortical regions, activity was significantly greater when attending to location than to phoneme identity. In the same regions, Krumbholz et al. (2007) also found activity was significantly greater when attending to motion changes than to pitch changes. In summary, although there is evidence for feature-specific attentional enhancement in the auditory cortex when attending to spatial features, there is no strong evidence for non-spatial features.

Attentional modulation in primary cortices. Does attentional modulation occur in primary sensory regions, or are these regions are only involved in sensory processing? In the visual system, there is growing evidence for the former (Posner & Gilbert, 1999), although the size of the modulation effect is of smaller magnitude than in higher-order regions (Tootell et al., 1998). This evidence comes from both physiological recordings (Motter, 1993; Roelfsema, Lamme, & Spekreijse, 1998) and human neuroimaging studies (Kamitani & Tong, 2006; Smith, Cotillon-Williams, & Williams, 2006; Somers, Dale, Seiffert, & Tootell, 1999; Tootell et al., 1998). As Posner and Gilbert (1999) indicate in their review, modulation effects in V1 depend on several factors, such as the nature of the stimuli, the difficulty of the task and the competition from nearby objects.

In the auditory system, there is evidence from physiology that attentional modulation occurs in the primary auditory cortex of ferrets and macaques (Brosch, Selezneva, & Scheich, 2005; Fritz, Elhilali, & Shamma, 2007). Although some human neuroimaging studies have also shown evidence for attentional enhancement in the primary auditory cortex (Jancke, Mirzazade, & Shah, 1999; Woldorff et al., 1993; Woodruff et al., 1996) other studies have shown that it occurs mainly in the non-primary auditory cortex (Ahveninen et al., 2006; Degerman et al., 2006; Petkov et al., 2004). Additionally, the studies that did show enhancement in the primary, did not demonstrate a *feature-specific* enhancement. There are many possible reasons why, beyond the fact that very few studies on humans on auditory attention have been conducted up to now. It is possible that the stimuli and tasks used were not appropriate (Posner & Gilbert, 1999). Imaging the primary auditory cortex is generally difficult (see Section 3.1) and the magnitude of attentional modulation in primary cortices is relatively small, and cannot be detected if sensory response is already high (i.e. close to the saturation of the BOLD signal). By using appropriate stimuli, paradigm and scanning parameters, one might be able to demonstrate attentional modulation in the primary auditory cortex. All these parameters are investigated in Chapter 3. In summary, although in the visual system there is evidence for attentional modulation in V1, it is still unclear whether attentional modulation occurs in the primary auditory cortex.

Enhancement of the baseline response As discussed earlier, Luck et al. (1997) showed evidence that baseline enhancement occurs while expecting for the stimulus to appear. Evidence for baseline enhancement when attending has also been shown in fMRI studies on humans (for a review see Driver & Frith, 2000). For example, Ress et al. (2000) showed that this baseline enhancement occurs in a feature-specific manner, and it is highly correlated with performance, as well as with task difficulty. Specifically, the more demanding the task, the better the performance and the higher the amplitude of the baseline enhancement.

The question of interest here is this: what is the relationship between baseline enhancement and stimulus-specific attentional enhancement? The first is measured when the participant is expecting the stimulus. The second is measured when attending to the stimulus, when the stimulus is actually present. There is evidence that baseline enhancement contributes to the stimulus-specific enhancement in an additive way (Buracas and Boynton 2007; Williford and Maunsell 2006). On the other hand, there is some contradictory evidence that baseline enhancement does *not* contribute to stimulus-driven attentional enhancement (Kastner et al., 1999; McMains, Fehd, Emmanouil, & Kastner, 2007). McMains and colleagues (2007) showed that although this baseline enhancement was location-specific, it was not feature-specific, while the stimulus-specific attentional enhancement *was* feature-specific. An auditory fMRI study by Voisin et al. (2006) was also successful in differentiating between baseline shift when listening attentively in silence (while expecting a tone to appear) and stimulus-driven enhancement. The former occurred in HG, while the latter in posterior and antero-lateral auditory cortex. However, their study was not designed to show whether this baseline enhancement was feature-specific. The design of Experiments 1 and 2 does not allow differentiating between baseline enhancement and stimulus-specific enhancement. However, it is important to keep in mind that there is a potential confound of a baseline enhancement.

1.8.3 Suppression of irrelevant/unattended information

In Section 1.7.5, neural suppression for the ignored frequency was discussed in the context of the auditory physiological study of Fritz et al. (2003). Several human fMRI studies show evidence for suppression of unattended information. In this event-related fMRI study (Slotnick, Schwarzbach, & Yantis, 2003), participants were presented with checkerboard stimuli in the inner (yellow), middle (green) and outer (blue) visual field and they were cued to covertly switch their attention to stimuli at the contralateral side, or to continue attending to a stimulus in the same side. First, as Figure 1.25 shows, there was location-specific enhancement (pink circle) in retinotopic areas sensitive to the attended location (indicated by the stars). Additionally, there was suppression (blue circles) in the surrounding retinotopic areas surrounding the areas, as well as in many regions of the (unattended at all times) lower hemifield ('dorsal' region, upper half of Figure 1.24).

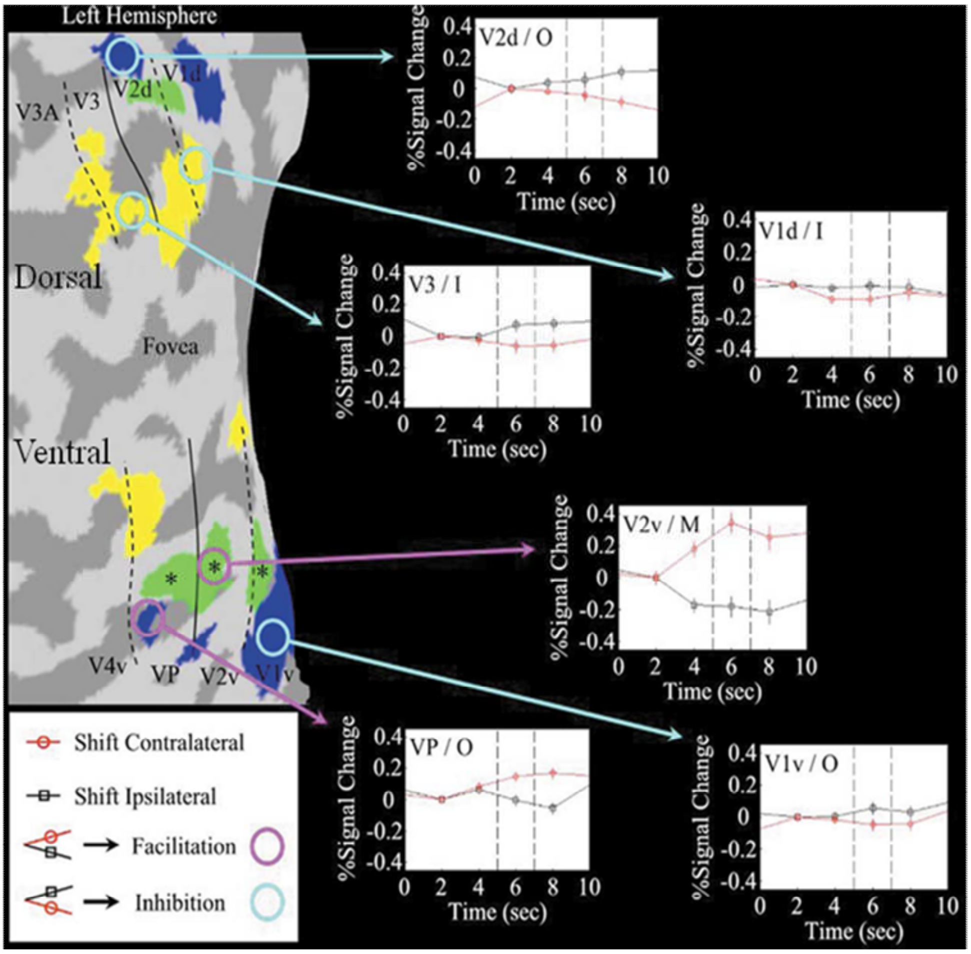


Figure 1.24 Attentional enhancement and suppression shown in the left hemisphere of a participant (Slotnick et al., 2003). The upper fields represent the dorsal (superior) visual regions. Lower ones represent the ventral (inferior) visual regions, which are sensitive to locations in the upper visual field, which is where the attention of the participants was always directed to. Yellow, green and blue colors represent sensory response for the stimuli in inner, middle and outer locations, respectively. The starts indicate the visual region sensitive to the attended location. Solid lines denote the right horizontal meridian, while dashed lines denote vertical meridia. The event-related activity for certain regions are shown on the right, for the two attend conditions 'shift ipsilateral' (blue) and 'shift contralateral' (red), and the effects of facilitation (red>blue) and inhibition (blue>red) are defined on the lower left box. The dotted lines on the graph indicate 6 s after the beginning of the event, where statistical analysis was performed.

A recent fMRI study showed that the demands of the attentional task are correlated with the suppression of response for the unattended stimulus (Sylvester et al., 2008). When participants expected a low-contrast stimulus (difficult task), there was more suppression in regions that are sensitive to the unattended stimulus than when expecting a high-contrast stimulus (easy task). Furthermore, there was more suppression in the unattended regions when participants responded correctly, than when they responded incorrectly for the low-contrast task. Conversely, enhancement of response was shown in frontal eye fields and posterior inferior frontal sulcus when expecting the low- than when expecting the high-contrast stimuli. The results suggest that, for demanding tasks, top-down control signals from frontal eye fields and posterior inferior frontal sulcus inhibit response for regions sensitive to the unattended information, thus enhancing processing of the attended stimulus.

It is possible that these fMRI studies measure the same mechanisms of suppression as reported from electrophysiology. However, given that neural inhibition requires metabolic energy, neural suppression could still be associated with a positive BOLD response, and so one should be cautious when linking the results from electrophysiology to the results from fMRI.

Finally, some visual studies report deactivation in supramodal regions, such as TPJ, during an active task when compared to a passive listening condition (Shulman, Astafiev, McAvoy, d'Avossa, & Corbetta, 2007; Shulman et al., 2003). It is hypothesized that this deactivation is linked to the filtering of irrelevant/unattended information.

1.8.4 Enhancement and sharpening

Studies on humans using electroencephalography (EEG) and magnetoencephalography (MEG) have shown that attending to pure tones masked by noise operates by two mechanisms: sharpening the receptive fields of neurons, and enhancement (gain) of the response (Kauramäki, Jääskeläinen, & Sams, 2007; Okamoto, Stracke, Wolters, Schmael, & Pantev, 2007). In an MEG study by Okamoto et al. (2007) participants had to attend to a tone in noise and detect a deviant tone or watch a silent movie and ignore the sounds. The noise was 8000-Hz low-pass filtered white noise with a frequency notch around the target tone. There were four ‘notch’ (band-eliminated-noise, BEN) conditions; 20, 40, 80 and 160 Hz bandwidth. When the notch was wide, the tone was expected to stimulate many new neurons. Therefore, the N100 response was expected to have high amplitude. Conversely, when the notch was narrow, many neurons sensitive to the tone would be responding to the noise and so the N100 amplitude was expected to decrease. The authors argued that if attending to the tone sharpened tuning, there should be no significant difference between the four different notch conditions when participants attended to the tone. In contrast, when ignoring sounds, the amplitude of the N100 should increase as a function of the notch bandwidth and these conditions were used for measuring the sensory representation of the sounds.

The results are plotted in Figure 1.25, as a function of the ratio between each band-eliminated noise width (BEN) over the condition where there was no noise (no-BEN). Response of the N100 amplitude was greater when attending to the tone than when ignoring it (black vs white circles respectively), which is evidence for attentional enhancement. The effect of sharpening of the RFs when attending to the tone is shown by the greater steepness of the slope for the ‘distracted’ condition than for the ‘attend’ condition and by the increasingly greater difference between attend-distracted conditions for narrower notches. As expected, the ‘ignore’ condition varied as a function of the notch, while the ‘attend’ condition varied much less. Therefore, these results provide evidence that there is both enhancement and sharpening.

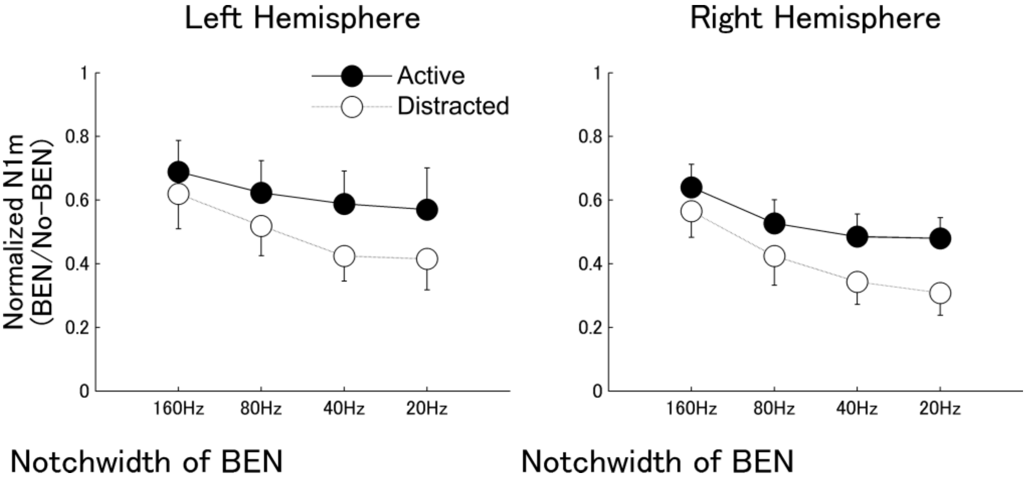


Figure 1.25 Group average normalized N100 amplitude for the left and right hemisphere, as a function of notch widths. Error bars represent 95% confidence intervals. Wight circles: ignore sounds. Black circle: attend sounds. BENs= band-eliminated noises. (adapted from Okamoto et al., 2007)

1.9 Summary

The mechanisms of selective attention in both visual and auditory systems were reviewed. There is much more research in visual attention, and this provides an inspiration and guide for auditory research. Selective attention appears to operate by many different mechanisms, such as sharpening of receptive fields, enhancement of firing rates, synchronization of firing between neurons of a regions and between regions. There is evidence that attentional modulation in the sensory regions is implemented by top-down bias from regions in the frontal and parietal cortices. In the visual system, there is evidence for location-specific attentional modulation in the primary sensory cortex. On the other hand, frequency-specific attentional modulation in the human primary auditory cortex has not clearly been demonstrated yet, and this is what will be investigated in Experiment 1. Experiment 2 attempts to confirm that spatial and non-spatial feature-specific attention occur in non-primary auditory cortex. Finally, Experiment 3 tests whether eccentricity influences the size of benefit from auditory spatial attention.

Chapter 2: Coding of sound features within the auditory system

2.1 Introduction

This chapter reviews the literature on the sensory representation of four key auditory features. The features of interest are sound frequency, FM, sound location and auditory motion. Evidence from human neuroimaging and animal electrophysiology will be discussed. These features were used in Experiments 1 and 2, described in subsequent chapters, to investigate feature-specific attentional modulation. The sensory representation of these stimulus features in the brain has previously been investigated and so the responses to these features in different auditory brain regions are relatively well characterized. This provided a necessary stepping stone for the studies in this thesis.

Relationship between the different neuroimaging measures of neural activity To investigate coding of auditory stimuli in humans, several techniques have been used; namely, EEG, MEG, PET and fMRI. Since each of these techniques measures a different aspect of neural activity, it is important to briefly describe how each technique works. For a description of fMRI see Section 1.6.

PET. Positron emission tomography is similar to fMRI in that it does not measure changes in neuronal activity directly, but rather measures changes in blood flow. In contrast to fMRI, PET is invasive. The participant is injected with a radioactive isotope of oxygen diluted in water ($H_2{}^{15}O$). This binds with glucose molecules in the blood stream and reaches the brain in 30 s, but takes 30 more seconds for the isotope to build up in a brain region. The isotope decays and the resulting signals are detected by coincident detectors placed around the participants' head.

One advantage of PET over fMRI is that it not noisy, and thus auditory stimuli can be presented in total silence. However, spatial specificity is not as good as in fMRI. This is because the signal that is detected by the coincident detectors has travelled by up to 3 mm from the actual locus of activation. Another reason for poor spatial resolution is that, to achieve statistical reliability, brain activation from many participants has to be averaged and the data have to be smoothed typically by 14 mm. The final resolution is about 18-20 mm, which is rather poor for spatiotopic mapping.

EEG/ERP and MEG. EEG directly measures the electrical activity produced by neuronal action potentials. The main origin of the EEG signal is the synchronous inhibitory and excitatory post-synaptic potentials of cortical pyramidal cells that are perpendicular to the scalp. Synchronous activity acts like a dipole, with a negative and positive potential at either end. Note that the dendrites are oriented towards the scalp and the positive electrical potential is outwards facing. These electrical potentials can be measured by placing electrodes on the scalp (Kandel, Schwartz, & Jessel, 1991). Electrical potentials can also be measured directly from the auditory cortex as part of a clinical evaluation for epilepsy (e.g. Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; Liégeois-Chauvel, Musolino, & Chauvel, 1991). EEG reveals both the spontaneous and the stimulus-evoked activity of the brain. Event-related potentials (ERP) refer to the response that is time-locked to a stimulus. In order to obtain a good signal-to-noise ratio, many ERP trials have to be averaged, since the EEG trace is very noisy and has much greater signal amplitude (20-40 µV) than the ERP amplitude (5-10 µV).

The source of the signal recorded by MEG is exactly the same as the one recorded with EEG. However, MEG measures the magnetic field created by the electrical activity in the brain. Very small changes in the magnetic field can be measured by the many detectors placed around the head of the participant.

There are two things that researchers measure when using EEG and MEG techniques. One is the timing of the neural response, and the other is the location of the neural activation. The particular strength of EEG and MEG lies in their excellent temporal resolution (in the order of ms). The location of the response is often computed as a current dipole. Similar to the notion of each neuron representing a dipole, with negative potential on one side and positive potential on the other side, a population of neurons can be modeled as a single dipole too. However, note that dipole modeling relies on many assumptions and activity of many brain regions might contribute to the signal recorded. In fact there is an infinite number of ‘inverse solutions’ that can be fitted to explain the source of the response. For EEG it is difficult to determine the precise locus of the response since the scalp and the cerebro-spinal fluid blur the distribution electric activity on the scalp. MEG has the advantage that it is not influenced by such effects of volume conductance and thus it has much better spatial resolution than EEG. The highest spatial resolution that can be reached is about 70 mm for EEG and 7 mm for MEG (Huettel, Song, & McCArthy, 2004). However, this resolution is not adequate for investigating tonotopy in the human primary auditory cortex because the resolution is not good enough to make distinctions within the tonotopic fields.

2.2 Frequency coding

Frequency is the most fundamental sound feature for the auditory system. Some researchers claim that frequency is to audition what space is to vision (Kubovy & Van Valkenburg, 2001). This is because while the retina and the primary visual cortex are organized spatiotopically, the cochlea and primary auditory cortex are organized tonotopically. Humans are exquisitely sensitive to sound frequency. Using frequency-discrimination paradigms, it has been shown that normally hearing listeners are able to discriminate between two tones that are separated by only a few Hz (Sek & Moore, 1995; Wier, Jesteadt, & Green, 1977). For example, for a pure tone of 1000 Hz at 60-70 dB, the minimum difference that can be detected is about 3 Hz.

2.2.1 Frequency coding in the cochlea

The simplest sound wave is a pure tone, for which pressure varies sinusoidally over time. The frequency of a pure tone refers to the number of the cycles of the pure tone that happen during a particular length of time. For example, if one cycle occurs within 1 s, then this sound is 1 Hz (1 cycle/1s=1 Hz). The groundbreaking work of von Békésy (1960) revealed how the cochlea responds to sounds of different frequency. Variations in pressure that occur in the fluid of the cochlea induce displacements in the basilar membrane, which appear to have a wave-like motion. The place on the basilar membrane at which this displacement peaks in amplitude depends on the frequency of the sound (Figure 2.1). For high-frequency sounds, the peak occurs towards the base and for low-frequency sounds it occurs towards the apex. In other words, each place in the basilar membrane is tuned to a particular characteristic frequency, or 'best' frequency (BF). Additionally, pure tones that are near in frequency to each other make neighboring places in the basilar membrane show peak amplitude at nearby regions of the basilar membrane. That is, the basilar membrane is organized tonotopically.

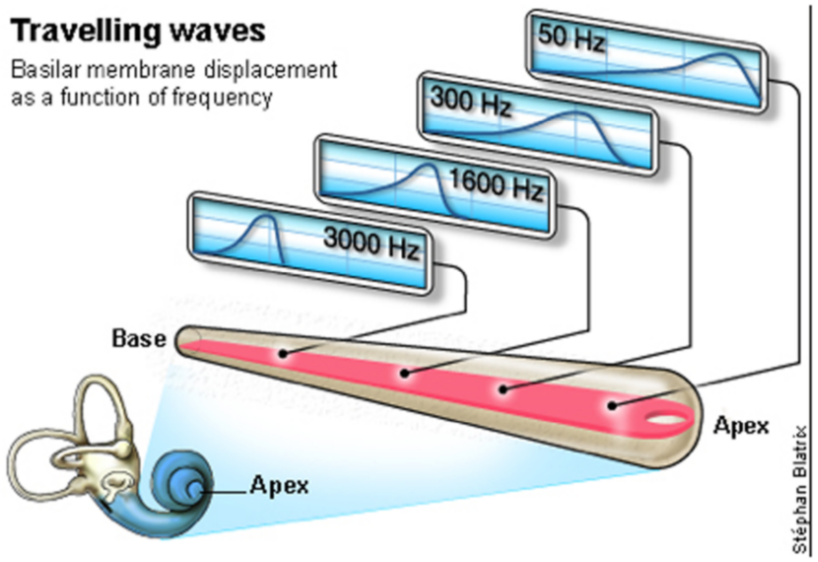


Figure 2.1 Illustration of the tonotopic organization of the basilar membrane. Source: http://iua share.upf.es/wikis/seminaris/images/1/12/RS_Hendrik_3.png

After the sound has caused the basilar membrane to vibrate in particular places, the inner ear translates the mechanical vibrations of the basilar membrane into electrical activity in the auditory nerve. The inner hair cells within the Organ of Corti are responsible for this process (Pickles, 2008). Specifically, when the basilar membrane vibrates, it makes the tectorial membrane above it to vibrate as well. This causes the stereocilia on the tops of the inner hair cells to bend from side to side at the same rate as the vibration of basilar membrane. This bending of the stereocilia opens ion channels through which flow ions, causing the voltage across the cell to depolarize (i.e. there is an increase in potential inside the cell relative to outside the cell). This causes the discharge of a chemical neurotransmitter, which creates electrical activity in neighboring auditory nerve fibers.

The auditory nerve consists of many nerve fibers. Each hair cell is contacted with 10 to 30 of the auditory nerve fibers (Pickles, 2008). Since each auditory nerve fiber is connected to an inner hair-cell at a certain place along the basilar membrane, it represents the response to the sound at that location. Therefore, each auditory nerve fiber has a best frequency. This type of coding is known as the place coding of frequency information because there is a systematic spatial arrangement of tuning both along the basilar membrane and within the auditory nerve fiber bundle. The nerve fibers that are near the centre of the auditory nerve have a low best frequency, while fibers at the periphery have a high BF (Pickles, 2008). This kind of tonotopic organization is preserved throughout the ascending auditory system up to the auditory cortex. Specifically, research over a range of mammals showed tonotopic organization in cochlear nucleus (Bourk, Mielcarz, & Norris, 1981), superior olive (Tsuchitani, 1977), inferior colliculus (Malmierca et al., 2008), in the ventral part of medial geniculate body (Anderson, Wallace, & Palmer, 2007) and in the auditory cortex (e.g. Merzenich & Brugge, 1973).

Additionally to place coding, frequency is represented in the phase locking. When the basilar membrane vibrates due to a low-frequency sound, the hair cells will depolarize only when their stereocilia bend towards a particular direction. That is, the hair cells depolarize only at a particular phase of the waveform. This indicates that frequency is coded by the timing of auditory nerve activity. This type of code cannot be measured by fMRI or PET, but synchronous oscillatory activity can be detected using EEG and MEG.

2.2.2 Frequency coding in the auditory cortex: Primate studies

Non-human primate auditory cortex has perhaps the greatest homology with human auditory cortex and so where a lot of evidence is available on primates, the review will focus on this research rather than other mammalian species.

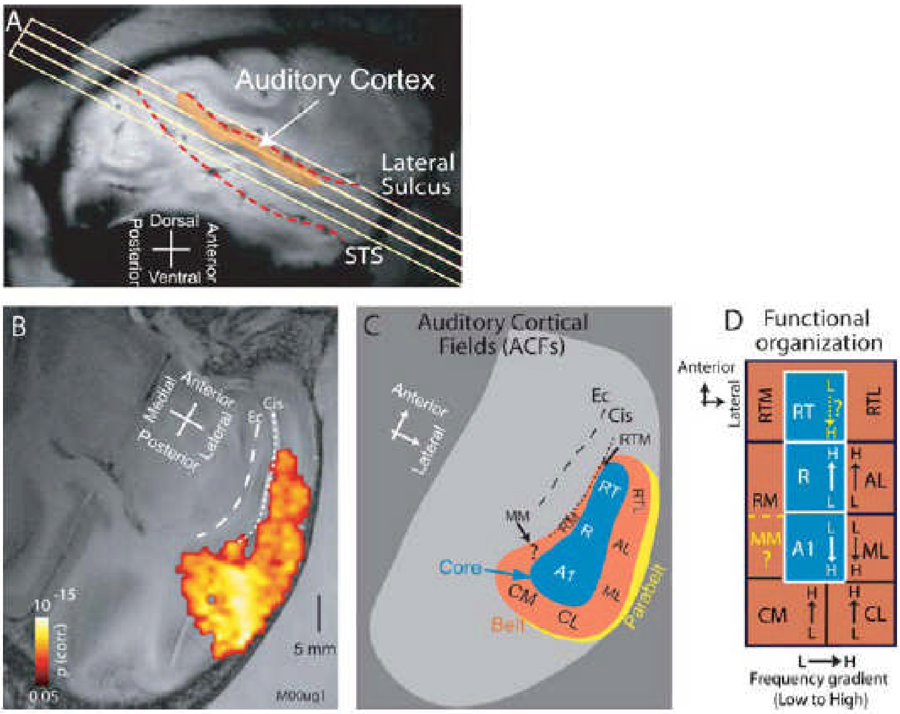


Figure 2.2 Tonotopic organization of the auditory fields in macaque monkeys. (A) sagittal view of the brain of a macaque monkey (STS: superior temporal sulcus). (B) fMRI response to broadband noise from one of the anesthetised animals (C) Primary (core) fields can be viewed in the middle (blue), surrounded by the belt region (orange). (D) functional organization of the core and belt auditory fields. H:high frequency, L: low frequency, RT: rostrotemporal, R: rostral, C: caudal, M:medial, L: lateral. Adapted from Petkov, Kayser, Augath, & Logothetis (2006).

Within the primate auditory cortex, three anatomical subdivisions have been identified: the core, the belt and the parabelt fields, each corresponding to three levels of auditory cortical processing (for a review see Hackett, 2003).

Figure 2.2 shows the different auditory fields across the auditory cortex. These are results from an fMRI study on macaques (Petkov et al., 2006). The core is the primary auditory region in primates, which has three subdivisions: the most caudal A1 which is the largest primary region, a rostral region R, and a rostrotemporal region RT. The core is referred to as ‘koniocortex’ due to the high density of cells in layers II and IV. The rostrotemporal region exhibits characteristics of both primary and non-primary regions, and it is debatable whether it is a primary or a non-primary region. The belt region appears to have eight fields: four medial: rostrotemporal medial (RTM), rostromedial (RM), medio-medial (MM), caudomedial (CM) and four lateral: caudolateral (CL), mediolateral (ML), anterolateral (AL), rostrotemporal lateral (RTL). The parabelt is thought to have two fields: rostral parabelt (RP) and caudal parabelt (CP).

The core, belt and parabelt are distinguished anatomically by several factors (Hackett, 2003). Firstly, the core regions have a high density of myelin, which is reduced in the belt, and even more reduced in parabelt. Secondly, the enzymes acetylcholinesterase and cytochrome oxidase, which are indicative of metabolic demands of brain tissue, have high expression in the core, intermediate in the belt, and low in the parabelt. Thirdly, the core has no pyramidal cells in layer III and only a few in layer V, while the belt has large pyramidal cells. The parabelt has also large pyramidal cells, which are more uniform in size than the belt and resemble organ pipes because they are places next to each other in columns.

In terms of the tonotopic organization of these fields, electrophysiological recordings and functional MRI in mammals have revealed many tonotopic maps (Kosaki, Hashikawa, He, & Jones, 1997; Merzenich & Brugge, 1973; Petkov et al., 2006). Within each map, neurons tuned to the same sound frequency are co-localized in a strip across the cortical surface, with an orderly progression of frequency tuning across adjacent strips. Frequency tuning is sharper in the primary auditory fields than it is in the surrounding non-primary fields, and so the most complete representations of the audible frequency range are found in the primary fields. Neurons in the primary auditory cortex are responsive to pure tones, while non-primary neurons are more responsive to more spectrally complex stimuli, such as broadband noise (Rauschecker, Tian, & Hauser, 1995). It is possible to demonstrate tonotopy by fMRI as well as by electrophysiology, even though frequency selectivity deteriorates at the moderate to high sound intensities required for fMRI sound presentation (Phillips, Semple, Calford, & Kitzes, 1994). As a recent example, mirror-symmetric frequency gradients have been confirmed across primary auditory fields using high-resolution fMRI at 4.7 and 7 Tesla in six anesthetized macaques and at 7 Tesla in one awake behaving macaque (Petkov et al., 2006). The results of this study confirmed and extended previous electrophysiological findings, also in macaques (Kaas & Hackett, 2000). As shown in Figure 2.2D, each of the three primary fields contains a tonotopic map, whose axis shows mirror symmetry with the tonotopic map of the adjacent fields. In the belt, four of the eight non-primary fields showed a clear tonotopic organization. Again, each of the four fields contains a tonotopic map, each of which is the mirror image of the adjacent field (and thus shares a common low- or high-frequency border).

2.2.3 Frequency coding in the auditory cortex: Human studies

The relation between anatomical and functional results is less straight forward than in animal studies, as most human tonotopic studies are non-invasive and so do not identify individual anatomical boundaries. Thus, one can only speculate about the structure-function correspondence with reference to probabilistic maps of different anatomical fields.

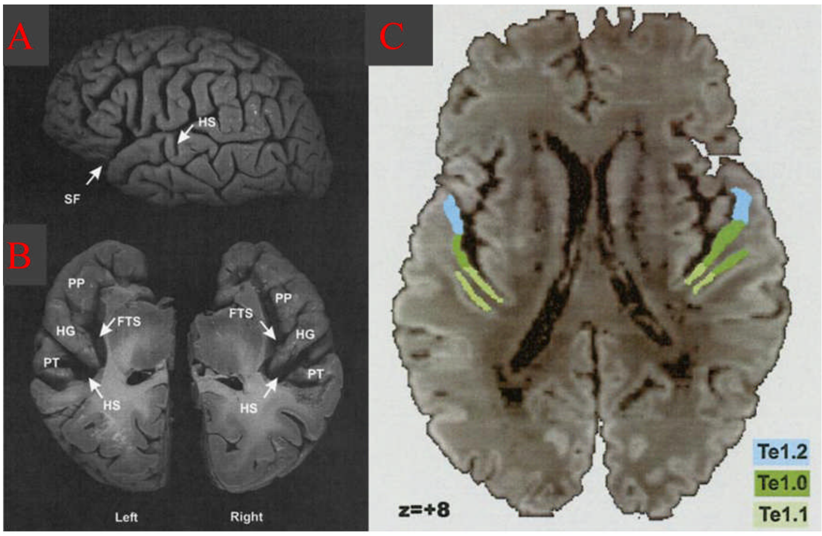


Figure 2.3 Anatomical subdivisions of the primary auditory cortex (Morosan et al., 2001). (A) Sagittal view of the brain (B) Oblique slice of one post-mortem brain showing the different auditory regions. (C) Axial anatomical slice showing the three different anatomical subdivisions of HG; Te 1.1 (most medial), Te 1.0 (central), Te 1.2 (lateral). FTS=first transverse sulcus;

Figure 2.3 shows an oblique cross section of an individual human brain post-mortem (Morosan et al., 2001). The primary auditory cortex is situated in Heschl’s gyrus (HG), defined by Heschl’s sulcus (HS) posteriorly and the first transverse temporal sulcus anteriorly (Figure 2.3B). HG has an anterior-lateral to posterior-medial orientation, oblique to the lateral surface of the brain (Figure 2.3B). The volume of the primary auditory cortex is between 1 and 4 cm^3 in each hemisphere (Penhune, Zatorre, Macdonald, & Evans, 1996). The non-primary auditory cortex is situated posteriorly and anteriorly of HG; including planum temporale (PT) and planum polare (PP) respectively. According to Morosan and colleagues (2001), Nissl staining showed that the primary auditory cortex is anatomically divided in three parts: medial (Te1.1), central (Te 1.0) and lateral parts (Te 1.2) (Figure 2.3C). On the other hand, Wallace and colleagues (2002) found a completely different scheme in primary auditory cortex. They found two primary regions, the larger one located in medial and central HG, while the smaller one just behind this region, along HG. Furthermore, Wallace and colleagues (2002) provided evidence that the lateral part of HG (Te 1.2) is actually a non-primary auditory region. Although their results from Nissl staining agreed with those of Morosan et al. (2001), other measures showed different results. Specifically, lateral HG did not show a high density band of high cytochrome oxidase activity, as in the primary cortex. Instead, it contained a relatively high-density of pyramidal cells containing achetylcholine, similarly to non-primary auditory regions. However, since different measures qualify this area as either primary or non-primary, it is possible that this region is actually a transitional region. In fact, the results for this region resemble the rostrotemporal (RT) field found in monkeys, which also seems to have mixed cytoarchitectonic characteristics.

In this thesis, Morosan's scheme will be adopted; with the exception that region Te 1.2 will be considered as a non-primary region. The anatomical scheme reported by Wallace is perhaps more reliable, because they used a variety of methods to stain for anatomical and metabolic markers. On the other hand, Morosan's scheme has been incorporated into probability maps that have been transformed into the same standard brain space as fMRI data and are therefore easy to use (Eickhoff et al., 2005).

Evidence that lateral HG (Te 1.2) is not a primary auditory region, also comes from direct recordings from patients suffering from epilepsy (Liégeois-Chauvel et al., 1994). Electrodes implanted in their superior temporal gyrus and nearby brain regions of the non-dominant hemisphere, helped not only to determine which parts of the brain are responsible for their seizures, but also to investigate the structure and function of the auditory cortex. In an earlier study, Liégeois-Chauvel and colleagues (1991) had shown that the response amplitude to sound clicks was greater when recording from posteromedial HG, compared to central HG. The amplitude of the response decreased when recording from lateral HG. In a later study, Liégeois-Chauvel and colleagues (1994) demonstrated that a different portion of HG showed particular latency potentials in response to a sound of 1000 Hz. That is, the lateral part showed a latency of 60-75 ms, the central part, 50 ms, and the posteromedial part, 13-30 ms. These results support the notion that only the posteromedial part of HG coincides with the primary auditory cortex, because shorter latencies are expected from the primary region, due to the more direct connection of this region to the auditory periphery. Note that these results limit the primary auditory cortex to an even smaller region than the scheme proposed by Wallace and colleagues.

The rare studies of direct recording from the human auditory cortex have not only been informative of its structure as shown above, but also of its tonotopic organization. Howard et al., (1996) recorded from multi-unit depth electrode that was inserted along the axis of HG of one patient. The patient listened to 24 pure tones ranging in frequency from 200 to 10000 Hz. By determining the BF of each of the electrodes, they showed that low frequencies were represented more laterally and higher frequencies more medially in HG.

Most human electrophysiology has been conducted using non-invasive EEG and MEG methods. To demonstrate evidence for tonotopic organization, variations in the following parameters have been examined: the latency of the transient and steady-state waveform responses, as well as the location and orientation of the dipole source related to the frequency of the sound. Steady-state responses are sustained fluctuations in electrical activity throughout the stimulus period. Transient responses are short-lived responses, usually at stimulus onset and offset. In terms of waveform latency, tonotopic organization can be demonstrated by the timing of the N100 peak. Stufflebeam et al. (1998) showed that for the lower frequencies, the N100 peak appeared later (159-180 ms) than the peak for the high frequencies (110-150 ms). It is uncertain how the orientation of this tonotopic gradient maps onto human auditory fields.

In terms of waveform source estimates, dipole fitting for a middle latency peak (10-50 ms, ‘Pam’) showed that sources become more medial the lower the frequency (Pantev et al., 1995). In contrast, a later component (the N100/M100 that peaks 100 ms after the stimulus onset), has its dipole source more medially the higher the sound frequency (Pantev et al., 1995; Pantev et al., 1988; Verkindt, Bertrand, Perrin, Echallier, & Pernier, 1995). It is possible that the N100 and the Pam represent two different tonotopic maps. Evidence from dipole fitting of steady-state responses also seem to support the same medial-high and lateral-low orientation of frequency sensitivity (Pantev, Roberts, Elbert, Ross, & Wienbruch, 1996), although not as clear as the transient N100 component (Pantev, Eulitz, Elbert, & Hoke, 1994). Furthermore, in an EEG study, Verkindt et al. (1995) showed that the frequency of the sound does not affect the location of the dipole source, but that its orientations appears to vary as a function of the foldings of different parts of HG.

In a critical review of MEG studies, Lütkenhöner and colleagues (2003) stated that two decades of MEG studies using a single dipole analysis of the wave N100 have failed to provide robust evidence for tonotopy. As the authors suggest, the main reason for these shortcomings is that the N100 wave originates from many fields of the auditory cortex. It is highly unlikely that a dipole analysis can distinguish between fields as they are too close to one another. Furthermore, the N100 is more likely to originate from PT than from primary auditory cortex. Lütkenhöner and colleagues (2003) suggested that a solution to the problem of single dipole modeling would be to implement multiple dipole modeling, but with certain constraints. Another proposed reason for the failure is the high inter-subject variability of the N100 source that is concealed by the grand average. These criticisms of MEG can also be applied to EEG.

PET has also been used to study human tonotopy. Lauter et al. (1985), suggested that regions more sensitive to 4000 Hz tones were located deeper and more posterior in the auditory cortex than areas more sensitive to 500 Hz tones. However, Johnsrude et al. (2002) concluded that PET is not an appropriate tool to study the fine-detailed functionality of the auditory cortex because of its poor spatial specificity.

In contrast to the shortcomings of EEG, MEG and PET, fMRI seems to be an ideal tool for exploring the spatial distribution of the frequency-dependent responses across the human auditory cortex because it provides good spatial resolution (up to about 0.7 mm) and the analysis requires only a few *a priori* modeling assumptions. In addition, it is possible to detect statistically significant activation using individual participants. This is important when determining fine-grained spatial organization because averaging data across different listeners would inevitably blur the subtle distinctions. A number of recent studies have sought to determine the organization of human tonotopy, but here only the most important ones will be considered (Formisano et al., 2003; Langers, Backes, & van Dijk, 2007; Schönwiesner, Von Cramon, & Rubsamen, 2002; Talavage, Ledden, Benson, Rosen, & Melcher, 2000; Talavage et al., 2004). To avoid the problem of neural adaptation (i.e. a decrease in the neural response to a stimulus due to repeated presentations of the stimulus), experimenters tend to choose stimuli that elicit robust auditory cortical activation. For example, Talavage and colleagues (2000) presented amplitude-modulated (AM) signals, while Schönwiesner and colleagues (2002) presented sine tones that were frequency modulated across a narrow bandwidth. Langers and colleagues (2007) used a signal-detection task in which the tone targets at each frequency were briefly presented (0.5 s). In agreement with the primate literature, evidence for the presence of tonotopic organization is most apparent within primary auditory cortex while frequency preferences in surrounding non-primary cortex are broader (Langers et al., 2007). For this reason, the primary auditory cortex will be the main focus of this section.

To identify a low frequency-sensitive region in fMRI, a condition in which participants are presented with high-frequency tones is typically subtracted from a condition in which they are presented with low-frequency tones (e.g. Talavage et al., 2000). Care is taken that the frequencies of the sounds are far enough apart in frequency and so they do not activate overlapping regions. It is hypothesized that a region sensitive to low-frequency tones will show a greater BOLD response to low than to high frequencies.

Figure 2.4 provides a summary of the most reliable tonotopic gradients and endopoints as revealed by four studies (Formisano et al., 2003; Schönwiesner et al., 2002; Talavage et al., 2000; Talavage et al., 2004). In their first study, Talavage and colleagues (2000) contrasted pairs of low (< 66 Hz) and high (> 2490 Hz) frequency stimuli of moderate intensity. These activated regions were considered as the endpoints of frequency gradients. In total, Talavage and colleagues identified eight frequency-sensitive regions across HG (primary auditory cortex) and the surrounding superior temporal gyrus (STG, non-primary auditory cortex). Each region was reliably identified across listeners and the regions were defined by a numerical label (1-8). Regions 1-4 occurred around the medial two-thirds of HG and are good candidates for representing frequency coding within primary auditory cortex (Figure 2.4). Specifically, regions 2 and 4 seem to be in Te 1.1, region 1 in Te 1.0, and region 6 in Te 1.2. Finding several endpoints does not provide direct confirmation of tonotopy because tonotopy necessitates a linear gradient of frequency sensitivity. Nevertheless, Talavage et al. argued that the regions 1-3 were at least consistent with predictions from primate electrophysiology. The arrangement of the three regions encompassed the primary auditory cortex, suggested a common low-frequency border and had a mirror-image reversed pattern. This interpretation was criticized by Schönwiesner et al. (2002) who stated that it was wrong to associate these regions with specific tonotopic fields because pairs of low- and high-frequency regions could not clearly be attributed to specific frequency axes nor to anatomically-defined fields. Indeed, in their own study, they did not observe the predicted gradual decrease in the amplitude of the frequency response at locations away from the best-frequency region, but instead found a rather complex distribution of response profiles. The authors speculated that the regions of frequency sensitivity reflected, not tonotopy, but distinct cortical regions that each preferred different acoustic features associated with a limited-bandwidth signal. Additionally, Schönwiesner et al. (2002) found evidence that low-frequency region 1 was actually two distinct low-frequency regions, one in the medial first transverse temporal gyrus (1a) and one in medial HS (1b).

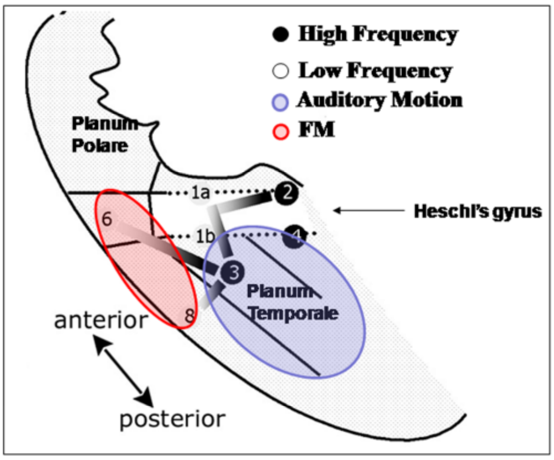


Figure 2.4 Schematic representation of the most consistently found tonotopic gradients in the auditory cortex, on an oblique slice of the brain, cutting across the auditory cortex. Gradients depict the tonotopic gradients as identified by three different studies (Schönwiesner et al., 2002; Talavage et al., 2000; Talavage et al., 2004). The approximate locations of FM-sensitive (red) and motion-sensitive (blue) regions are also shown. Adapted from Hall (2005).

The BOLD signal measured in fMRI reflects only a very small part of the MR signal. The MR signal includes fluctuations due to heartbeat, respiration, temperature, movement and other sources. Typically, the BOLD response due to a stimulus presentation ranges between 2 and 5% (Hall & Paltoglou, 2008; Huettel et al., 2004). If the signal-to-noise ratio (SNR) is low, then it is impossible to statistically separate the signal due to the stimulus from the noise. This is very important for any fMRI research, but especially for investigating tonotopy, since there are additional difficulties due to the position and the small size of the auditory regions. One way to increase BOLD SNR is by using a high strength of the static magnetic field of the MR scanner (Turner et al., 1993). By doing so, more spins align with the static magnetic field, and thus more spins are then tipped to produce the MR signal.

Frequency sensitivity in the primary auditory cortex was studied using a 7 Tesla ultra-high field MR scanner to improve the BOLD SNR and to provide fine grained (1 mm^3) spatial resolution (Formisano et al., 2003). Formisano and colleagues sought to map the progression of activation as a smooth function of frequency across HG. Computing the locations of the best response to six single frequency tones (0.3, 0.5, 0.8, 1, 2, and 3 kHz). Flattened cortical maps of BF revealed two mirror-symmetric gradients (high-to-low and low-to-high) travelling along HG from an anterolateral point to the posteromedial extremity. In general, the amplitude of the BOLD response decreased as the stimulating tone frequency moved away from the best frequency tuning characteristics of the voxel. A receiver coil placed close to the scalp over the position of the auditory cortex is another way to achieve a good BOLD SNR and this was the method used by Talavage et al. (2004). Talavage and colleagues measured BF responses to an acoustic signal that was slowly modulated in frequency across the range 0.1 to 8 kHz. Again, the results confirmed the presence of two mirror-symmetric maps that crossed HG (extending from the anterior first transverse temporal sulcus to the posterior HS) and shared a low-frequency border.

Similarly to physiological studies to non-human primates, human fMRI studies have produced evidence for the existence of multiple tonotopic maps in the auditory cortex. Both sets of evidence have shown that the primary auditory cortex is more narrowly tuned to sound frequency (Te 1.1 and Te 1.0 in humans, AI and Rostral in macaques), while neurons in the non-primary cortex have much broader tuning. This is important for Experiment 1, as it indicates that it would be more fruitful to focus on frequency-sensitive regions of the primary auditory cortex to investigate attentional modulation when attending to sound frequency.

2.3 Coding of spatial and non-spatial auditory features

There is evidence that the auditory cortical system is organized into two streams: an anterior stream, including anterior non-primary auditory cortex, which appears to code the non-spatial features (i.e. pitch or FM) and a posterior stream, including the posterior auditory non-primary cortex coding spatial attributes, such as spatial location and motion (Barrett & Hall, 2006; Degerman, Rinne, Salmi, Salonen, & Alho, 2006; Hart, Palmer, & Hall, 2004; Lomber & Malhotra, 2008). This idea originates from the visual system, where this organization is very well established. A meta-analysis of several studies revealed that there is partial support for such a segregated organization (Arnott, Binns, Grady, & Alain, 2004). Experiment 2 used FM as the non-spatial feature, and spatial motion as the spatial feature, to activate different parts of the auditory cortex. Thus, a review of these two features follows. The review includes spatial (non-moving) processing, because this feature was used in

Experiment 3, and also because localization of static sounds and motion processing appear to be closely linked, both in terms of the mechanisms as well as in terms of which brain regions are involved.

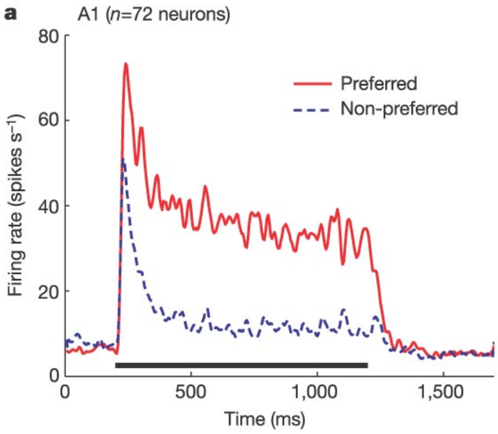
2.3.1 Coding of temporal modulation

Sounds in the environment rarely consist of steady state tones. Rather, they have some kind of modulation over time either in frequency (FM) or in amplitude (AM). Slow modulations are important for speech and melody, while fast modulations are important for pitch perception. The coding of temporal modulation has been extensively studied in the auditory system of cats, rats, monkeys and humans, typically using a controlled sinusoidally modulated signal.

Human neuroimaging (Hart, Palmer, & Hall, 2003) and electrophysiological studies on the auditory cortex of awake marmoset monkeys (Liang, Lu, & Wang, 2002) as well as of anaesthetized rats and cats (Eggermont, 1994; Gaese & Ostwald, 1995) have shown that sinusoidal AM and FM sounds have a similar representation in the auditory cortex. This result indicates that cortical neurons extract the temporal profiles of modulated tones by the same mechanism, regardless of the spectral content of the sounds. For this reason, evidence for both FM and AM will be considered here, although it is the former that is of interest for Experiment 2.

The higher in the auditory system, the less accurately is the temporal modulation encoded. In the auditory nerve, temporal modulation is represented faithfully in temporal discharge patterns (Joris & Yin, 1992). The further along the ascending auditory pathway [cochlear nucleus (Rhode & Greenberg, 1994), inferior colliculus (Langner & Schreiner, 1988; Liu, Palmer, & Wallace, 2006), medial geniculate Body (Bartlett & Wang, 2007) and auditory cortex (Eggermont, 1991, 1994; Liang et al., 2002)], the more the representation of the upper limit of temporal modulation decreases. Similarly, evidence from human neuroimaging suggests that there is a hierarchical organization in encoding of temporal modulation, whereby the higher the level, the lower the preferred modulation frequency (Giraud et al., 2000). Specifically, inferior colliculus responds best to high modulation rates (32-256 Hz) in a transient manner, while HG responds best to low modulation rates (8 Hz) in a sustained manner. Furthermore, non-primary auditory cortex responds best to even lower modulation rates (4-8 Hz) again in a sustained manner. This successive reduction is partly due to the temporal integration of inputs that occurs from one processing stage to the next (Wang & Sachs, 1995). In the auditory cortex, the response to low modulation rates is typically sustained whereas the response to high modulation rates is typically transient (Giraud et al., 2000; Harms & Melcher, 2002).

Evidence that cortical neurons have preferred or ‘best’ frequency modulation rates comes also from electro-physiological studies. Specifically, Wang and colleagues (2005) recorded from neurons in the primary auditory cortex (A1) and non-primary lateral belt of awake marmoset monkeys. In both regions, when neurons were stimulated with their ‘best modulation frequency’, there was a high firing rate and a sustained response for the duration of the stimulus (Figure 2.5). When stimulated with modulated sounds which were not at the neuron’s best modulation frequency, neurons showed a high onset response, but the sustained response was much smaller than for the preferred modulation frequency.



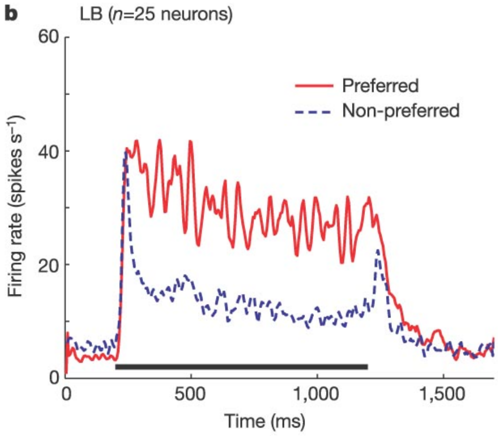


Figure 2.5 ***a***, ***b***, *Mean firing rates derived from neurons of primary auditory cortex A1 (**a**) and lateral belt area (LB) (**b**) for each neuron's preferred stimulus at preferred (red) and non-preferred (blue) modulation frequencies (10−20 repetitions at each modulation frequency). Stimuli were pure tones AM and FM, as well as broadband AM sounds. The preferred modulation frequency, shown in red, was the best modulation frequency (BMF) of a neuron. The non-preferred modulation frequency (shown in blue) is the one that produced the minimum firing rate above the BMF. A similar trend is observed at the non-preferred modulation frequency below the BMF. The line above the x axis denotes the duration of the stimulus (1 s) (Wang et al., 2005).*

Furthermore, physiological studies have shown that cortical neurons of the primary auditory cortex (AI) appear to encode temporal modulation in terms of both the temporal firing pattern and the mean firing rate, depending on the rate of modulation. Lu and colleagues (2001) presented awake marmoset monkeys with repetitive tones, at various rates. Although this was not a modulated stimulus, this stimulus elicits a similar phase-locked response at slow presentation rates. They showed that the primary auditory cortex (AI) coded slow modulations (up to 16 Hz) in an explicit manner, as a temporal discharge code, while it codes fast modulations in an implicit manner, using a discharge *rate* code. In the first case, neurons fired synchronously at a particular phase of the cycle, while in the second case neurons fired asynchronously but modulation was represented by the average sustained discharge rate of the neurons. Liang et al. (2002) also showed that neurons in AI of the same animals preferred modulation rates at 16-32 Hz (range 8- 64 Hz).

Is there a particular region in the auditory cortex that codes FM of a particular rate? Most electro-physiological studies have only examined responses in primary fields, and thus there is no strong evidence of whether the non-primary fields respond differently to FM. Wang et al. (2005), who recorded from both primary and non-primary auditory cortex, appears to have found similar results for both regions.

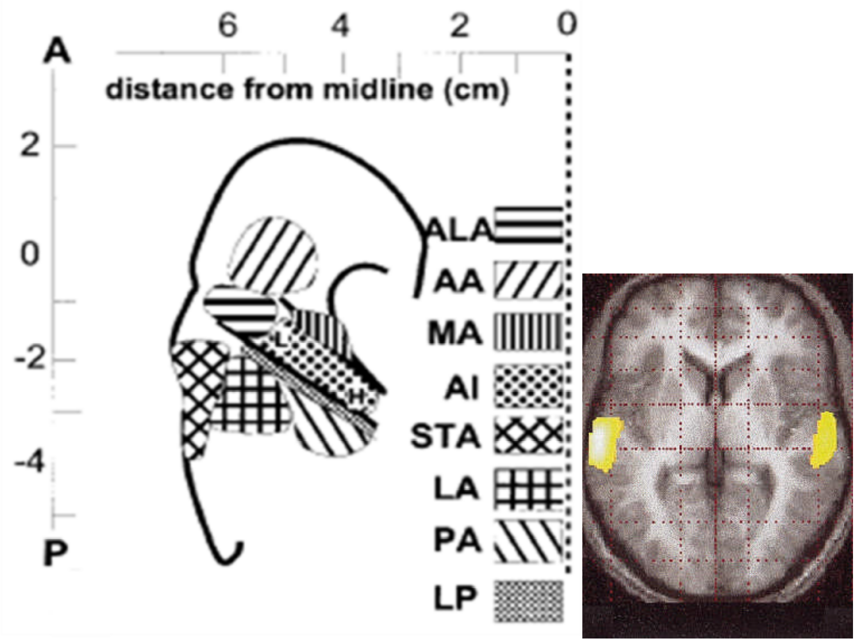


Figure 2.6 Left: Anatomical regions of the auditory cortex. A=anterior, P=posterior, S=superior, L=lateral, AI-primary auditory cortex. Originally from (Rivier & Clarke, 1997), modified by (Wallace et al., 2002). Right: Response in the auditory cortex when contrasting FM with steady state sounds in the study by Thivard et al. (2000) (right).

On the other hand, human fMRI and PET studies have provided evidence for specific region that responds more to slow-rate modulated sounds than to steady state sounds (Hall et al., 2002; Hart et al., 2003; Hart et al., 2004). Hart et al. (2004) showed that sinusoidal FM at a slow rate (5 Hz) elicits activity mainly around the lateral portion of HG as well as lateral PT. These regions appear to correspond to the anatomical regions of anterolateral HG (ALA), identified by Wallace et al (2002), as well as the lateral PT region, namely STA and possibly LA, identified by Rivier and Clarke (1997) (Figure 2.6, left). A similar result was shown by a PET study by Thivard et al. (2000). In this study, synthetic sounds with spectral maxima that were modulated in time were used as FM sounds. They were similar to vocal sounds in structure. Compared to sounds that were spectrally stationary, FM sounds showed bilateral activation in the lateral PT and lateral HG (Figure 2.6, right). These are the regions where sensory response for slow-rate FM sounds is expected in Experiment 2 (Chapter 5).

2.3.2 Coding of auditory spatial location

To enable us to localize sounds, the auditory system has to calculate particular cues from the sounds that arrive at both ears. There are three auditory spatial cues: interaural time differences (ITD), interaural level differences (ILDs) and spectral cues. ITDs are most suitable for spatial localization of low-frequency sounds, while ILDs for high-frequency sounds. These two cues are binaural cues, and enable localization of sounds in the azimuth. The spectral cues are monoaural cues, that enable localization in the vertical dimension (Butler & Belendiuk, 1977). The ridges and cavities of the pinna, as well as the rest of the body, introduce ‘spectral signatures’ to the sound wave of frequencies higher than 4000 Hz, which is received by higher auditory centre as localization cues (Plack, 2005).

Spatial location coding-animal physiology. ITD and ILD cues are initially processed in superior olivary complex (SOC) (Yin & Chan, 1990). There is evidence that the medial part of SOC codes mainly ITDs while the lateral part codes both ITDs and ILDs (Joris, 1996; Joris & Yin, 1995). Specifically, the neurons in the medial part of SOC receive excitatory signal from both the cochleae, and their response appears to be facilitated when the signals from both ears coincide, thus coding ITDs. The lateral part of the SOC, as well as containing neurons that receive excitatory input from both cochleae, also contains neurons that receive excitatory input from one and inhibitory input from the other cochlea. These latter neurons appear to be coding ILDs, while the former neurons code ITDs.

Visual spatial encoding is mediated by a systematic representation of space in the visual cortex (Warnking et al., 2002). Individual neurons are tuned to particular locations. Neural spatial tuning is narrower in the primary visual area than in non-primary regions. In contrast, sound location is not coded by single auditory cortical neurons that are sharply tuned to individual locations. Rather, it is coded by a network of broadly spatially-tuned neurons (Stecker & Middlebrooks, 2003). One theory is that auditory space is coded by opponent channels (Harper & McAlpine, 2004; McAlpine, 2005). According to this model, there are two neuronal populations, one tuned to contralateral, and one tuned to ipsilateral space. Each hemisphere contains both. Each channel codes stimuli that are in their field by graded changes in neuronal responses, in a panoramic way (i.e. they don’t represent a particular location). So, spatial location is coded by comparing the response of the two channels.

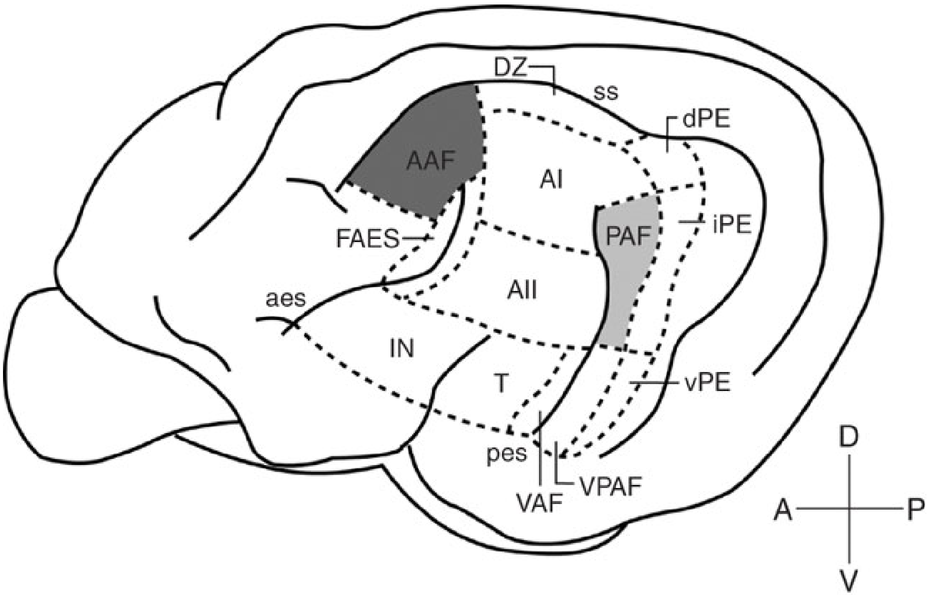


Figure 2.7 Illustration of the auditory cortex of a cat (Lomber & Malhotra, 2008). AAF=anterior auditory field, AI=primary auditory cortex, AII=second auditory cortex, dPE=dorsal posterior ectosylvian area, DZ=dorsal zone of auditory cortex, FAES=auditory field of the anterior ectosylvian sulcus, IN=insular region, iPE=intermediate posterior ectosylvian area, PAF=posterior auditory field, T=temporal region, VAF=ventral auditory field, VPAF=ventral posterior auditory field, vPE=ventral posterior ectosylvian area, aes=anterior ectosylvian, pes=posterior ectosylvian sulcus, ss=suprasylvian sulcus,

There is evidence that certain auditory fields more sensitive to spatial coding than others (Stecker, Harrington, Macpherson, & Middlebrooks, 2005; Stecker & Middlebrooks, 2003). Stecker and colleagues (Stecker et al., 2005; Stecker & Middlebrooks, 2003) studied anesthetized cats while they were presented with sounds in free field, at various locations. They showed that neurons in the posterior auditory field (PAF, see Figure 2.7) and dorsal zone (DZ) were more sensitive to spatial location with narrower spatial tuning than AI. Especially in PAF, an increase in stimulus intensity did not broaden the spatial tuning of neurons, as much as in AI. Furthermore, in PAF and DZ, spike counts depended on stimulus location more than in AI. Finally, AI appears to be most sensitive to contralateral space, while posterior non-primary field appears to be sensitive to both contralateral and ipsilateral space.

Lomber & Malhotra (2008) gave further support for PAF as a space-sensitive region, this time in awake-behaving cats. They showed that when deactivating PAF by cooling it down, cats showed deficits in free-field sound localization, but not in sound identification. Electrophysiological study on awake macaque monkeys showed that the caudo-medial non-primary area had the sharpest spatial tuning (Recanzone, Guard, Phan, & Su, 2000; Woods, Lopez, Long, Rahman, & Recanzone, 2006). In the belt region, narrower spatial tuning was also shown in anesthetized macaques for the caudo-medial field, compared to the anterior-lateral and medio-lateral fields (Tian et al., 2001). These fields are situated posterior to the primary auditory cortex, thus possibly a homologue of PAF (and human PT).

Spatial location coding- human neuroimaging. A human fMRI study illustrates that a hierarchy of cortical and subcortical regions are involved in ITD coding; inferior colliculus, medial geniculate body and primary auditory cortex (Krumbholz, Schonwiesner et al., 2005) (Figure 2.8). The right parietal cortex also appears to be crucial for spatial processing, as studies on individuals with spatial neglect has shown. Furthermore, patients with lesions in right inferior or superior parietal cortex are impaired at attending to objects in the left hemifield. Note that due to the differences in the extent of lesions, it is not clear which particular part of the parietal cortex causes neglect. Although the deficit is typically associated with visual stimuli, several studies have shown that patients show impairment for sound stimuli as well (Bisiach, Cornacchia, Sterzi, & Vallar, 1984; Eramudugolla et al., 2007; Pavani, Husain, Ládavas, & Driver, 2004; Soroker, Calamaro, Glicksohn, & Myslobodsky, 1997). Furthermore, Karnath and colleagues (2001) showed some evidence that lesions in posterior STG on the right hemisphere are responsible for spatial neglect.

2.3.3 Cortical representation of auditory motion

It has been established by both human and animal studies that neurons in region MT/V5 are sensitive to direction and speed of motion (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Maunsell & Van Essen, 1983; Orban et al., 1995). A quest of auditory studies has been to find out whether sub-cortical or cortical auditory fields show a similar sensitivity.

Spatial motion coding-animal physiology. Recording from the inferior colliculus of anesthetized guinea pigs, Ingham et al. (2001) showed that, although some neurons preferred different motion direction (clockwise or anti-clockwise) relative to each other, there was no clear evidence that neurons were selective to a particular motion direction or velocity. Their results suggest not a motion-sensitive system, but rather a coding of motion through adaptation. Specifically, neurons showed high levels of activity for the first sound of the moving sequence, and much less for the subsequent sounds. When the time between the sounds was increased, neurons showed an enhanced response for the sounds that followed the first in the sequence, consistent with the recovery from adaptation. This is against the notion that neurons are motion-sensitive. However, as the authors note, these results could be due to the fact that they did not use continuous motion.

Other studies have found neurons in primary field AI (Poirier, Jiang, Lepore, & Guillemot, 1997) and in an anterior non-primary field (AAF) (Jiang, Lepore, Poirier, & Guillemot, 2000) of cats, which appear to be sensitive to a specific motion direction. These results are surprising, as one might expect PAF to be sensitive to motion, because it has been found to be sensitive to static localization. Furthermore, tentative evidence for motion sensitivity comes from a study on awake monkeys that were presented with sounds varying in location showed that response of AI neurons to a sound was influenced by the location of the preceding sound (Malone, Scott, & Semple, 2002). What is more, neurons showed preference for particular directions.

In summary, it is still not clear whether motion is computed by comparing ‘static’ snapshots of sounds at different location, or whether there are actually neurons that are sensitive to dynamic features of space, such as motion direction and velocity, just as in the visual system (Middlebrooks, Harrington, Macpherson, & Stecker, 2008).

Spatial motion coding-human neuroimaging. Some fMRI studies have examined motion coding by either presenting sound sequences that contain discrete shifts in spatial location, (Krumbholz et al., 2007; Krumbholz, Schonwiesner et al., 2005; Krumbholz, Schönwiesner et al., 2005), or continuous motion (Baumgart, Gaschler-Markefski, Woldorff, Heinze, & Scheich, 1999; Pavani, Macaluso, Warren, Driver, & Griffiths, 2002)

Different studies have used different cues to create the percept of motion; such as ITDs (Krumbholz et al., 2007; Krumbholz, Schonwiesner et al., 2005; Krumbholz, Schönwiesner et al., 2005) and ILDs (Baumgart et al., 1999; Griffiths & Green, 1999; Griffiths, Green, Rees, & Rees, 2000; H.C. Hart et al., 2004) and head-related transfer functions (HRTFs) which includes spectral cues that contribute to an externalized spatial percept for a sound presented over headphones (Pavani et al., 2002; Warren, Zielinski, Green, Rauschecker, & Griffiths, 2002). To localize a response to moving sounds, neuroimaging studies typically compare a condition in which the sounds are moving, with a condition that the sounds are stationary, presented at the midline (Krumbholz et al., 2007) or at various locations (Poirier et al., 2005)

Krumbholz, Schonwiesner et al. (2005) examined which auditory regions are involved in the localization of static sounds and which in auditory motion processing. To investigate the processing of static lateralized sounds, monoaurally presented sounds (only left and only right) were contrasted with diotic sounds (identical in both ear, perceived in the midline). This contrast showed activation in inferior colliculus, medial geniculate body, primary auditory cortex, and anterior PT bilaterally (Figure 2.8, red). On the contrary, when contrasting moving sounds (varying ITD from -1000 µs to 1000 µs) with stationary diotic sounds (ITD=0µs), there was a response in PT, extending to the TPJ (Figure 2.8, blue). That is, there was a relatively clear segregation of the response whereby static location processing engaged earlier sub-cortical and primary cortical regions, while motion processing engaged non-primary regions. The authors speculate that a reduction in the response for the stationary sounds in PT occurred because neurons in non-primary auditory cortical regions produce mainly a phasic, rapidly adapting response, i.e. they respond only to change. On the other hand, neurons in earlier auditory regions produce mainly a sustained, slowly adapting response, i.e. they respond for as long as the stimulus is present (Harms, Guinan, Sigalovsky, & Melcher, 2005). This suggests that the response in PT for moving stimuli possibly reflects adaptation of the phasic PT neurons to stationary sounds (Krumbholz, Schonwiesner et al., 2005). The fact that motion processing takes place so late in the auditory pathway hierarchy, indicates that the auditory system analyses individual binaural representations of sounds in consecutive places in space, that are relayed from lower auditory regions rather than by creating a smooth continuous representation of auditory motion (Krumbholz, Schonwiesner et al., 2005).

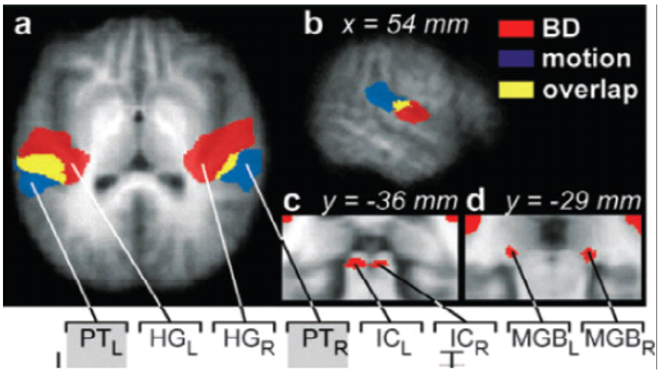


Figure 2.8 (a) Axial and (b) sagittal view of the brain, showing activation related to processing of location cues (red monaural left/right vs diotic) and auditory motion (blue, moving sounds>diotic) in the auditory cortex. Overlap is shown in yellow. Processing of location in c) IC and d) MGB (Krumbholz, Schonwiesner et al., 2005).

Other neuroimaging studies have also shown a response to moving sounds in PT (Baumgart et al., 1999; H.C. Hart et al., 2004; Krumbholz, Schonwiesner et al., 2005; Pavani et al., 2002; Warren et al., 2002) and there is some consensus that this is the motion centre of the auditory system, especially on the right (Baumgart et al., 1999). Krumbholz, Schönwiesner et al.(2005) have shown that the right PT responds to auditory motion in both hemifields, while the left responds only in the contralateral hemifield. This is consistent with right hemisphere dominance for spatial coding (c.f. literature on visual neglect). Note that, which particular anatomical field of PT is involved in motion processing is still not clear, as different studies appear to show variable results. In terms of the anatomical regions identified by Rivier and Clarke (1997), sometimes response appears to be in medial region PA, while in other studies it appears to be in LA (Figure 2.6, left).

In the visual system, it has been suggested that there are two distinct stages of motion processing. The first stage takes place in the visual region V5, (Braddick et al., 2001; Ffytche, Skidmore, & Zeki, 1995) and the second, cognitive stage, in right parietal cortex (for a review see Battelli, Pascual-Leone, & Cavanagh, 2007). There is evidence for a similar scheme for the auditory motion processing. Warren and colleagues (2003) suggest that the PT reflects first stage, while parietal cortex reflects the second stage. Specifically, several studies suggest that the inferior parietal cortex, TPJ (Bremmer et al., 2001; Griffiths, Buchel, Frackowiak, & Patterson, 1998; Griffiths & Green, 1999; Griffiths et al., 2000; Griffiths, Rees et al., 1998; Krumbholz et al., 2007; Krumbholz, Schönwiesner et al., 2005) and operculum (Warren et al., 2002) are involved in aspects of motion processing. Furthermore, there is tentative evidence that neglect patients with a right parietal lesion show deficits in the perception of both static and moving spatial cues (Battelli et al., 2001). More studies are needed to confirm the link between neglect and a deficit in auditory motion processing. Additionally, there is evidence that the *superior* part of right parietal cortex is involved in auditory motion processing (Griffiths, Buchel et al., 1998; Griffiths & Green, 1999; Griffiths et al., 2000; Griffiths, Rees et al., 1998; Pavani et al., 2002), especially on the right. Lewald et al. (2002) showed that bilateral inhibition of the posterior parietal cortex with transcranial magnetic stimulation (TMS), shifted the perception of sound location, while it did not affect ITD discrimination acuity. This result indicates that this region is involved in changes in spatial location, rather than the processing of spatial cues per se.

Poirier et al. (2005) have shown that moving sounds contrasted with static sounds induced a response not only in the right PT, premotor and parietal regions bilaterally, but additionally in visual motion regions V5. This is tentative evidence to suggest that V5 is involved in motion processing of auditory stimuli, which is supported by a TMS study, which showed impairment of auditory spatial judgment when TMS was applied in occipital cortex (Lewald, Meister, Weidemann, & Töpper, 2004).

In summary, auditory spatial coding appears to involve certain sub-cortical nuclei and the primary auditory cortex, while motion coding appears to involve PT. Parietal cortex appears to be involved in both static localization and motion processing. There is very little support from the animal literature of a motion-sensitive region, equivalent to PT (and visual V5/MT), which is partly due to the small number of motion-processing studies. Although posterior non-primary area PAF (and CM belt region in monkeys) appears to be sensitive to static location, there is no evidence that it is also sensitive to motion.

2.4 Summary

In this chapter, the coding of auditory features in the auditory system was discussed. There is considerable evidence from both human and animal studies that the main organizing principle of the auditory cortex is tonotopicity, which is particularly prevalent in primary auditory cortex. Furthermore, there is evidence for regions sensitive to FM in antero-lateral non-primary auditory cortex, and regions sensitive to motion in posterior-medial non-primary auditory cortex. Note that this evidence comes mainly from human studies, possibly because these techniques are able to have an overview of the activity in the whole auditory cortex. This hierarchical organization of auditory cortical processing resembles very much the organization of the visual cortex. The question investigated in subsequent experimental chapters is whether selective attention to these features is also mediated in a way similar to the visual selective attention, i.e. in a feature-specific way.

Chapter 3: Optimizing the methods for Experiment 1

3.1 Introduction

There is considerable evidence that selectively attending to a visual feature of an object enhances the neural response in the region of the visual cortex that is sensitive to the attended feature (Corbetta et al., 1990; Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991; Schoenfeld et al., 2007; Tootell et al., 1998). Conversely, suppression of the neural response occurs in the regions of the visual cortex that are sensitive to the unattended visual features (Tootell et al., 1998). Enhancement and suppression combine to increase the contrast between stimulus and background (Treue & Trujillo, 1999). In its strictest interpretation, selective attention is described as feature-specific when enhancement acts only in those brain regions that are sensitive to the attended attributes of the stimulus, and not elsewhere.

To test for feature-specific attentional modulation, visual researchers have typically located the visual brain regions that are sensitive to a particular feature, and then investigated the effects of attention in that region. To shed light at how fMRI researchers investigate selective attention, it is informative to consider in detail the method used by Tootell et al. (1998).

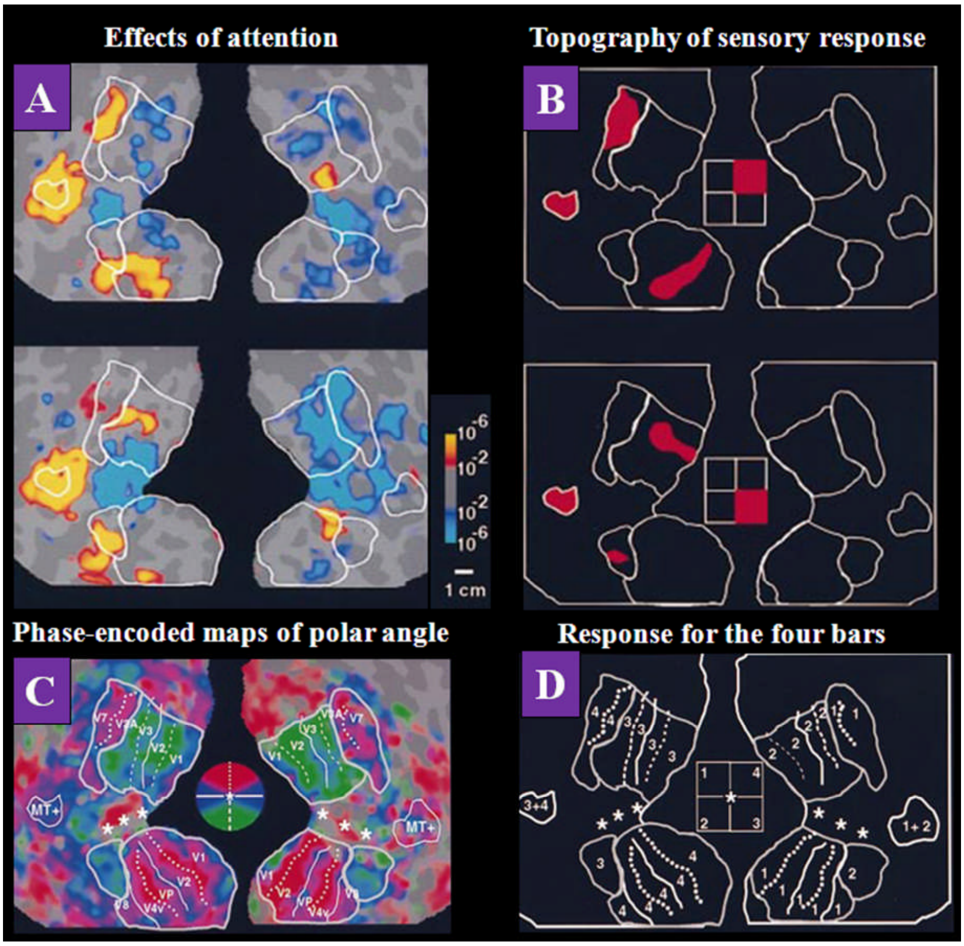


Figure 3.1 Sensory activation and attentional modulation of visual stimuli presented in four different quadrants of space. All results are presented on flattened maps of the visual cortex. The borders of different regions of the visual cortex are illustrated by the white lines. All data are from one participant. (A) Loci of the effect of attentional modulation when selectively attending to a bar, when contrasted with the three conditions whereby participants attended to one of the other three bars. Red through yellow: enhancement of activity when attending to bar in that location defined by the contrast 'attend one bar>average of attending to other 3 bars; blue: suppression of response when attending to the bar compared to when attending to the other three bars. (B) Topography of the sensory activation for stimuli at the location indicated by the red rectangles in the middle of the image. (C) Phase-encoded maps of polar angle (D) Summary of the topography of sensory responses for each of the four bars. Adapted from Tootell et al., (1998).

The visual cortex is retinotopically organized; that is, there is a systematic representation of space where visual objects that are adjacent in the visual field are represented by adjacent populations of neurons within a region of the visual cortex (Warnking et al., 2002). Tootell et al. (1998) investigated whether attention to specific spatial locations enhances the response in regions of the visual cortex that are sensitive to those locations. In the first part of the study, the authors used a phase-encoding method to map the cortical response to four different quadrant of the visual field. To do this, they presented participants with simple visual stimuli (bars) that appeared asynchronously in all of the four quadrants, in the context of a passive viewing task. Panel B in Figure 3.1 demonstrates the retinotopic mapping of the upper right and the lower right quadrants. The visual response is represented in certain regions of the contralateral visual cortex, including parts of V1, V2, V7 and other visual regions that are defined in panel D of Figure 3.1. Area V1 is known as the primary visual cortex and areas V2-V7 refer to different non-primary visual fields. In the second part of the study, the authors used a target discrimination task to investigate the effects of selective attention to stimuli presented in the different quadrants. The targets were horizontal bars and the distractors were vertical bars. Participants were signaled to attend to a particular quadrant in which the bars were a different color to the other three.

In their analysis, the authors contrasted a condition in which attention was directed to one quadrant with the other three conditions in which attention was directed to each one of the other quadrants. A relative increase in the BOLD response is evidence of enhancement while a relative decrease in the BOLD response is evidence of suppression. The results are shown in panel A (Figure 3.1). Enhancement is shown in yellow, suppression is shown in blue. Enhancement by attention predominantly occurred in similar regions as those responding to the sensory representation of the stimulus. This spatial correspondence is termed the ‘retinotopy of visual spatial attention’. Suppression by attention predominantly occurred in the adjacent regions of the visual cortex. These results provide powerful evidence for feature-specific attentional enhancement and suppression in the visual system. In contrast, in the auditory system there is little support for the existence of such a mechanism. However, this issue has not yet been examined in a systematic manner.

Note that, in the case of Tootell et al. (1998), enhancement and suppression occur in separate regions. To examine the effect of suppression, Tootel et al. simply reversed the enhancement contrast (Figure 3.1). In the current study, a slightly different approach for examining suppression is adopted. Enhancement and suppression are investigated within the same region, by directing participant's attention to the preferred or non-preferred part of the stimulus, respectively.

The auditory cortex is organized tonotopically and there is evidence that distinct auditory regions respond best to different frequencies (Formisano et al., 2003, Talavage et al., 2000, Talavage et al., 2004, Schonwiesner et al., 2002). This organization makes sound frequency a good candidate for the study of selective attention. Frequency-specific attentional enhancement is defined as a relative increase in the BOLD response within the high-frequency-sensitive regions when listeners attend to high-frequency sounds. In contrast, frequency-specific attentional suppression is defined as a relative decrease in the BOLD response within the high-frequency-sensitive regions when listeners attend to low-frequency sounds. The same patterns of enhancement and suppression are expected for the low-frequency sounds. A two-step procedure, similar to the one reported by Tootell et al. (1998), was considered appropriate for investigating the issue of frequency-specific auditory selective attention. First, the experiment mapped the sensory response for low- and high-frequency sounds while participants passively listened to them. Second, the experiment measured attentional modulation of the frequency-specific response while participants were performing a task.

Although some human neuroimaging studies have also shown evidence for attentional enhancement in the primary auditory cortex (Jancke et al., 1999; Woldorff et al., 1993; Woodruff et al., 1996) they were not designed to demonstrate a *feature-specific* enhancement. Furthermore, a recent by Petkov et al. (2004) failed to show any attentional enhancement in the primary auditory cortex. Electrophysiological recordings in awake ferrets demonstrate feature-specific attentional modulation in neurons of the primary auditory cortex (Fritz et al., 2003). Therefore the cortical site of selective attention to sound frequency is still unclear. One might expect selective attention to sound frequency to exert a stronger modulatory influence on primary than non-primary auditory cortex since neurons in primary auditory cortex have the sharpest frequency tuning (Kosaki et al., 1997). The null results in the primary auditory cortex from some of the fMRI studies could be due more to a lack of sensitivity than an absence of the attentional effect. Furthermore, Tootell and colleagues (1998) have shown that the magnitude of enhancement is significantly smaller in primary than in non-primary visual cortex (Figure 3.2). Interestingly, the magnitude of suppression is reasonably constant across all visual regions.

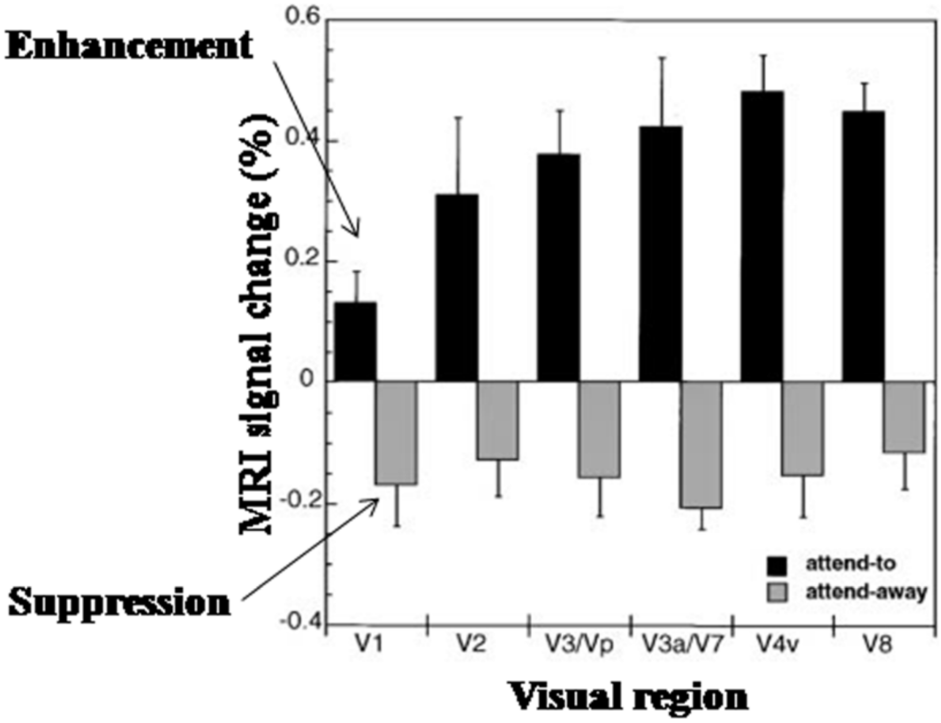


Figure 3.2 Attentional modulation in areas of the visual cortex. Black bars indicate enhancement of response when attending to a location that the specific brain area is sensitive to, and grey bars denote suppression of response when attending away from that location (adapted from Tootell et al., 1998).

Several pilot studies were conducted to optimize Experiment 1 and to ensure that it is sufficiently sensitive to investigate the effects of feature-specific attentional enhancement and suppression in the human auditory cortex. This chapter reports three pilot studies. The issues explored were the following::

i) design an effective stimulus for mapping sensory response and directing attention to one frequency or another,

ii) confirm the spatial specificity of the BOLD response to sound frequency,

iii) design an effective task that is sufficiently difficult to ensure a benefit for selective attention and a cost for divided attention,

iv) optimize a number of scanning parameters, namely voxel resolution and TE.

The rationale for exploring these particular issues is presented in the following sections.

3.1.1 Design of an effective stimulus

It was essential to have the same stimulus for both steps in the experiment. To effectively manipulate attention, stimuli should contain both target and distracting sounds, since the effect of attention is greatest when the task was more difficult (Boudreau, Williford, & Maunsell, 2006). Therefore each stimulus condition contained two different sets of sounds, i.e. low and high frequency. The high-frequency sound had a centre frequency of 4000 Hz and the low-frequency sound had a centre frequency of 250 Hz. These centre frequencies were chosen to evoke activity in non-overlapping regions along the tonotopic axis (Schönwiesner et al., 2002). The low- and high-frequency sounds were added together in a 15.5s sequence of two interleaved streams; a low-pitched stream and a high-pitched stream (Figure 3.3). Within one stimulus condition, one stream contained a majority (80%) of one frequency (fast rhythm) and the other stream contained a minority (20%) of the other frequency (slow rhythm). This is illustrated in Figure 3.3 where there is a cluster of four of the majority sounds to every one of the minority sounds. The majority and minority streams were counterbalanced across the two frequencies so that one stimulus condition contained a majority of high-frequency sounds and the other contained a majority of low-frequency sounds.

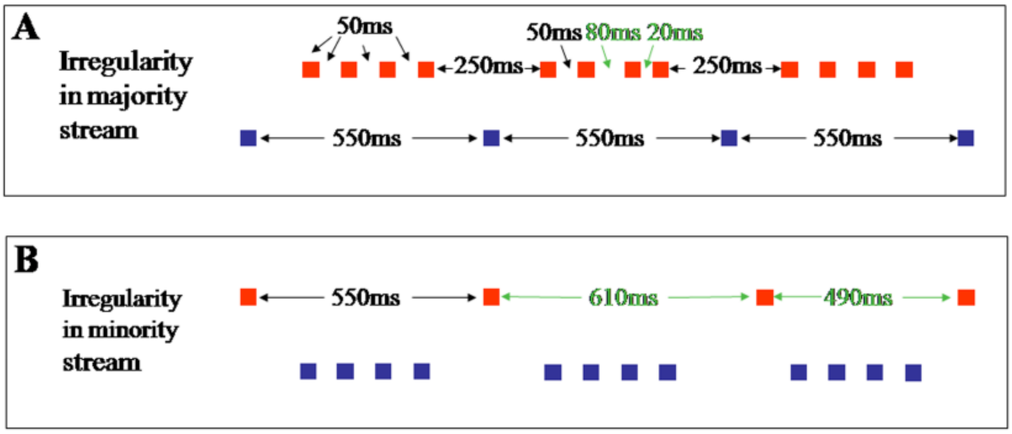


Figure 3.3 Timings of the sounds in A) high-frequency majority and B) low-frequency majority sequences. Low frequency sounds are depicted in blue, high frequency sounds in red. Irregularities are indicated in green font.

One particular issue for investigation in pilot study 1 was whether stimuli containing both low and high frequencies could be used to detect frequency-sensitive regions. Previous fMRI studies of tonotopy have favored single-frequency stimuli rather than mixed-frequency stimuli. For example Formisano et al (2003) used pulsed tones at 300, 500, 800, 1000, 2000 and 3000 Hz in six separate stimulus conditions.

Each sound was a narrow-band noise burst rather than a pure tone to generate more reliable activation in auditory cortex (Hall et al., 2002). The width of each frequency band was a third octave. For high-frequency sounds, the bandwidth was therefore 3564-4490 Hz and for low-frequency sounds, the bandwidth was 223-281 Hz. Note that the width of auditory filter at the centre frequencies is 27 Hz (for the 250 Hz sound) and 432 Hz (for the 4000 Hz sound), as estimated by the equation for the equivalent rectangular bandwidth (ERB)= 24.7(4.37x+1) (Glasberg & Moore, 1990). So, the width of the experimental stimuli was wider than the auditory filters, but far enough apart spectrally to not overlap. All noise bursts had an interaural correlation of unity, so that they were perceived in the centre of the head. Each noise burst was 50 ms in duration with 5 ms onset and offset ramps.

3.1.2 Spatial specificity of the BOLD response to sound frequency

Confirmation of spatial specificity was evaluated in pilot studies 1 and 3 by comparing the results to those of previous studies (Formisano et al., 2003; Schönwiesner et al., 2002; Talavage et al., 2000; Talavage et al., 2004). Based on these previous fMRI studies, it was expected to find multiple frequency-sensitive regions in the auditory cortex. Figure 3.4 summarizes the most consistent frequency-sensitive regions as numbered by Talavage et al. (2000) and Schönwiesner et al. (2002). Two high-frequency-sensitive regions (regions 2 and 4, Talavage et al., 2000) are expected in the medial most portion of HG (Te 1.1, Morosan et al., 2001). Low-frequency-sensitive regions are less consistent across studies. Talavage et al. (2000) reported one low-frequency-sensitive region, region 1. However, Schönwiesner et al. (2002) have suggested that region 1 is actually two spatially separate low-frequency-sensitive regions and they labeled these 1a and 1b. These two low-frequency-sensitive regions are expected in the central portion of HG (Te 1.0, Morosan et al., 2001). For practical reasons, the pilot studies use the scheme reported by Talavage et al. (2000) in which a single peak coordinate is given for region 1. All the regions described above are considered to be within the primary auditory cortex. Another low-frequency-sensitive region (region 6) is expected in the lateral most portion of HG (Te 1.2, Morosan et al., 2001). This region is thought to be a transitional region between primary and non-primary auditory cortex

(Wallace et al., 2002). Detailed analysis was performed only for those voxels in primary auditory cortical regions 1, 2 and 4 and not in region 6.

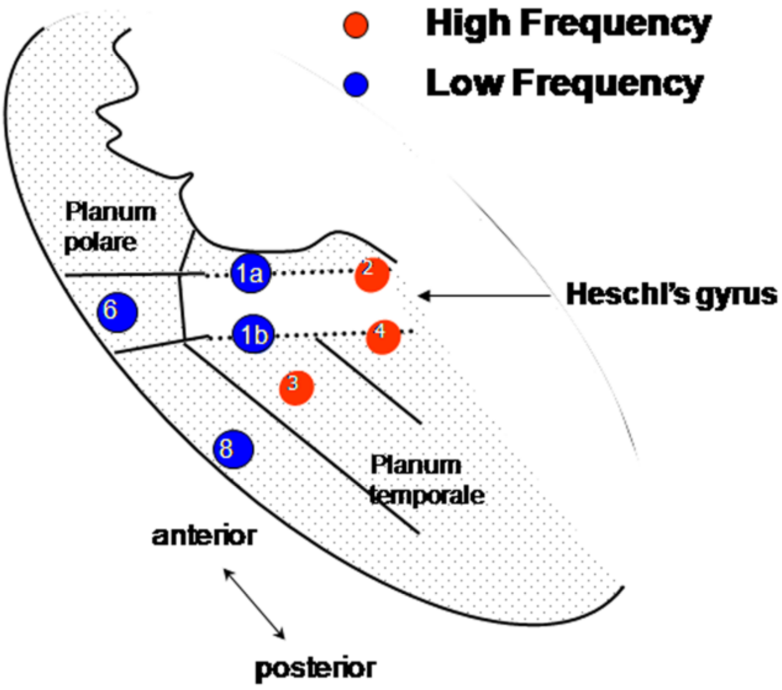


Figure 3.4 Illustration of an oblique brain slice which depicts the location of different high- and low-frequency endpoints of tonotopic maps and their location within different auditory regions. Red: high-frequency region: blue; low-frequency region. Adapted from Hall (2005).

3.1.3 Design of an effective task

It was considered important to employ a difficult listening task requiring participants to continuously monitor one frequency stream while ignoring the other. It was also considered important to monitor performance as evidence that listeners attended to one stream or another. Targets were defined by an irregularity in the rhythm. Specifically, in the majority stream, there were groups of four noise bursts, having a 50-ms interval between them (Figure 3.3). Irregularities were created always within a group of four narrow-band noise bursts, such that the target appeared 30 ms earlier or later than expected. In the minority stream, there was a 550-ms interval between each noise burst. Irregularities were created by making the target appear 60 ms earlier or later than the standard sounds (Figure 3.3). Irregularities in one stream did not affect the timings of the other stream. Participants were instructed to press a button whenever they heard a target in the attended stream. The effect of listening expertise on target-detection performance was explored in pilot studies 1 and 2. The outcome of these pilot studies determined whether or not listeners in Experiment 1 were trained in the task before the scanning began.

3.1.4 Optimization of voxel resolution and echo time

The size of human primary auditory cortex is fairly small given the rather gross scale voxel resolution of fMRI. The available fMRI evidence (discussed in Section 2.2.3) suggests that a small voxel size would be the best choice for separating low- and high-frequency responses within primary auditory cortex. However it is not clear whether sufficient BOLD SNR is available at 3 Tesla, despite using a new design of 8-channel SENSE Head Coil for improved BOLD contrast sensitivity. Therefore pilot studies 1 and 3 investigated the statistical reliability of sound-evoked activation at two different voxel sizes: the standard resolution (3x3x3 mm, pilot study 1) and a high resolution (1.5x1.5x2.5 mm, pilot study 3).

Another key scanning parameter is that of TE. TE varies with voxel resolution. The magnitude of the BOLD signal change measured in fMRI is primarily dependent on two factors: the TE and the rate of decay of transverse magnetization. TE refers to the time between the excitation by the radiofrequency pulse and the MR data acquisition. Note that the radiofrequency excitation pulse applies a 90° flip on the transverse magnetization component of the MR signal. The rate at which the transverse magnetization component of the MR signal decays after the initial radiofrequency pulse is determined by the inherent decay rate of the local brain tissue caused by the paramagnetic properties of deoxygenated hemoglobin (the T2 effect) and other static inhomogeneities in the magnetic field. These combined effects are known as the T2* effect. Through this complex relationship, changes in blood oxygenation associated with neural activity influence the MR decay parameter, T2*, leading to changes in image intensity in the T2*-weighted functional images that are acquired during an experiment. The largest BOLD signal changes occur when TE is approximately equal to the value of T2* (see Figure 3.5). From the description given above, it can be seen that T2* can differ over different anatomical regions because of the dephasing effects of static inhomogeneities in the magnetic field. Measurement of T2* in the brain region of interest is therefore important for determining the optimal value of TE to use in an fMRI experiment in order to achieve maximal BOLD SNR (Clare, Francis, Morris, & Bowtell, 2001). If TE is too long, then the transverse magnetization component has fully decayed, and thus the magnitude of the BOLD signal will be close to zero. On the other hand, if TE is too short, then the transverse magnetization component will not yet have exhibited much decay and thus there would be little difference in the BOLD signal between stimulus conditions (Huettel et al., 2004). The auditory cortex is a special case because the medial parts of the STG are located close to the nasal passages where variations in magnetic susceptibility between the air, bone and soft tissue increase MR signal dephasing (Hall et al., 1999; Henkelman & Bronskill, 1987). Thus, the value of optimal TE in the auditory cortex may be very different to that in the visual cortex. In pilot study 3, the value of T2* in the auditory cortex was measured for one participant.

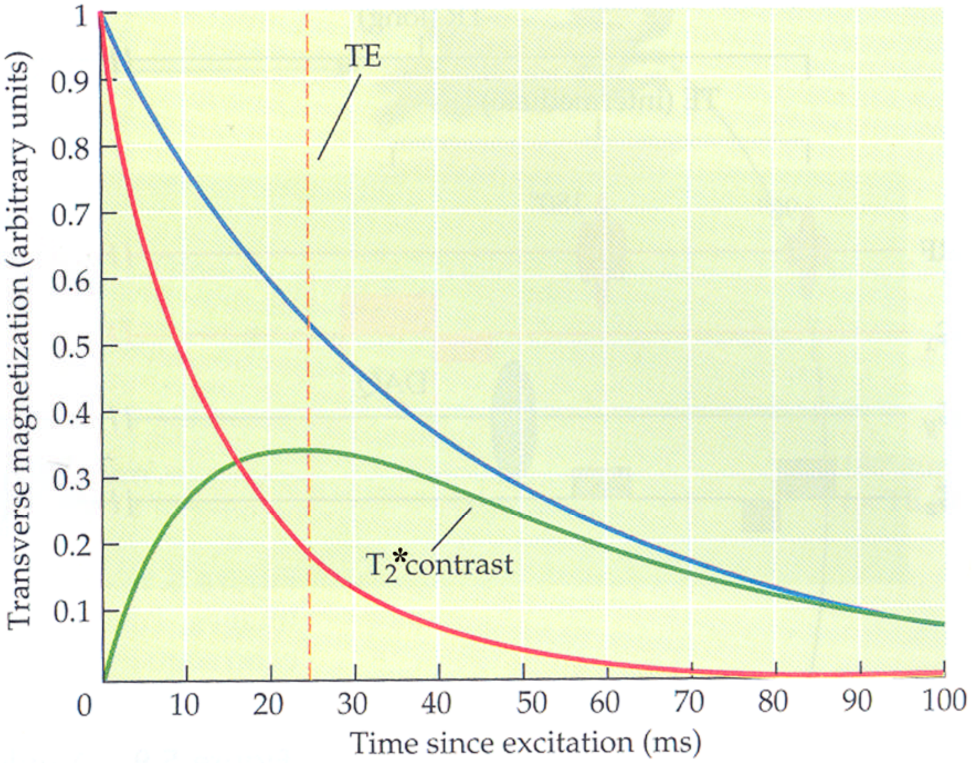


Figure 3.5 Graph showing the T2 curve, which derives from the difference between the blue and red curves. The blue and red curves represent different transverse relaxation times. In blue is the relaxation time for active brain tissue, containing a greater proportion of oxygenated hemoglobin and in red is the less active tissue that has a greater proportion of deoxygenated hemoglobin. The dashed line represents the optimal TE value at which the difference in transverse magnetization between the two tissues is greatest. Adapted from Huettel et al. (2004).*

3.2 Pilot study 1: Activation mapping using a standard voxel resolution (3x3x3 mm)

Pilot study 1 had four aims. The first aim was to confirm whether the stimuli were suitable for separately identifying activity for the two different frequency sounds. The second aim was to examine whether the spatial location of the low- and high-frequency-sensitive regions matched the location of the regions found in previous studies of tonotopy (Talavage et al., 2000; Talavage et al., 2004). The third aim was to explore the range of listeners' performance on the task. The fourth aim was to investigate attentional enhancement and suppression in the frequency-sensitive regions. To investigate these four issues both the behavioral and BOLD responses were examined.

3.2.1 Methods

Participants

Three participants were tested (2 males and 1 female, 21, 23 and 30 years old, participants #1, #2 and #3 respectively). All participants were right-handed and had normal hearing (≤25 dB for 250 to 8000 Hz). None of the participants had a history of audiological or neurological impairment. All participants were recruited via a poster at the University of Nottingham campus and were paid for their participation. Informed consent was obtained before MR scanning. The experimental procedures were approved by the local NHS trust service ethics committee (A/1/2005).

Stimuli and task

High-frequency majority and low-frequency majority stimuli were presented using high-fidelity headphones, modified to be MR compatible. The headphones were driven by a specially engineered fMRI sound system using custom-written software (Palmer, Bullock, & Chambers, 1998). High-frequency majority stimuli were presented at 94 dB SPL while low-frequency majority stimuli were presented at 91 dB SPL. This, and subsequent stimulus presentation levels were calibrated by mounting the headphones on a KEMAR manikin equipped with a free-field response microphone (Brüel and Kjær, Type 4134) and Zwislocki Coupler (Brüel and Kjær, Type DB-100), and connected to a measuring amplifier (Brüel and Kjær, Type 2636). The two stimulus conditions were crossed with three listening conditions, in a 2x3 factorial design (Table 3.1). The three listening instructions were ; i) just listen, ii) attend to the high-pitched stream and iii) attend to the low-pitched stream.

Listening instructions	Stimuli	
	High-frequency majority	Low-frequency majority
Just listen	A	B
Attend high	C	D
Attend low	E	F

Table 3.1 2x3 factorial design.

Pilot study 1 contained 15 repetitions of each of the six conditions, split across two runs. The first run was passive listening and this required participants to just listen to the two stimulus conditions, plus interleaved silent baseline conditions. The second run required participants to attend to the stimuli and press a button when they heard irregularities in the high- or the low-frequency stream. Again, these 'attend' conditions were interleaved with silent baseline conditions. The order of the conditions in each session was randomized in a latin square design. The order of the runs (passive listening followed by attend) as well as the order of the sound sequences within each session were the same for all participants. The passive listening run always came first to ensure that participants were not aware of the target detection task and so were not listening out for the timing irregularities.

Visual instructions informed participants about which task to perform during each stimulus condition. The instructions appeared throughout each sound sequence. In the passive listening run, the instruction for all conditions was ‘just listen’ while keeping eyes open. In the attend session, the instruction for attending to the targets in the high-frequency stream was ‘high sounds’. The instruction for attending to the targets in the low-frequency stream was ‘low sounds’. Participants were required to press a button with their right thumb when they detected a target in the attended stream. Again, during the silent baseline condition the instruction was ‘just listen’ and no response was required. The instructions were presented via a visual display presented using fibre-optic goggles (SV-7021, Avotec Incorporated, "Silent Vision" High Resolution Visual System). The goggles were mounted on the Philips SENSE head coil. Participants were instructed to stay as still as possible during scanning and relax. Button presses were recorded for offline analysis of target-detection performance.

Imaging protocol

Scanning was performed on a Philips 3 Tesla Intera MR scanner (Achieva/Intera Release 1.2/11) equipped with an 8-channel SENSE Head Coil for improved BOLD contrast sensitivity compared to a standard quadrature receiver coil. Scanning took place at the Sir Peter Mansfield Magnetic Resonance Imaging Centre, University of Nottingham. All participants were scanned in one session (varying from 1 hour and to 1 hour and 30 minutes). The scanning session consisted of eight stages.

i) Survey scan. This scan confirmed the participant’s head was in the isocenter of the scanner bore where the magnetic field is most homogenous. This scan took 30 s.

ii) Reference. This sequence calibrated the parallel channels on the SENSE coil to maintain image signal consistency. This sequence lasted 40 s.

iii) T1-weighted anatomical scan. This scan provided a detailed whole brain image of the cortical anatomy. The parameters of this sequence were as follows: voxel resolution 1x1x1 mm, matrix size 256x256, 160 sagittal slices, repetition time (TR)=8.2 ms, TE=3.7 ms. This scan lasted 4.5 mins.

iv) Single 18-slice T2*-weighted functional scan. From the display of the anatomical scan, an 18-slice functional scan was positioned over the centre of the superior temporal gyrus. The scan was oriented in an oblique axial position avoiding the eyes (see Figure 3.6A). Visual inspection of this single scan was used to check the orientation and position of the scans in the subsequent functional runs. The parameters of this sequence were as follows: voxel resolution 3x3x3 mm, matrix size 64x64, TR=2000 ms, TE=35 ms, flip angle=90º. All functional scans were acquired using a clustered volume acquisition sequence (Edmister, Talavage, Ledden, & Weisskoff, 1999) in which all 18 slices were acquired as rapidly as possible within the TR period. This minimum acquisition time was 1852 ms. The total time for this scan was 1.5 mins.

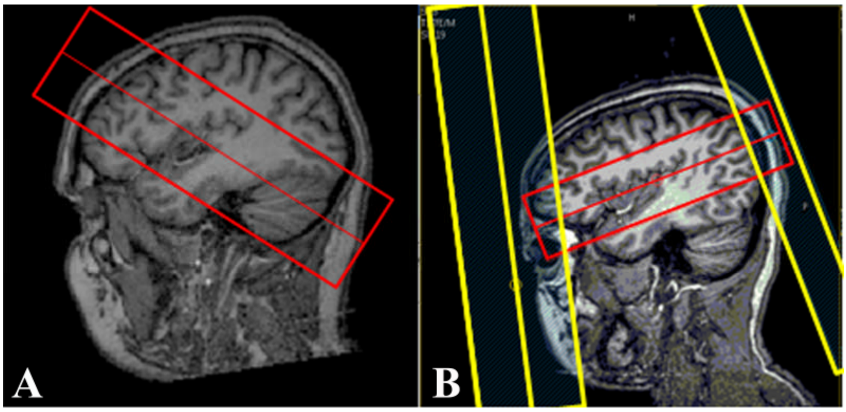


Figure 3.6 Anatomical scans showing the orientation and extent of the functional slices. (A) position of scans for Pilot study 1, shown in red. The voxel resolution was 3x3x3 mm and so the field of view extends beyond the head. The scans are oriented at 90º to the axis of the supratemporal plane to avoid Nyquist artefacts from the eyes. (B) position of scans for Pilot study 3, shown in red. The voxel resolution was 1.5x1.5x2.5 mm and so the field of view is contained within the head, including the eyes. The orientation is parallel to the axis of the supratemporal plane. Nyquists artefacts are removed by placing saturation bands, shown in yellow, at the anterior and posterior edges of the scan.

<u>v) T2* weighted passive listening run</u>. A time-series of 92 functional scans was acquired using the same parameters as in iv). The sparse sampling method was used to reduce the effects of scanner noise on the pattern of auditory activation (Hall et al., 1999). This enabled the sound sequence to be delivered mainly during the quiet period between scan acquisitions thus reducing acoustic masking by the background scanner noise. Because each sound sequence was 15.5 s in duration, scans were acquired at approximately the middle and the end (Figure 3.7A). The passive listening run lasted 12 mins.

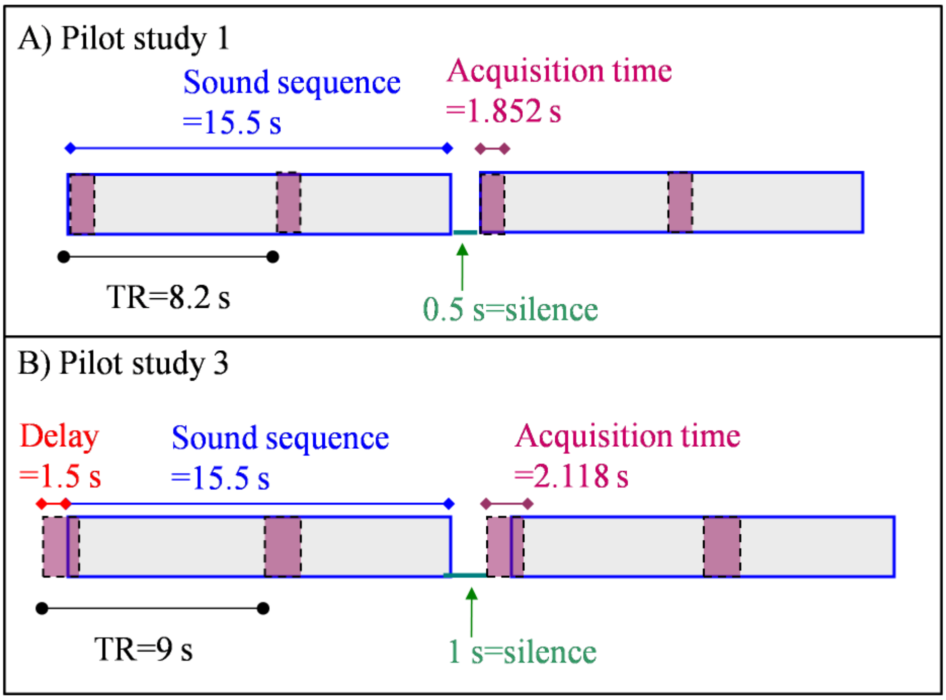


Figure 3.7 Sparse sampling (Hall et al., 1999) in pilot studies (A) 1 and (B) 3. The grey rectangular bar represents the duration of the sound sequence (15.5 sec), and the purple rectangular represents the acquisition of a single volume of data.

<u>vi) Single 36-slice T2*-weighted functional scan</u>. This scan was used to facilitate image analysis. The 36-slice scan covered a larger part of the brain than the 18-slice scan, giving more landmarks for accurate reorienting. The parameters and duration were the same as in iv). The centre of the 36-slice scan corresponded to the centre of the 18-slice scan.

<u>vii) Training</u>. Participants were given verbal instructions for responding to the timing irregularities. Practice trials contained four sound sequences for each of the four attend conditions. At the end of the practice, general comments about accuracy were provided. Training lasted between 5 and 10 mins.

<u>viii) T2*-weighted attend run</u>. A time-series of 152 functional scans was acquired using the same parameters as in v). This run lasted 22 mins.

Analysis of target detection data

Both Hit minus false alarm (FA) rate and d prime measures of performance were calculated. These are both measures that take into account individual response bias in a simple yes-no decision task. The two measures use slightly different calculations of FA rate. In the former measure, hit rate is defined by the number of hits divided by the total number of targets (#hits/# targets) while FA rate is defined by the number of FAs divided by the total number of responses (#FA/(#FA+#Hits). A participant's response was considered to be correct (hit) if he/she pressed the button within two seconds from the offset of the target. Any other button response was considered a false alarm (FA). In contrast, d prime was calculated by subtracting the z transform of the false alarm rate (#FA/(#FA+#correct rejections) from the z transform of the hit rate (# hits/# targets). To calculate the correct rejections, each sound sequence was divided into eight bins. The size of the bin was determined by the time window for calculating hits (2 s).

The examiners suggested that I consider a different way to calculate false alarm (FA) rate, as discussed in the viva. I acknowledge that the calculation of FAs for the Hit minus FA rate was incorrect. Specifically, I used a definition of FA rate as #FA/(#FA+#Hits). The more appropriate definition for FA rate is the following: FA rate = #FA/(#FA+#correct rejections). This holds for every calculation of hit minus FA rate in this thesis. Note that the calculation of FAs for the *d* prime was correct.

Image analysis

Due to individual variability in the patterns of cortical folding and layout of the tonotopic fields of the auditory cortex (Penhune et al., 1996), averaging the normalized image data was not considered appropriate. Therefore image analysis was conducted at an individual subject level. Image pre-processing and analysis were performed on a Sun Ultra 2 computer (Sun Microsystems) using SPM2 software (http://www.fil.ion.ucl.ac.uk/spm/) running in MATLAB v6.5 (Mathworks, Natick, MA, USA). The functional time series were motion corrected to account for head movements both within and between the two runs (Friston et al., 1995), using the central scan as a reference. Head movements did not exceed 3 mm (translation) and 3° (rotation). The anatomical scan was used to transform the functional data into standard 'normalized' brain space using a set of automated algorithms. This step is necessary to report the results in three-dimensional space and to compare between different subjects, as well as with different studies, as brains can be very different in shape (Brett, Johnsrude, & Owen, 2002). There are a number of different standard brain spaces, including the one defined by Talairach and Tournoux (1988), and the one defined by the Montreal Neurological Institute (MNI) (for a description see Brett et al., 2002). The brain space used in SPM is defined by the MNI. The first step was to co-register the anatomical scan with the mean of the realigned functional scans so that they were matching in orientation, using a mutual information algorithm (Collignon, Maes, Delaere, Vandermeulen, & Suetens, 1995; Studholme, Hill, & Hawkes, 1998). The second step was to segment the anatomical scan into its grey, white matter and cerebro-spinal fluid components. This was achieved by an algorithm which first transformed the image to match the T1-weighted anatomical template, and then used cluster analysis with mixture model and *a priori* information about the tissue in which a voxel is located (Ashburner & Friston, 1997). The third step was to spatially transform the grey matter image to match the SPM2 grey-matter template. An automated algorithm minimized the sum of squares difference between the grey-matter image and the SPM2 grey-matter template, first by determining the optimum 12-parameter affine transformation, followed by estimating nonlinear deformations, whereby the deformations are defined by a linear combination of three dimensional discrete cosine transform basis functions (Ashburner & Friston, 1999). Finally, the same transformation parameters were applied to both the anatomical and the functional scans for that individual. The normalized anatomical scan preserved its voxel resolution of 1 mm^3, while the functional scans were upsampled to 2 mm^3. To improve the SNR ratio, functional scans were then smoothed by a Gaussian kernel of 4 mm full width at half maximum (FWHM). The value of 4 mm was chosen because the current study needed fine spatial resolution. Wider smoothing kernels would blur the image data over a larger area, which would not be ideal for the purposes of this study, as the frequency-sensitive regions that are of interest in this study are very small.

Individual time series data were modeled within the framework of the general linear model (GLM). The formula for the GLM is $X=\beta*G+\varepsilon$, where X is the total variability of the MR signal in a voxel, β represents the parameter weights assigned to each condition that would make the error as small as possible, G represents the conditions of the experiment (design matrix), and ε is the error (Huettel et al., 2004). In other words, this equation calculates the optimal value of β for which the G can explain best the total variability (X) so that the error (ε) is eliminated as much as possible. This univariate ANOVA is applied to each voxel separately and so a t-statistic is computed for each voxel. The design matrix (G) is inputted to the model in the form of a text file. In this text file, the different conditions of the experiment are coded in arrays of zeros and ones, so that each row corresponds to each scan of the experiment, and each column corresponds to a condition of the experiment. For example, the design matrix for the passive listening run contained 92 rows by 6 columns; one scan corresponding to the 'just listen high frequency majority' condition would be coded as '1 0 0 0 0 0' and one scan for the 'just listen low frequency majority' condition would be coded as '0 1 0 0 0 0'. Columns 3 to 6 corresponded to the other attend conditions. Each column of the design matrix is called a regressor. Regressors can also code for systematic variations in the MR signal that are associated with factors that are unrelated to the experimental paradigm, such as head movement and between-session differences. Regressors for head movement were not included in the design matrix here because the output of the realignment algorithm showed that there was very little head movement in individual participants. Two regressors, accounting for the differences between the mean signal intensity in the passive listening run and the attend run, were also included in the design matrix. So in summary the design matrix that was specified included eight regressors, one for each of the experimental conditions and two for the mean signal of each run. The silent baseline condition was implicitly modeled. This means that it was not coded as a separate condition. Low-frequency artifacts that are associated with physiological fluctuations such breathing and heart-beat were removed by high-pass filtering the time series. The high-pass filter cut-off is defined by twice the length of the greatest cycle time between two occurrences of the same condition. The high-pass filter cut-off was 192 s for the passive listening run and 320 s for the attend run. The cut-off was different for the two runs because the passive conditions cycled at a faster rate than did the attend conditions. The final step was the estimation of the goodness of fit of the data to the model.

The examiners recommended that the optimal cutoff values (period in s) are usually much lower than several minutes. To calculate my high-pass filter I used a ‘standard’ rule of thumb (i.e. 2*TR(in secs)* maximum number of scans between two repeated presentations of a condition). I note that this is only a rule of thumb. I used a sparse sampling fMRI sequence with a relatively long TR. The above calculation of the value of the high-pass filter cut-off therefore led to one that was extremely high (several minutes). The consequence is that this choice of filter removes very little of the extremely low-frequency temporal noise in the data. A better choice of filter cut-off value would have been a much lower one. However, although the choice of filter affects how much noise there is in the signal, it does not invalidate the activation results. This comment holds true for all high-pass filters used in fMRI analysis of Experiments 1 and 2.

To identify significant responses to the low- and high-frequency sounds, two t-contrasts were performed between the two ‘just listen’ conditions. The first contrast identified high-frequency responses and was defined by the following pairwise comparison ‘just listen high frequency majority>just listen low frequency majority’ (A>B, Table 3.2). The second contrast identified low-frequency responses and was defined by the reverse pairwise comparison (B>A). Voxel responses were considered to be significant only if they exceeded a voxel-level threshold of $p<0.01$ ($t>2.34$), uncorrected for multiple comparisons. Each pairwise comparison involves a great number of individual t-tests, one t-test for every voxel in the normalized functional scan. Multiple comparisons increase the possibility of making false positive errors. A correction is therefore typically applied (e.g. Genovese, Lazar, & Nichols, 2002). This correction is typically used in those fMRI studies that investigate patterns of brain activity across the whole brain that do not have any hypothesis-driven predictions about the expected locus of activation. Alternatively, for those fMRI studies that do have a hypothesis-driven prediction, the number of t-tests can be reduced to the number of voxels within the predicted region, and so it is common practice not to apply the correction for multiple comparisons (Hall & Plack, 2008; Petkov et al., 2004).

	Conditions contrasted	Functional response
A>B	Just listen high-frequency-majority stimulus> Just listen low-frequency-majority stimulus	High-frequency-sensitive response
C>A	Attend high-frequency targets in the high-frequency-majority stimulus> Just listen high-frequency majority stimulus	General enhancement when attending to high-frequency-majority stimulus
C>E	Attend high-frequency targets in the high-frequency-majority stimulus> Attend low-frequency targets in the high-frequency-majority stimulus	Attention-specific enhancement when attending to high-frequency-majority stimulus
A>E	Just listen high-frequency majority stimulus> Attend low-frequency targets in the high-frequency-majority stimulus	General suppression when ignoring the high-frequency-majority stimulus
B>A	Just listen low-frequency-majority stimulus> Just listen high-frequency-majority stimulus	Low-frequency-sensitive response
D>B	Attend low-frequency targets in the low-frequency-majority stimulus> Just listen low-frequency majority stimulus	General enhancement when attending to low-frequency-majority stimulus
F>D	Attend low-frequency targets in the low-frequency-majority stimulus> Attend high-frequency targets in the low-frequency-majority stimulus	Attention-specific enhancement when attending to low-frequency-majority stimulus
B>D	Just listen low-frequency majority stimulus> Attend high-frequency targets in the low-frequency-majority stimulus	General suppression when ignoring the low-frequency-majority stimulus

Table 3.2 Contrasts performed between the six experimental conditions.

Subsequently, these frequency-dependent responses were superimposed on a probability atlas which contains maskers of the three anatomical subdivisions of HG (Te 1.0, Te1.1, Te1.2, Morosan et al., 2001), using an SPM toolbox (Eickhoff et al., 2005). This probability atlas also contains maskers for a number of other brain regions, such as operculum. This method was used to identify the clusters of activation that overlapped with HG. Localization of the peaks of interest also considered the tonotopic scheme reported by Talavage et al. (2000). The first column of data in Table 3.3 shows the peak coordinates that Talavage et al., (2000) reported. Talavage et al. (2000) measured auditory cortical response only in the left hemisphere. In the current study, the assumption was made that the coordinates for the frequency-sensitive regions in the right hemisphere are a mirror symmetric representation of those in the left hemisphere. These coordinates were transformed from Talairach space into MNI space, using a non-linear tranformation created by Matthew Brett and implemented by the Matlab code ‘tal2mni’ (http://imaging.mrc.cbu.cam.ac.uk/imaging/CbuImaing). Both the probability atlas and the peak MNI coordinates (including the standard error of the mean) were used to localize the high- and low-frequency-sensitive regions 1, 2 and 4 in individual brains.

Region	Talairach coordinates for Talavage et al. (2000)	MNI coordinates for Talavage et al. (2000)	MNI coordinates of frequency-sensitive regions for participants #3
1	*x* = -51.9 +/- 1.3 *y* = -16.3 +/-1.9 *z* = 9.0 +/- 0.7	*x* = -52.4 +/- 1.3 *y* = -17.2 +/-1.9 *z* = 8.9 +/- 0.7	*x* = -48 *y* = -20 *z* = 8
2	*x* = -35.5 +/- 1.2 *y* = -18.5 +/-1.8 *z* = 8.5 +/- 0.9	*x* = -35.9 +/- 1.2 *y* = -19.5 +/-1.8 *z* = 8.2 +/- 0.9	*x* = -40 *y* = -20 *z* = 6
4	*x* = -38.0 +/- 1.3 *y* = -34.0 +/-2.3 *z* = 12.0 +/- 1.3	*x* = -38.4 +/- 1.3 *y* = -35.6 +/-2.3 *z* = 11.2 +/- 1.3	*x* = -43 *y* = -32 *z* = 4

Table 3.3 showing peak coordinates for frequency sensitive regions 1,2 and 4, as found by Talavage et al. (2000). Column 2 shows the coordinates of the peak voxel of these regions. The standard error of the mean represents the variability between the six participants of the study. The 3rd column shows the same coordinates transformed in MNI space. Finally, the 4th column shows coordinates for these regions as found in the left auditory cortex of participant #3.

As an alternative to an averaged group analysis, incidence maps were created to show the distribution of activation across participants. This is a descriptive statistic that illustrates the percentage of participants that show response at a particular site in the brain. It is constructed by summing individual, thresholded statistical maps, typically thresholded at $p< 0.05$ (Keilholz, Silva, Raman, Merkle, & Koretsky, 2004), $p<0.01$ (Hall & Plack, 2008) or $p < 0.001$ (Hall, 2005), uncorrected for multiple comparisons. In the present study, a probability threshold of $p < 0.01$ was chosen because it contributed information about the distribution of frequency-related activation for every participant. This method has the advantage that it does not obscure the variability across participants.

To explore the effects of selective attention within the frequency-sensitive regions (regions 1, 2 and 4), the beta values were plotted for the six conditions using i) for a single participant, the peak voxel in those regions and ii) for all participants, the average of all voxels in a region. Remember that the beta value represents the parameter estimate of the contribution of each listening condition to variability of the MR signal.

3.2.2 Results

Target detection results

Mean performance was poor (Figure 3.8). The mean value of d prime never exceeded a value of 1, which corresponds to 69% correct for both *yes* and *no.* trials. Performance was particularly poor when measured using the hit-FA rate since the average hit-FA rate was always below zero.

Individual performance was variable. Two of the participants (#1 and #2) responded to targets in both frequency streams, irrespective of the task instructions suggesting that they did not perform the task correctly. Across all conditions the average false alarm rate (56%) was actually greater than the average hit rate (35%). One interpretation of this low performance is that the training was not adequate. Another interpretation is that the task was just too difficult and no amount of training would benefit performance. It is interesting to note that despite the best performance, participant #3 reported some difficulty following the task instructions. In particular, after the experiment, the participant reported that the word ‘high’ in the instruction ‘high sounds’ could be interpreted as high probability of occurrence, referring to the majority stream, instead of high frequency. Despite this comment, this was not reflected in the pattern of performance, but it was taken into account when modifying later versions of the task.

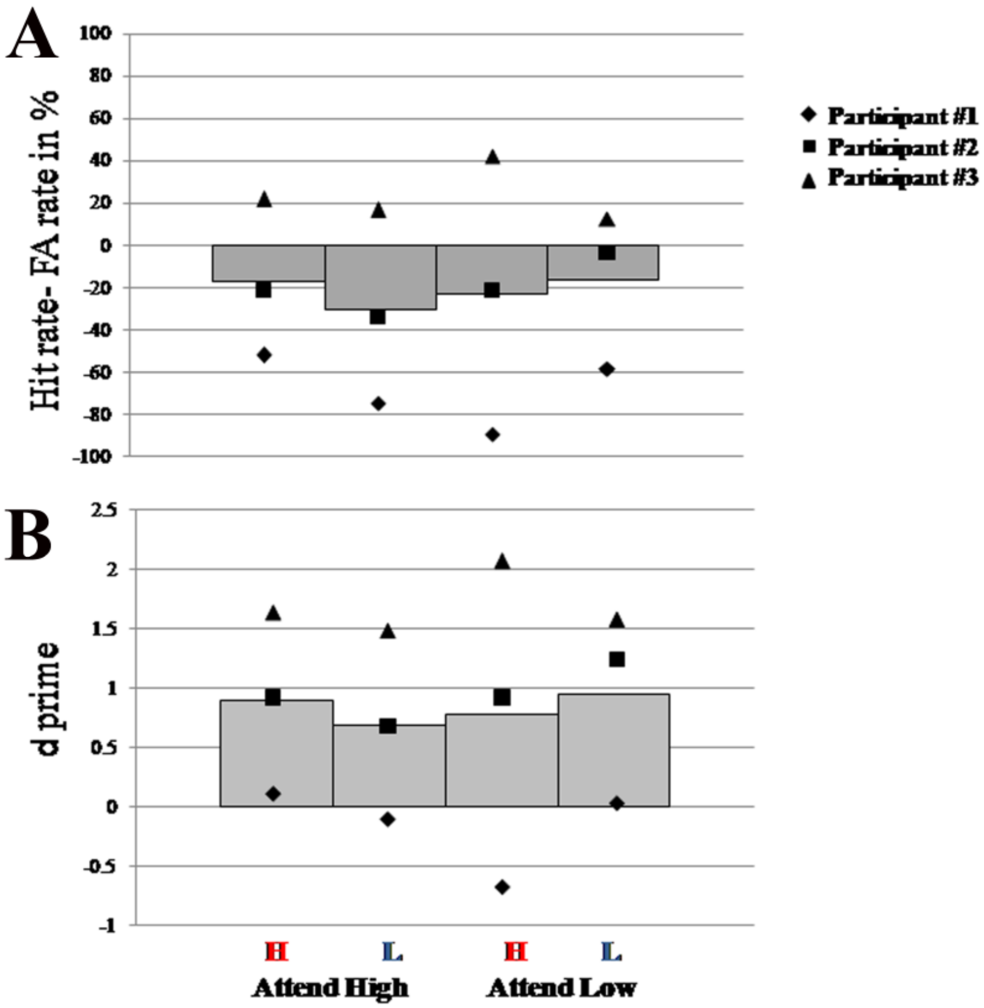


Figure 3.8 (A) hit rate-FA rate in % and (B) d prime of target discrimination performance for the four 'attend' conditions in pilot study 1. Bars indicate average across participants, symbols indicate individual performance. H: high-frequency-majority stimulus. L: Low-frequency-majority stimulus.

Spatial specificity of frequency-sensitive responses

The distribution of low- and high-frequency-sensitive responses that are shown by the incidence map broadly agrees with the scheme reported by Talavage et al. (2000). Figure 3.9 shows this distribution in four horizontal slices through HG. High-frequency-sensitive regions (shown in red) were located more medially in HG. Peaks corresponding to those reported by Talavage et al. (2000) are numbered 1, 2 and 4 in Figure 3.9. Low-frequency-sensitive regions (shown in blue) were located more laterally in HG. Peaks corresponding to those reported by Talavage et al. (2000) are numbered 1 in Figure 3.9. The incidence map did not show much evidence for overlap between participants. It is not clear from the incidence map whether or not adjacent voxels represent a frequency-sensitive response from the same or from different participants. The impression gained from visual inspection of the individual results is that individuals activate small and coherent clusters of voxels rather than a mosaic of isolated voxels. However the pattern of responses is driven by participant #3, and to a lesser extent by participant #1. Participant #2 did not show any low or high frequency-sensitive regions at this particular probability threshold. The individual extents of activation are reported later in Table 3.4.

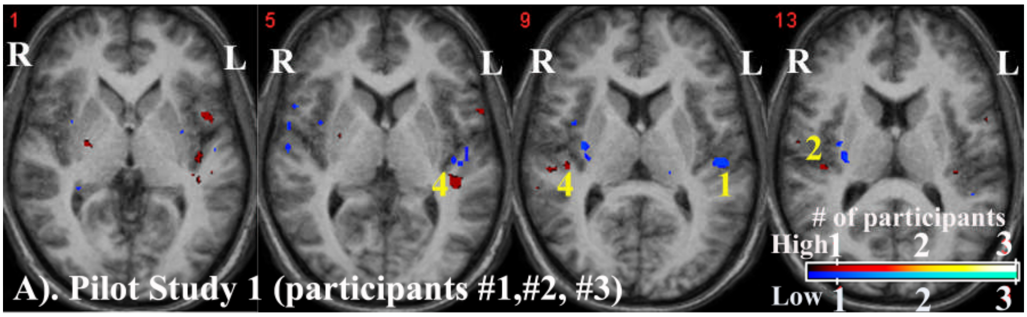


Figure 3.9 Incidence maps showing frequency-sensitive response around the auditory cortex in pilot study 1 Incidence maps are overlaid on the average anatomical image of (A) participants #1 to #3 and (B) participants #7 to #13, and shown in four axial slices. The z value shown on the top left of each slice The high-frequency regions are depicted in red and low-frequency sensitive regions in blue. Orientation: R to L.. Numbers on the figures represent the high- (2 and 4) and low-sensitive (1) regions that Talavage et al. (2000) identified.

Attentional modulation

First, the beta values were plotted for peak voxels in regions 1 and 2 in the left hemisphere for participant #3 (Figure 3.10). This example was chosen because the results were representative of the other hemisphere and the other two participants.

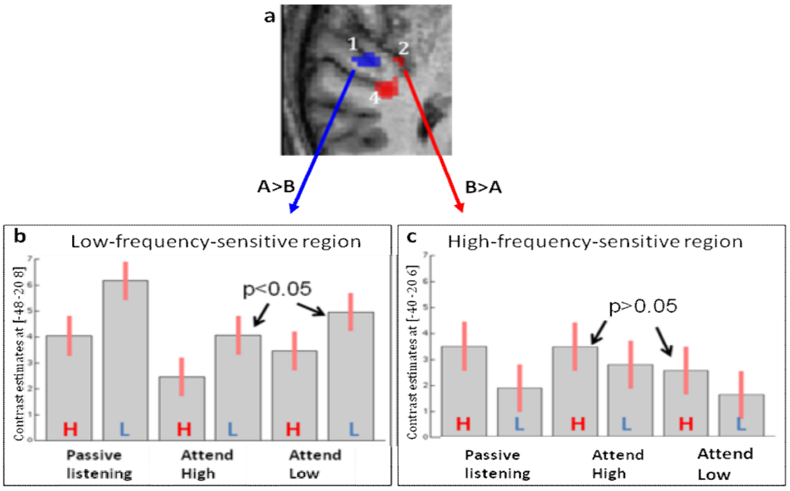


Figure 3.10 (A): Axial view (z=+6) of left part of the anatomical scan of participant #3, upon which are shown two high frequency sensitive regions (in red) and one low-frequency sensitive region (in blue). The contrast performed for the high-frequency-sensitive regions was: just listen high-frequency majority stimulus>just listen low-frequency majority stimulus (A>B). The reverse contrast was performed for the low-frequency-sensitive regions (B>A). These clusters are partly located in one of the subdivisions of HG (according to the probability maps (Eickhoff et al., 2005). The grey bars in the (B) and (C) show the beta values (arbitrary values) for each of the six conditions for the peak voxels of region 1 (B) and region 2 (C). The red bars represent the 90% confidence intervals.

In region 1 (Figure 3.10b), there was a greater response when participants were instructed to attend to the low-frequency targets in the low-frequency-majority stimulus (column 6) than when attending to the high-frequency targets in the same stimulus (column 4). The difference between them was significant at $p<0.05$. This attention-specific enhancement is consistent with the attentional enhancement hypothesis. Also consistent with the general suppression hypothesis, the results in region 1 showed a reduced response when participants were instructed to attend to the high-frequency targets in the low-frequency-majority stimulus (column 4) than when they were instructed to passively listen to the same stimulus (column 2). This difference was significant because the error bars are completely non-overlapping. However, an important finding that is inconsistent with the attentional enhancement hypothesis was that attending to the low-frequency targets in the low-frequency-majority stimulus (column 6) generated a smaller response than passively listening to the same stimulus (column 2). The attentional enhancement hypothesis would predict that attending to the low-frequency target increases the response in low-frequency-sensitive regions compared to both passive listening (general enhancement) and attending to the high-frequency targets in the same stimulus (attention-specific enhancement). For the high-frequency region 2 (Figure 3.10c), there was also a trend for a greater response for attend high-majority (column 3) than for attend low-frequency targets in the high-frequency-majority stimulus (column 5) but the response size was smaller and so these differences were not significant.

The second exploration of attentional modulation plotted the beta values for regions 1 and 2 across all three participants. The centre of region 1 was defined by the coordinate x=-46, y=-20 z=7 mm which showed overlap in two out of the three participants. The centre of region 2 was defined by the coordinate x=-42, y=-19 z=1 mm which showed response only for one participant. The beta values were extracted for all voxels contained within a sphere centred at this peak and with a 6 mm radius. Figure 3.11 shows the mean results and individual responses for these two regions plotted across the six experimental conditions. The pattern of attentional suppression was generally very similar to that seen previously for the peak voxel in participant #3, most notably in region 1 (shown in Figure 3.11B). In addition, the previous inconsistency with the attentional enhancement hypothesis was also seen here, namely there was a greater response when participants passively listened to the low-frequency-majority stimuli (column 2) compared to when they were detection low-frequency targets in the same stimulus (column 6). None of these effects reached significance, because there was a lot of variability across participants. In the high-frequency region, there was very little effect of the listening instructions on the magnitude of the response. The results showed less variability than in the low-frequency region.

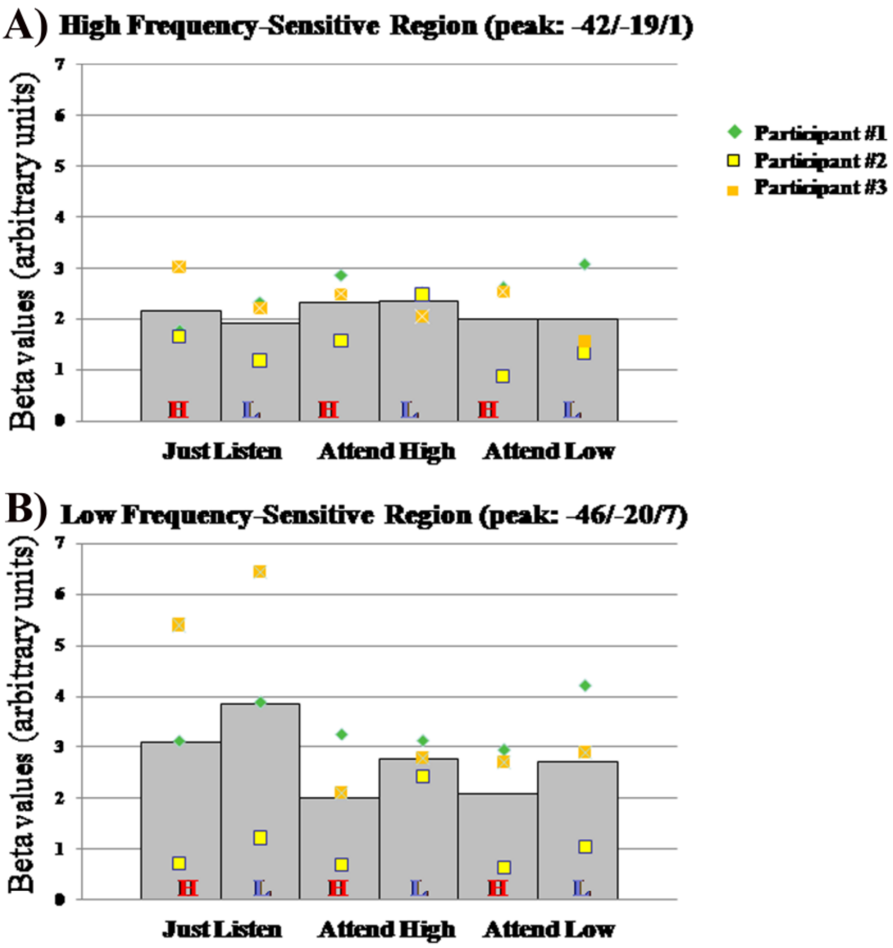


Figure 3.11 Beta values for all conditions in (A) high-frequency-sensitive region 2 and (B) low-frequency-sensitive region 1 for all three participants. Grey bars represent the average value, while the symbols indicate the values for each individual participant.

It is interesting to comment on the lack of association between target-detection performance and the effect of attention on the size of the response in regions 1 and 2. For example, participant #3, who obtained the best performance scores, showed the greatest reduction in the low-frequency-sensitive region for the ‘attend’ compared to the ‘just listen’ conditions.

3.2.3 Summary

Pilot study 1 had four aims. The first aim was to design a mixed-frequency stimulus that was suitable for separating high- and low-frequency-sensitive responses in the auditory cortex. While high- and low-frequency-sensitive regions were identified around HG, there was not very consistent overlap between individuals and participant #2 did not evoke reliable frequency sensitive activation. It is possible that the voxel resolution used to acquire the data was too coarse to detect the small volume of the frequency-sensitive activity. Higher voxel resolution can assist to avoid partial volume effects; a voxel that is relatively large in volume could contain tissue that does not contribute to the MR signal, which results in reducing SNR (Huettel et al., 2004). The second aim was whether the spatial location of these responses matched those found in previous studies of tonotopy. This was confirmed in participants #1 and #3.

The third aim was to investigate whether the task was effective for manipulating selective attention. The answer to that from pilot study 1 was ’no’, because two out of three participants did not seem able to conform to the task instructions and there were a relatively high number of false alarms. The reason was either because the training was inadequate or that the task was too difficult. Inadequate training could also be responsible for the lack of understanding of the task instructions. The issue of poor performance is investigated in pilot study 2.

The fourth aim was to investigate whether there was attentional modulation of the auditory responses when attending to different sound frequencies. In the high-frequency regions of the three participants, both feature-specific enhancement and suppression were generally confirmed. The low-frequency regions showed quite unexpected results. Specifically, there was a smaller response for the ‘attend’ conditions compared to the ‘just listen’ conditions. This result could reflect an effect of presentation order. For all participants, the ‘just listen’ conditions were presented in the first run of the fMRI experiment and the ‘attend’ conditions were presented in the second run. A smaller response in the second run than the first run might reflect adaptation to the sound stimuli over time, unrelated to the listening instructions. A solution to this problem would be to fully randomize the order of the conditions.

3.3 Pilot study 2: Task performed by expert listeners

To explore the issue of task performance three normal-hearing (≤25 dB for 250 to 8000 Hz) expert listeners (two females, mean age= 35, age range=29-40) were tested using the same stimuli and the same listening instructions as in pilot study 1. The task was performed in the MR scanner, but in the absence of scanning. The three expert listeners were myself (#4), a listener that was highly trained in psychoacoustical experiments (#5) and my first supervisor (#6). Participants #4 and #6 were already familiar with the task, while participant #5 was given five sound sequences for each ‘attend’ condition as training. The conditions were presented in a randomized order.

The button presses were analyzed in the same way as in pilot study 1 using both hit-FA rate and d prime measures. All three participants showed much improved performance compared to participants #1 to #3. For example, the performance of participants #4 to #6 had d primes above 1.5, while the performance of participants #1 and #2 had d primes consistently below 1 (see Figure 3.12B). The average hit rate for the four conditions was 92, 78, 83 and 65 % , while the average FA was 14, 11, 12 and 22 %. These results showed that the task was difficult, but could be performed satisfactorily by expert listeners. It is interesting to note that participants #5 and #6 seemed to score considerably lower in the condition ‘attend low-frequency targets in the low-frequency majority stimulus’ than participant #4 and this explains the lower mean score (see Figure 3.12A).

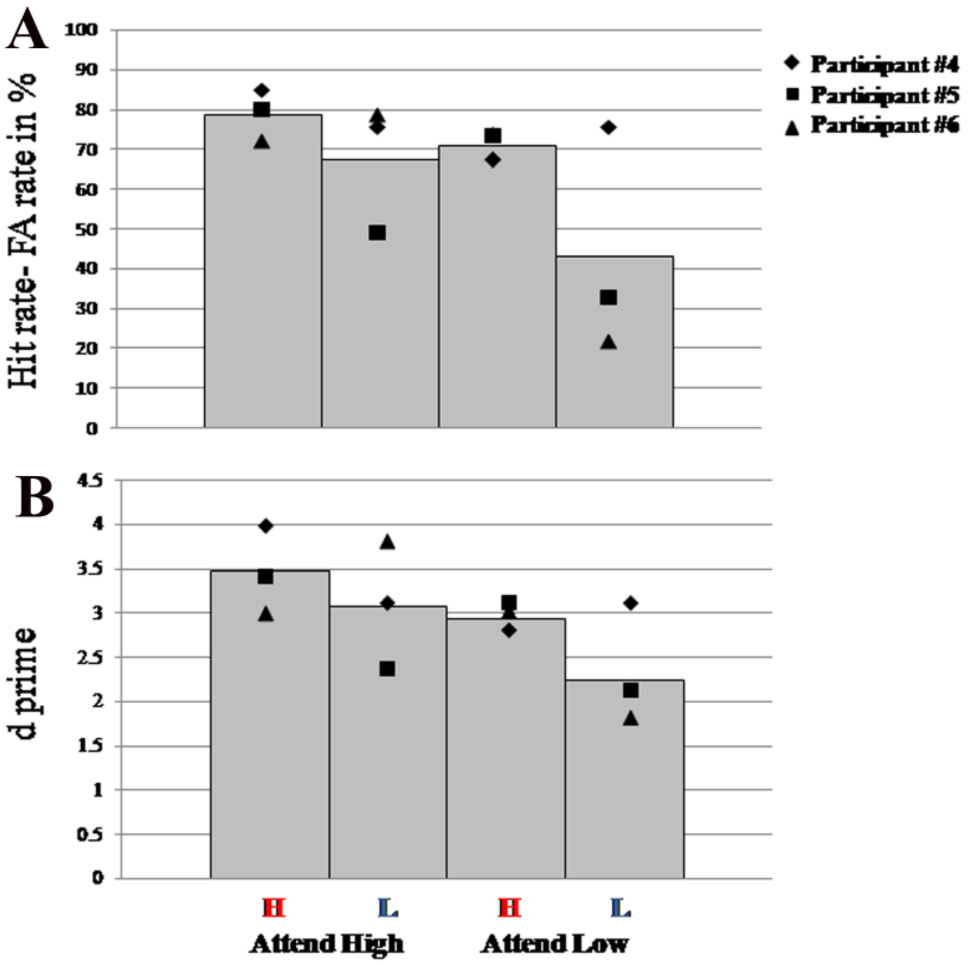


Figure 3.12 (A) hit rate-FA rate in % and (B) d prime of target detection performance for the four ‘attend’ conditions in pilot study 2. Grey bars indicate average across participants, dots indicate individual performance.

3.3.1 Summary

In pilot study 2, expert listeners were tested in the four ‘attend’ conditions to investigate the effect of expertise in task performance. Performance was sufficiently high, indicating that the tasks are possible to perform. Therefore, the low performance in pilot study 1 was possibly due to the fact that naïve listeners needed more explanation and more effective training on the task. This indicates that for naïve participants it is very important to improve training, by having many examples and more training sessions, and perhaps also a Power-point presentation which clearly explains the task. Pilot study 2 confirms that, although the task is difficult, it is not impossible to do well, if one understands the task. This result led me to incorporate a comprehensive training scheme in Experiment 1 before the scanning began.

3.4 Pilot study 3: Activation mapping using a high voxel resolution (1.5x1.5x2.5 mm)

Pilot study 3 had three aims. The first aim was to determine the optimal TE in the grey matter of HG when scanning using the high resolution sequence. The largest BOLD signal changes occur when the value of TE is approximately equal to the value of T2*. As stated previously, the value of T2* differs over anatomical regions (Clare et al., 2001). Due to the reduction in the field of view, the front and back of the head extended beyond the image matrix and so Nyquist artefacts were visible in the scan. To remove them, it was necessary to use saturation bands at the front and back of the head to reduce the MR signal from those brain regions (e.g. nose, eyes, scalp) at the edge of the field of view (Figure 3.6B). The saturation bands were defined by a 'pre-pulse' EPI (echo planar imaging) sequence. This sequence included an initial radiofrequency pulse that induced a 90-degree transverse magnetization component in the tissue that fell within the saturation band. So when applying the second 90º radiofrequency pulse during the acquisition of the functional scanning the 'pre-excited' tissue now had a 180º transverse magnetization component that was not picked up by the receiver coil and so had a zero MR signal. Adding two saturation bands increased the acquisition time.

The second aim of pilot study 3 was to investigate the statistical reliability of sound-evoked activation at a high voxel resolution. The primary auditory cortex covers a small area (~1-4cm^3) and its position highly variable across individuals. Finer spatial resolution should enable better separation of iso-frequency bands within a tonotopically organized region as well as a better separation of low-frequency-sensitive regions 1a and 1b which, in pilot study 1, could not be reliably separated within an individual participant. Furthermore, the iso-frequency bands across this region are narrow, certainly much smaller than the voxel resolution. It is therefore important to be able to ‘zoom in’ as much as possible, so that the voxel captures the spatially specific response to individual frequencies. Additionally, finer spatial resolution is advantageous in avoiding partial volume effects, as the larger the voxel, the more possible it is that it contains tissue that does not contribute to the MR signal, thus reducing SNR. However, high voxel resolution is not always advantageous, because it increases the total acquisition time of the pulse sequence. Furthermore, it can result in poorer SNR ratio due to the smaller volume of tissue being sampled. To investigate the effects of the standard versus the high voxel resolution, the extent and magnitude of the response in pilot studies 1 and 3 were quantified and compared.

The third, brief aim was to confirm whether general performance for naïve listeners could be improved by providing extensive training before the ‘attend’ run.

3.4.1 Methods

Participants

Seven participants were scanned (4 males and 3 females, mean age=22 years old, range 19-29). These participants are referred to as numbers #7 to #13. They had normal hearing (≤25 dB for 250 to 8000 Hz). None of the participants had a history of audiological or neurological impairment. Three more participants (#I to #III) were scanned, but their data were excluded from further analysis. For two of the participants (#I and #II), target-detection responses were not recorded due to technical problems. The third participant (#III) did not make any responses in the 'attend' conditions because, as they reported afterwards, the 'just listen' instruction in the 'silent' condition was taken as an instruction for all subsequent sound sequences.

Stimuli and task

Pilot study 3 was based on the design of pilot study 1. Therefore, only the parameters that differ will be reported. A greater emphasis was placed on training following the poor performance in pilot study 1. The length of the training session depended on each participant's needs and ranged from 10 to 25 mins. To supplement the verbal instructions, participants were presented with a PowerPoint presentation which explained in detail the sound conditions and the tasks using written instructions, diagrams and sound examples. Participants completed several practice runs for each of the four 'attend' conditions. Only when participants responded to more than half the number of targets in a practice run, did they proceed to the ‘attend’ run.

The visual instructions that informed participants about which task to perform during each stimulus condition were changed from ‘high sounds’ to ‘high pitch’ and from ‘low sounds’ to ‘low pitch’. This was done to avoid the possibility of participants confusing the word ‘high’ (referring to high-frequency stream) with the word ‘high’ referring to the high-majority sequence. Part way through pilot study 3, it also became apparent that the instruction to ‘just listen’ during the silent condition was misleading since one of the participants (#III) assumed that all the subsequent ‘attend’ conditions did not require any response. For this reason, from participant #9 onwards, the only cue that appeared on screen during the silent intervals was a fixation cross.

Imaging protocol

The imaging protocol was the same as that reported in pilot study 1, except for the following changes. The voxel resolution was reduced from 3x3 mm in plane to 1.5x1.5 mm, and from 3 mm through plane to 2.5 mm. The orientation of slices also changed. While in pilot study 1 the slices were vertical to the axis of the Sylvian fissure (Figure 3.6A), in pilot study 3 they were parallel to it (Figure 3.6B). This latter orientation maximizes the amount of HG included in the smaller field of view. Using a high-resolution sequence with the added saturation bands (see Figure 3.6B), increased the acquisition time of one scan from 1852 to 2118 ms (Figure 3.7). To avoid, as much as possible, overlap of the scanner noise at the beginning of each sound sequence, a delay of 1.5 s was added before the sound onset. With this arrangement, the sound sequence and the scanner noise overlapped only by 618 ms. So that the scanner noise did not overlap with the end of each sound sequence, the TR was increased from 8200 to 9000 ms. This choice of TR also increased the duration of each session by a few mins: it lasted 15 mins for the ‘just listen’ session and 25 mins for the ‘attend’ session. Note that for participant #7, the choice of delay and TR was as for pilot study 1 in which the TR was 8200 ms, and there was no delay. To assist in the pre-processing of these high resolution data which encompassed a smaller field of view across the brain, the number of slices acquired in the single T2*-weighted functional scan was increased from 36 to 60. The choice of TE for the single T2*-weighted functional scan was also decreased from 35 to 24 ms to increase MR signal intensity.

For the T2*-weighted ‘just listen’ and ‘attend’ runs, the data for participant #7, were acquired using TE=35 ms as in pilot study 1. Following the analysis and calculation of optimal TE, for participants #8 onwards, the data were acquired using the optimal TE (55 ms). The data used for calculating the optimal TE, were obtained from participant #I. Six 18-slice scans were acquired without any sound stimulation, using a different TE value. The order of acquisition was the following: TE=24, 64, 34, 54, 44 and 74 ms. During image pre-processing, the scans were realigned according to the TE=24 ms scan.

Image analysis

The pre-processing was the same as in pilot study 1. The normalized scans maintained their original resolution of 1.5x1.5x2.5 mm and were smoothed using a Gaussian kernel of 3 mm FWHM. The general linear model (GLM) analysis was as in pilot study 1. Pairwise t-contrasts were computed to identify the low- and high-frequency sensitive responses. From these contrasts, incidence maps for participants #7 to #13 were formed to investigate the overlap in terms of sensory response to the two sound frequencies (Figure 3.9C) ($p<0.01$, uncorrected). To investigate attentional modulation, the % MR signal change was plotted for voxels within two types of regions: high-frequency-sensitive regions ('just listen high-frequency majority>just listen low-frequency majority', A>B) and low-frequency-sensitive regions (B>A). Percentage (%) MR signal is a metric used in other studies (such as in Tootell et al., 1998 see Figure 3.2). It is a more direct measure of the BOLD response and easier to interpret than the beta values. Marsbar software (http://marsbar.sourceforge.net/) was used to extract the mean time series for clusters that had at least 10% probability of belonging to one of the three anatomical subdivisions of HG (Te1.1, Te1.0, Te1.2) as revealed by the probability maps (Eickhoff et al., 2005). More usually than not the clusters extended to other areas as well, such as the Operculum.

3.4.2 Results

Target detection results

Mean performance was improved compared to pilot study 1 which probably reflects the better training the participants received (Figure 3.13). Specifically, in pilot study 1 mean performance was between -30 and -8 % hit-FA rate (Figure 3.13A), and between 0 and 1 d prime (Figure 3.13B). In this study, it was between -20 and 40 % and 1.5 and 2.5 d prime, respectively. The only condition that did not show any reliable change in performance was ‘attend low-frequency targets in low-frequency majority stimulus’ (F). However, despite the general improvement, performance was still very low (especially in terms of hit-FA rate). Performance was highly variable between individuals, especially for the two ‘attend low’ conditions. One of the reason is that participants #9 and #10 have a lot of false alarm responses and very small number of hits (hence the negative performance). Interestingly, the two different measures of performance showed a different pattern of results, possibly because the hit-FA rate does not always produce a bias-free estimate of sensitivity. The d prime measure showed that the performance for the conditions ‘attend high-frequency targets in low-frequency majority stimulus (D) and ‘attend low-frequency targets in high-frequency majority stimulus (E) was broadly the same as for the condition ‘attend high-frequency targets in high-frequency majority stimulus’ (C), while the hit-FA rate measure for the same conditions showed that it was worse.

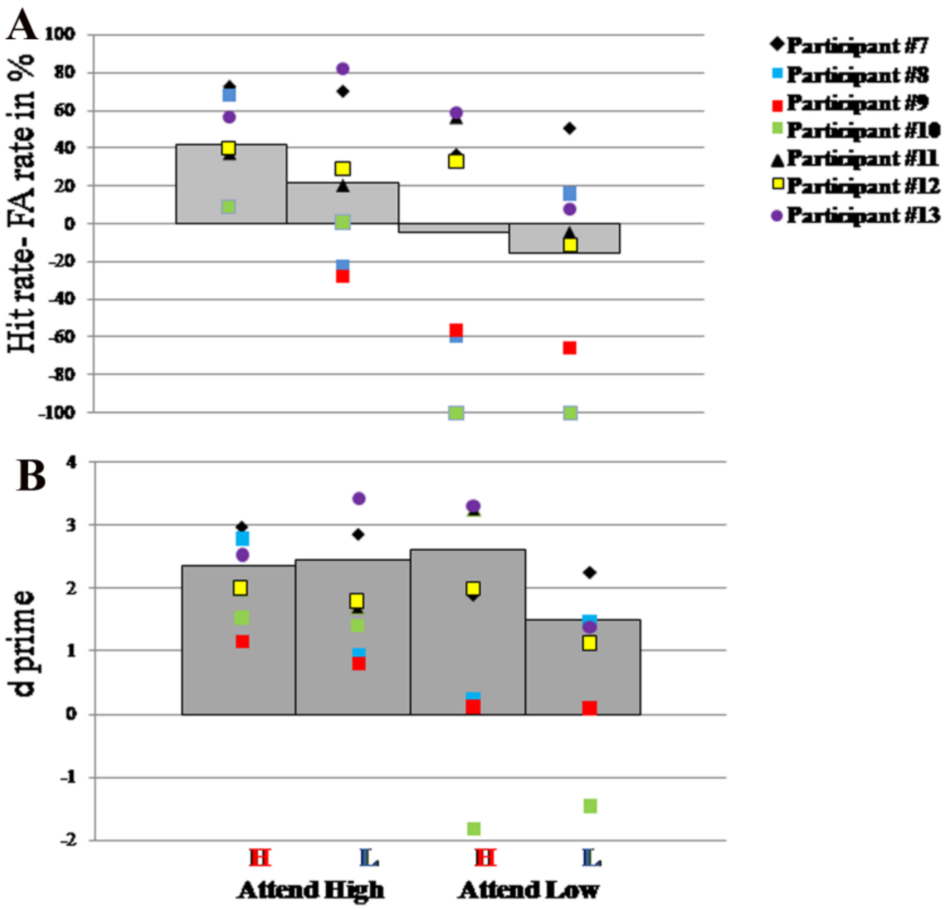


Figure 3.13 (A) hit rate-FA rate in % and (B) d prime of target detection performance for the four 'attend' conditions in pilot study 3. Bars indicate average across participants, symbols indicate individual performance. H: high-frequency-majority stimulus. L: Low-frequency-majority stimulus.

Frequency-sensitive response and attentional modulation

Figure 3.14 shows the incidence maps for the high- and low-frequency-sensitive regions for the seven participants of pilot study 3. Slice z=+9 shows a bilateral response within regions 1, 2 and 4. The high-frequency-sensitive regions (2 and 4) occur in the medial part of HG, and the low-frequency sensitive region (1) occurs in the lateral part of HG. The location of these frequency-sensitive regions agrees with previous studies. The results show a high degree of overlap, up to five participants in the high-frequency region 2.

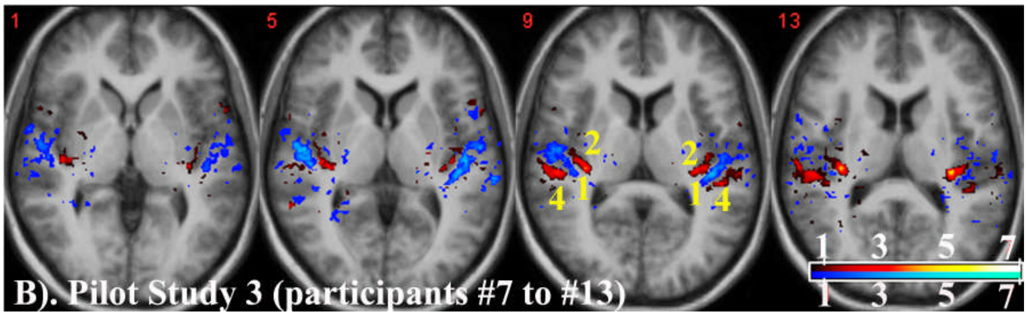


Figure 3.14 Incidence maps showing frequency-sensitive response around the auditory cortex in pilot study 3. Incidence maps are overlaid on the average anatomical image of participants #7 to #13, and shown in four axial slices. The z value shown on the top left of each slice The high-frequency regions are depicted in red and low-frequency sensitive regions in blue. Orientation: R to L. Numbers on the figures represent the high- (2 and 4) and low-sensitive (1) regions that Talavage et al. (2000) identified.

To investigate attentional modulation, the average % MR signal was plotted for voxels within high- and low-sensitive regions for each participant. Percentage (%) MR signal change was calculated for each condition against the relevant silent condition using the following formula: (signal+silence)/silence*100. To satisfy the normality requirements for parametric tests, the % MR signal change for each run was first log transformed. Statistical analysis (in SPSS) was then performed, to determine whether the differences between the conditions were significant. If the data satisfied parametric tests for normality (Shapiro-Wilk test of normality for the residuals) and for homogeneity (Levene test), then univariate analysis of variance was performed. This analysis of variance had six levels corresponding to the six listening conditions. Tukey post-hoc tests were then performed to test for significant differences between conditions (adjusted significance $p<0.05$). If the assumptions of the parametric tests were not met then a non-parametric Kruskal-Wallis test was performed. Mann-Whitney U post-hoc tests were performed to test for significant differences between the conditions (with

Bonferroni adjustment for multiple comparisons, so the significance level was 0.05/4=0.01). The four post-hoc tests will be described for a high-frequency region. The same principle applies to the low-frequency regions. The letters refer to the contrasts summarized in Table 3.2.

i)To validate the frequency-sensitivity contrast performed by SPM (i.e. A>B), the same contrast was performed by SPSS.

ii)To investigate the effect of general enhancement and attention-specific enhancement, two tests were performed between the three high-majority conditions: C>A and C>E respectively.

iii)To investigate the effect of attentional suppression, one test was performed: A>E.

Representative examples of the % MR signal change within three frequency-sensitive regions of one participant (#7) will be discussed. Although this participant was not scanned using the optimal TE of 55 ms and thus the data may not have the best SNR ratio, this should not influence the overall pattern of frequency-sensitive response nor the pattern of attentional modulation. The three examples are illustrated in Figure 3.15. These are the low-frequency-sensitive region 1 in both hemispheres and the high-frequency-sensitive region 4 in the left hemisphere. As in pilot study 1, there was some evidence for a greater response for the ‘just listen’ conditions than for the ‘attend’ conditions. For example, the response to the low-frequency-majority stimulus in region 1 of the right hemisphere was significantly smaller when participants were attending to targets than when they were passively listening. The same pattern was observed for high-frequency region 4 in the left hemisphere. Again, these results are inconsistent with the hypothesis of attentional enhancement and indicate that there might be adaptation effects due to presentation order.

There were also significant effects between the ‘attend’ conditions. In region 4 of the left hemisphere there was a greater response when attending to high-frequency targets in the high-frequency-majority stream than when attending to the low-frequency targets in the same stimulus (p<0.007). This trend was also seen in both regions 1, but it did not reach significance. This pattern could be consistent with the attentional enhancement hypothesis. However, due to the fact that responses for both of these ‘attend’ conditions were lower than for the ‘just listen’ condition, makes the interpretation unclear.

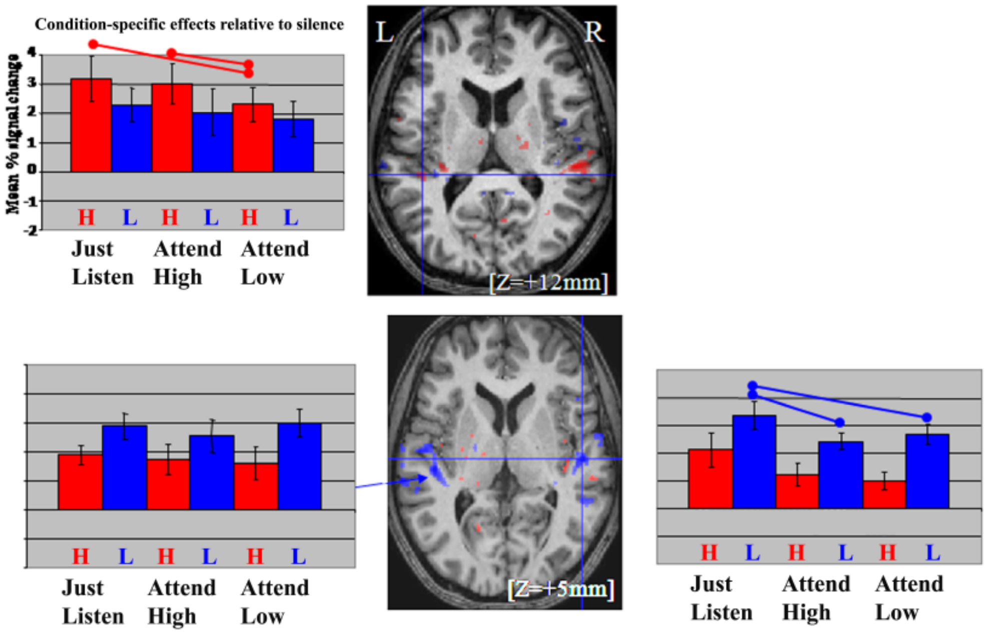


*Figure 3.15 % MR signal change in high-frequency and low- frequency regions for participant #7. The functional results were superimposed on the normalised anatomical scan of each participant. The orientation of all slices is left to right. The bars show the mean % signal change for each condition against silence. The vertical lines are 95% confidence intervals((standard deviation*1.96)/(square-root of number of observations)). The red lines above the bars indicate that there is a significant difference between the high-frequency majority sound conditions (p<.05 for parametric tests, and p<.007 for non-parametric tests), while the blue lines indicate a significant difference between the low-frequency majority sound conditions.*

Calculation of the optimal TE

To obtain optimal signal, it was important to determine the optimal TE for the high-resolution sequence. The optimal TE of a region is approximately equal to the T2* relaxation time (Clare et al., 2001). To calculate T2* relaxation within the auditory cortex, three grey matter regions of interest (ROIs) were specified from three different slices. Two of the ROIs were approximately in HG, while the third ROI was in the left STG (Figure 3.16). Subsequently, the mean MR signal was extracted from each ROI for each TE value.

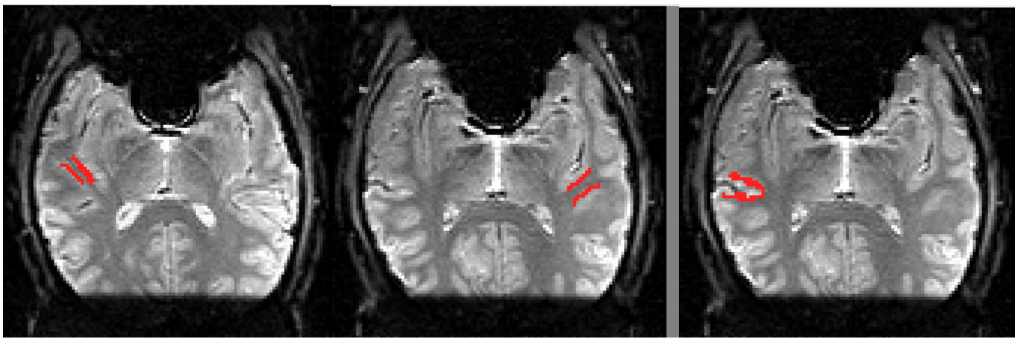


Figure 3.16 Axial/oblique slices of T2-weighted functional scans of participant #I. The red regions indicate the regions of interest (ROIs) selected for analysis.*

The MR signal was transformed to a natural logarithm (LN) and these were plotted as a function of TE on a single graph (Figure 3.17). A straight line was fitted to these data and the gradient of the line was determined (y=ax+b, where a is the gradient and b is the intercept). The T2* of the tissue is calculated as the following: T2* = -1/(gradient). For these three ROIs, the T2* was 60.2 indicating that an optimal TE is about 60 ms. However at a TE of 60 ms transverse magnetization component of the MR signal has decayed close to zero resulting in a darkened brain image Figure 3.5). As a compromise, a TE of 55 ms was chosen.

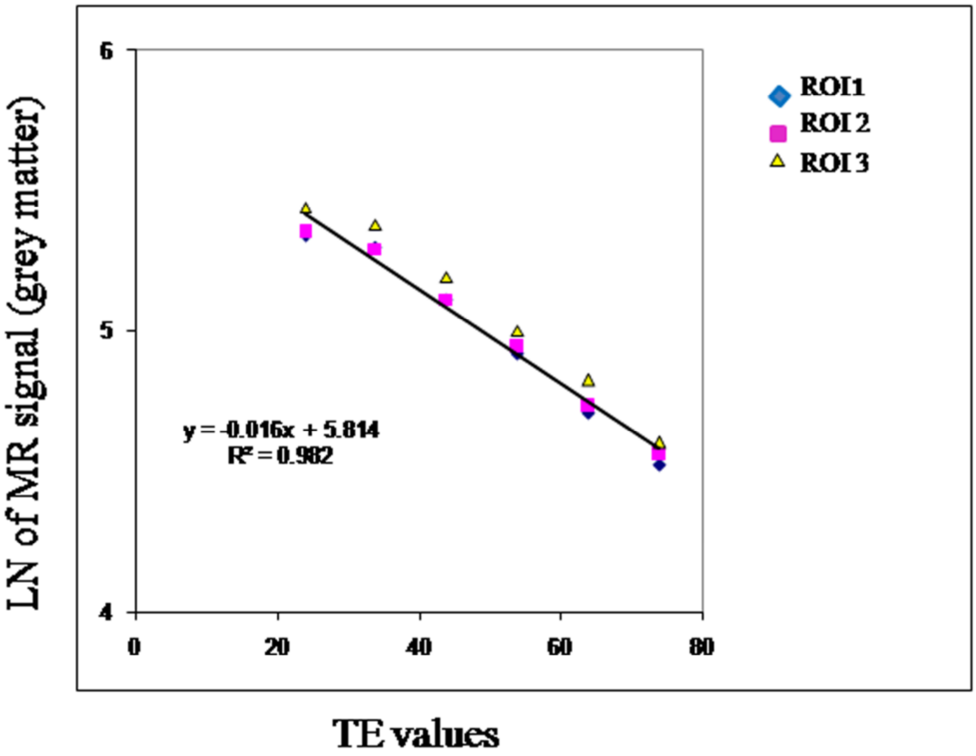


Figure 3.17 Plot of natural logarithm (LN) for the three ROIs across echo times (TEs).

The examiners noted that the value of T2* estimated from the MR data (the reciprocal of the slope of the regression line) was surprisingly high. I have discussed this issue with the MR physicists at the MR Centre and it appears that this is a data processing error. The default setting on the software that converts the raw data (PAR/REC formats) to image data (IMG/HDR formats) applies a scaling to the pixel intensities. This automatically adjusts the scale between the different TE scans. The optimal TE value should have therefore been lower than 60 ms and as a consequence the scanning parameters for Experiment 1 were not optimised for BOLD signal. This does not invalidate the results of the analysis, but makes them less sensitive to the signal of interest.

Comparison between pilot studies 1 and 3

To investigate the suitability of the high-resolution pulse sequence for addressing the experimental hypotheses, the distribution, the statistical reliability, the extent and magnitude of the frequency-dependent responses were compared to the data in pilot study 1, acquired using the standard resolution pulse sequence.

To examine the distribution of the frequency-sensitive responses, the incidence maps for pilot studies 1 and 3 (shown in Figure 3.9 and Figure 3.14) were visually compared. The centers of region 1, 2 and 4 were generally in the same place around HG. In pilot study 3, the incidence maps were more extensive, although this could be due to the greater number of participants.

To quantify the difference in sensitivity between the two pulse sequences, the statistical reliability of the frequency-sensitive response around the primary auditory cortex was measured (Table 3.4). Two measures of statistical reliability were the number of suprathreshold ($p>0.01$, $t>2.34$) voxels in an activated cluster and the maximum t value within that cluster. A cluster was included in Table 3.4 if it satisfied two additional criteria; i) a probability of at least 25% of the cluster lying within one of the three anatomical subdivisions of HG and ii) at least four suprathreshold voxels. The number of voxels was converted to a volumetric measure to equate the different voxel resolutions. Participant #7 was excluded from the analysis because he/she was not scanned using the optimal TE.

Table 3.4 shows that the frequency-sensitive response was detected much more reliably using the high-resolution sequence than the standard-resolution sequence. The average volume of the high-resolution sequence was 1173 mm^3 for the high- and 906 mm^3 for the low-frequency-sensitive regions, while for the standard-resolution sequence the averages were 0 mm^3 and 117 mm^3 respectively. The average t value for the high-resolution sequence was 11.5 for the high- and 5.5 for the low-frequency-sensitive regions, while for the standard-resolution sequence it was 0 and 3.5 respectively. It is also interesting to note that for the standard-resolution sequence, there were no high-frequency-sensitive responses for any of the three participants and no suprathreshold activation at all for participant #2. In contrast, for the high-resolution sequence, no significant high-frequency response was obtained in only one out of the six participants and all participants showed some suprathreshold activation.

Participant Nr	Cluster size (mm^3)	t value	Hemisphere	Cluster size (mm^3)	t value	Hemisphere
High-frequency-sensitive regions				Low-Frequency-sensitive regions		
Standard resolution sequence (3x3x3 mm)						
#1	0	-		152	5.1	L
#2	0	-		0		
#3	0	-		112	3.3	R
				240	2.7	R
				236	4.6	L
Average	0	0		117	3.5	
High resolution sequence (1.5x1.5x2.5 mm)						
#8	1671	4.1	R	186	4.1	L
	1457	7.7	L			
	551	4.9	L			
	124	3.5	L			
#9	158	3.7	L	208	4.1	R
				332	3.9	L
				141	3.6	L
#10	158	4.5	L	270	4.7	L
	309	4.8	R			
	186	3.6	R			
#11	951	7.8	R	698	6.7	R
	574	7.2	L	366	4.7	L
	574	8.2	L			
#12	0	-		900	4.9	L
				849	4.6	R
				214	3.8	L
				101	2.9	R
				51	3.3	R
#13	203	4.5	R	613	4.8	R
	124	4.5	R	84	3.7	R
				264	7.3	L
				118	4.6	L
				39	2.9	L
Average	1173	11.5		906	5.5	

Table 3.4 High- and low-frequency-sensitive responses, determined using the standard-resolution sequence (3x3x3 mm) and the high-resolution sequence (1.5x1.5x2.5 mm).

To compare the magnitude of the condition-specific responses across pilot studies 1 and 3, voxels of maximum overlap were chosen from the incidence maps for each pilot study (Figure 3.9). For pilot study 1, the high-frequency voxel of ‘maximum incidence’ was only present in one participant but it was chosen because it was located in the centre of region 2, had the coordinate x=-42, y=-19, z=1 mm. For pilot study 3, the high-frequency-sensitive voxel of maximum incidence was present in three out of six participants and was located in the centre of region 2. The coordinate was x= -32, y=-27, z=12 mm. For pilot study 1, the low-frequency voxel of maximum incidence was present in two out of three participants and it was located in region 1, with the coordinate x=-46, y=-20, z=7 mm. For pilot study 3, the corresponding voxel showed overlap for three out of seven participants and the coordinate was x=-47, y=-20, z=5 mm. A spherical region of interest (6 mm radius) was centered on each of these coordinates and the magnitude of response for the conditions of interest was extracted, for participants #1 to #3 of pilot study 1 and #8 to #13 of pilot study 3. The beta value is taken to represent the magnitude of the response. To investigate the difference in frequency-sensitive responses for the two pilot studies, the beta values of the following conditions were subtracted. For region 2 'just listen low-frequency majority' from 'just listen high-frequency majority'(A>B); the reverse subtraction was performed for region 1 (B>A). The subtracted beta measure for these two contrasts were collapsed across the two regions to give a measure of the magnitude of the 'best-frequency' response (Figure 3.18A). To investigate the effect of attention-specific enhancement when attending to the 'best' frequency of a region the beta values of following conditions were subtracted: for region 2, 'attend low-frequency targets in the high–frequency majority stimulus' was subtracted from 'attend high-frequency targets in the high-frequency majority stimulus' (C>E). For region 1 'attend high-frequency targets in the low-frequency majority stimulus' was subtracted from 'attend low-frequency targets in the low-frequency majority stimulus' (F>D). The subtracted beta measure for these two 'attend' contrasts were collapsed across the two regions to give a measure of the magnitude of the frequency-specific attentional modulation (Figure 3.18B). Both the mean estimated response to frequency and the mean estimated response to selective attention were greater for pilot study 3 than for pilot study 1, although the effect of attention in pilot study 1 was highly variable across participants.

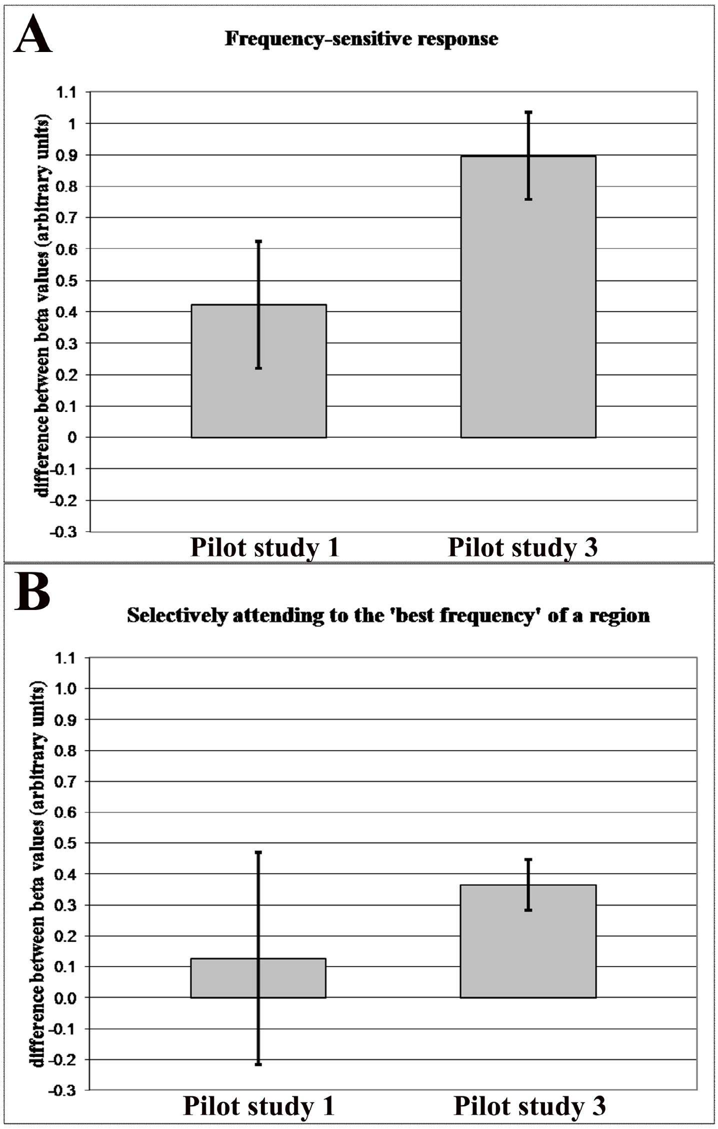


Figure 3.18 Graphs showing magnitude (arbitrary values) of BOLD response for (A) pilot 1 (for participants #1 to #3) and (B) pilot 3 (for participants #8 to #13). The error bars depict the standard error of the mean.

3.4.3 Summary

Pilot study 3 had three aims. The first aim was to determine the optimal TE in the grey matter of HG when scanning using the high resolution sequence. Calculations showed that 55 ms was the optimal value.

The second and most important aim of pilot study 3 was to investigate the statistical reliability of sound-evoked activation using image data acquired at a high voxel resolution. The results showed that the high-resolution sequence was more appropriate to use than the standard-resolution sequence because it resulted in more statistical reliability, i.e.greater t values for suprathreshold voxels and a greater extent of activation, as well as a larger magnitude of response to frequency and selective attention.

The final aim was to test whether a more extensive and improved training scheme would result in better target detection performance. Performance was considerably better than for pilot study 1. However it was still much lower than performance of the expert listeners in pilot study 2. It is possible that even this amount of training was still not enough and that participants should receive further training before scanning to make sure that they can perform the task satisfactorily.

In final conclusion it is informative to comment on the nature of the frequency-sensitive and attentional responses obtained using high voxel resolution data. Feature-specific responses were found in similar regions to pilot study 1, although much more extensive. In terms of attentional modulation, there was a greater response for the ‘just listen’ than for the ‘attend’ conditions. The most likely interpretation for this pattern is that it reflects an adaptation effect over the 30-45 mins duration of the experiment. This potential confound could be solved by mixing ‘just listen’ and ‘attend’ conditions, instead of presenting them in two separate runs.

3.5 General discussion

Three pilot studies explored optimal parameters for stimuli, task, TE and voxel resolution. It was considered important to design effective mixed-frequency sound stimuli that would serve both to map the sensory response for the high- and low-frequencies, and to direct listeners’ attention to one frequency or the other. The stimuli were proven successful in terms of eliciting patterns of frequency sensitivity, especially in pilot study 3, as the high-frequency-sensitive regions 2 and 4 and the low-sensitive region 1 were clearly located around HG. Neural effects of selective attention were found, although they were difficult to interpret because there was often a greater response for the ‘just listen’ than for the ‘attend’ conditions, in both pilot studies 1 and 3. This might be an effect of the order of presentation and will be discussed below.

It was also important to confirm whether the frequency-sensitive regions found by the current paradigm agreed with previous studies on tonotopy that have used single-frequency sound stimuli (Schönwiesner et al., 2002; Talavage et al., 2000; Talavage et al., 2004). The results showed good correspondence with those studies.

In addition to an effective stimulus, it was considered important to design an effective task which would be difficult, but at the same time enable reasonably good performance. The task provides evidence that participants are attending to one frequency stream or the other. The results of pilot study 1 revealed that the task was difficult. However, the results of pilot study 2 indicated that this low performance was more likely to be due to the fact that training was not effective than to the fact that the task was impossible to perform. In pilot study 3 training was more extensive and as a result performance improved. Further training might still be necessary in the subsequent experiment, although careful consideration should be paid to the length of time in the scanner. In pilot study 3, participants themselves indicated they were kept in the scanner too long and this contributed to fatigue. For this reason, it was considered important to train participants outside of the scanner, in a separate behavioral session. This would ensure that they receive appropriate training and that the scanning does not last too long.

3.6 Methodological issues arising

Taking into account the limitations highlighted by the three pilot studies, several issues were considered. As mentioned above, the key problems were i) poor performance in the behavioral task, ii) long training session in the scanner, iii) order effects in the activation patterns across the passive listening and attend conditions due to absence of counterbalancing.

Poor performance can be addressed by providing even more extensive training. Therefore, in Experiment 1, participants will be trained more extensively, and I will exclude those participants that cannot do the task. Performance might also be improved by having fewer switches of the task instructions and this could easily be done by changing the experimental condition after every two sound sequences, rather than after every single sound sequence. These modifications address the first problem listed above. It may be the case that being in the scanner is not an ideal environment for participants to focus and understand the task. Therefore in Experiment 1, I will train participants in a session separate before they go in the scanner. This will reduce the time that participants need to be in the scanner, which, in pilot study 3, extended up to two hours. This modification addresses the second problem listed above.

Counterbalancing the ‘just listen’ and ‘attend’ conditions across the two functional runs would address the problem of adaptation. One danger of mixing up the ‘just listen’ and ‘attend’ conditions is that participants will be aware of the task and they may listen out for targets in the ‘just listen’ condition as well, despite the instructions not to do so. This strategy could be resolved by removing the irregularities from the ‘just listen’ sound sequences, and informing the participants about it. This would ensure, as far as possible, that participants will not be looking for irregularities during passive listening. The stimuli and presentation order were revised in Experiment 1 as discussed above.

Chapter 4: Experiment 1: Selective attention to low- and high-frequency sounds studied with fMRI

4.1 Introduction

In the current study, the feature-specific attentional modulation hypothesis is revisited. Experiment 1 employs the optimal parameters in terms of stimuli, task, voxel resolution and TE informed by the three pilot studies discussed in Chapter 3. The study has three aims which are described in the following three sections.

4.1.1 Localization of high- and low-frequency-sensitive regions across primary auditory cortex

The first aim of the study was to localize specific high- and low-frequency-sensitive regions around the primary auditory cortex and compared them with the organization identified by previous studies (Schönwiesner et al., 2002; Talavage et al., 2000; Talavage et al., 2004). The high-frequency regions corresponded to regions 2, 3 and 4 (Talavage et al., 2000) and low-frequency regions corresponded to regions 1a and 1b (Schönwiesner et al., 2002), as shown in Figure 4.1B. To localize the high-frequency-sensitive regions, a condition in which participants were passively listening to the low-frequency majority stimulus was subtracted from a condition in which they passively listened to the high-frequency majority stimulus (A>B, Table 3.2). The reverse contrast was computed to find the low-frequency-sensitive regions (B>A).

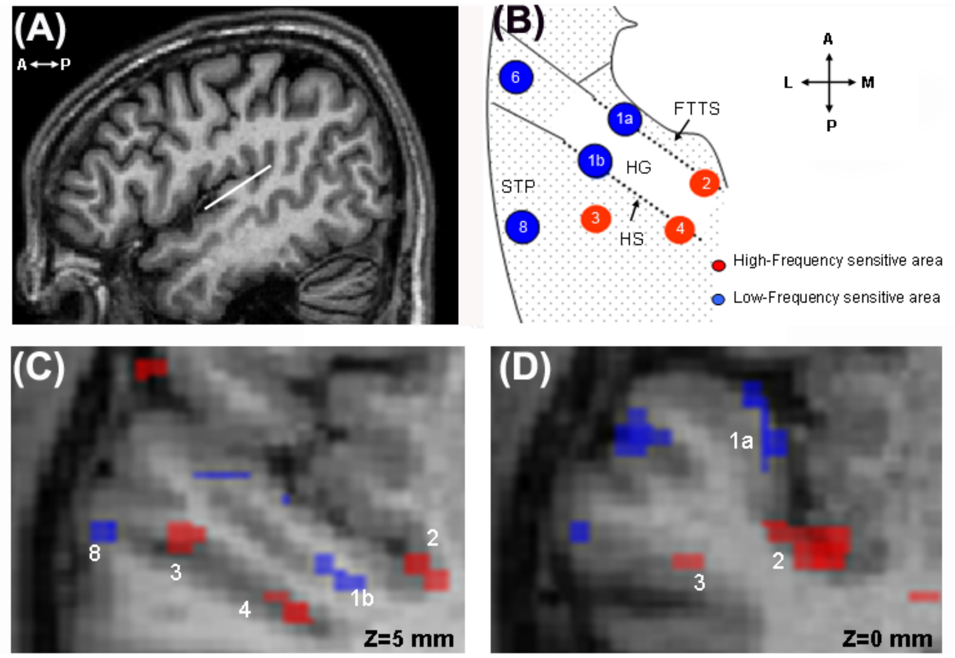


Figure 4.1 (A) Sagittal view of the brain with the oblique white line denoting the approximate location and orientation of the schematic view shown in panel (B) along the supratemporal plane. (B) Schematic representation of the most consistently found high (red) and low (blue) frequency-sensitive regions across the human auditory cortex reported by previous studies (Schönwiesner et al., 2002; Talavage et al., 2000; Talavage et al., 2004). The primary area is shown in white and the nonprimary areas are shown by dotted shading. Panels (C) and (D) illustrate the high- (red) and low- (blue) frequency sensitive areas across the left auditory cortex of participant #3. Two planes in the superior-inferior dimension are shown (z=5 mm and z=0 mm above the CA-CP line). Abbreviations: A: anterior, P: posterior, M: medial, L: lateral, FTTS: first transverse temporal sulcus, STP: supratemporal plane (Hall & Paltoglou, 2009).

A secondary goal for frequency-dependent localization was to investigate whether there was a similar pattern of frequency-dependent responses when contrasting high-frequency majority and low-frequency majority conditions, irrespective of the attentional instructions. This would determine whether the pattern of tonotopy was context insensitive. Thus, two more contrasts were performed for each of the ‘attend’ conditions. First, the condition in which participants attended to the high-frequency targets in the low-frequency-majority stimulus was subtracted from the condition where they attended to the high-frequency targets in the high-frequency-majority stimulus (C>D, Table 3.1). Second, the condition where participants attended to the low-frequency targets in the low-frequency-majority stimulus was subtracted from the condition in which they attended to the low-frequency targets in the high-frequency-majority stimulus (E>F). Subsequently, conjunction analysis was performed across the passive and the two attend contrasts. Conjunctions identify those voxels that are activated in several different pairwise comparisons (Price & Friston, 1997). Hence, it makes it possible to identify those voxels showing a response to the feature of interest, irrespective of the listening instructions. To investigate the same question for the low-frequency sensitive regions, conjunction analysis was performed on the equivalent three contrasts (B>A, D>C, F>E). The conjunction results are reported across participants using an incidence map approach.

4.1.2 Evidence for feature-specific attentional modulation

The second and most important aim was to investigate whether there was attentional modulation in the frequency-sensitive regions identified, i.e. a feature-specific effect of attention. An effect of ‘general enhancement’ would occur if there was greater response for attending to targets in the majority stream, than passively listening to the same stimulus (C>A, F>B). Note that the effect of general enhancement is poorly controlled as it is confounded by other factors which are not of interest here such as performing the detection task. In contrast, an effect of attention-specific enhancement would occur if there was greater response for the condition in which participants attend to targets in the majority stream, compared to attending the targets in the minority stream of the same stimulus (C>E, F>D, Table 3.2). Note that this contrast carefully controls for the task-related factors, as the focus of attention is the only difference between the two conditions contrasted.

An effect of suppression would occur if there was a reduced response for the condition in which participants attend to the targets in the minority stream than for when they passively listen to the same stimulus (A>E, B>D, Table 3.2). This effect was expected to be found in the regions that process the majority stimulus. For instance, in a high-frequency-sensitive region, one would expect less response for ‘attend low-frequency targets in high-majority stimulus’ than ‘passively listen to high-majority stimulus’.

4.1.3 Evidence for widespread attentional modulation

The third aim was to investigate attentional modulation across the auditory cortex, and not just in the frequency-sensitive regions. In the strict sense, frequency-specific attentional enhancement and suppression should occur only in the regions sensitive to the attended frequency. Widespread enhancement, even if it includes the feature-sensitive regions, does not support the hypothesis of feature-specific attentional modulation. Given that previous studies of auditory selective attention showed widespread enhancement (Degerman et al., 2006; Petkov et al., 2004) it was considered important to explore the data for this possibility. A related issue is whether attending to either high- or low- frequency targets in the majority stream leads to an increased response in the same regions of the auditory cortex, compared to passively listening to either high- or low-frequency majority stimulus respectively. To this end, conjunction analysis was conducted between the two effects of general enhancement contrasts (C>A and F>B, Table 3.2).

To examine the above research questions, several analyses were conducted on the magnitude, distribution and extent of the frequency-sensitive response and the attentional modulation of individual participants, and the results were pooled across the group.

First, a region-of-interest (ROI) analysis investigated the magnitude of the response to the conditions of interest within specific frequency-sensitive regions. One possible reason why other fMRI studies have not found feature-specific modulation is that group averaged analysis does not account for variability in location of these regions across participants (Degerman et al., 2006). The solution to this problem is to analyze participants individually and measure attentional modulation within individually specific ROIs.

Averaging the signal across the ROI can conceal variation in the patterns of individual voxels. This point can be demonstrated using an example from data acquired during pilot study 3. In the upper graph of Figure 4.2, the mean response of region 1 (averaged across all voxels) in participant #10 is plotted across conditions. Note that there is a greater response when passively listening to the low-frequency majority stimulus, than when attending to high- or low-frequency targets in the same stimulus. The responses of two individual voxels (a peak and a non-peak voxel) within this region are shown in the two lower graphs. The peak voxel has generally a similar pattern of response across conditions to that shown by the mean activation. However, the non-peak voxel shows a different pattern of results. There is a slightly greater response when attending to the low-frequency targets in the low-frequency majority stimulus, than when passively listening to them, although the difference does not reach significance. An analysis that takes into account the variability across individual voxels would be much more sensitive than an analysis performed on the mean data alone.

To avoid these pitfalls, first individual analysis was employed to localize the frequency-sensitive regions within the auditory cortex of each participant. Then, the time-courses of each voxel in selected ROIs were extracted and analyzed to investigate the effects of attention on the magnitude of response.

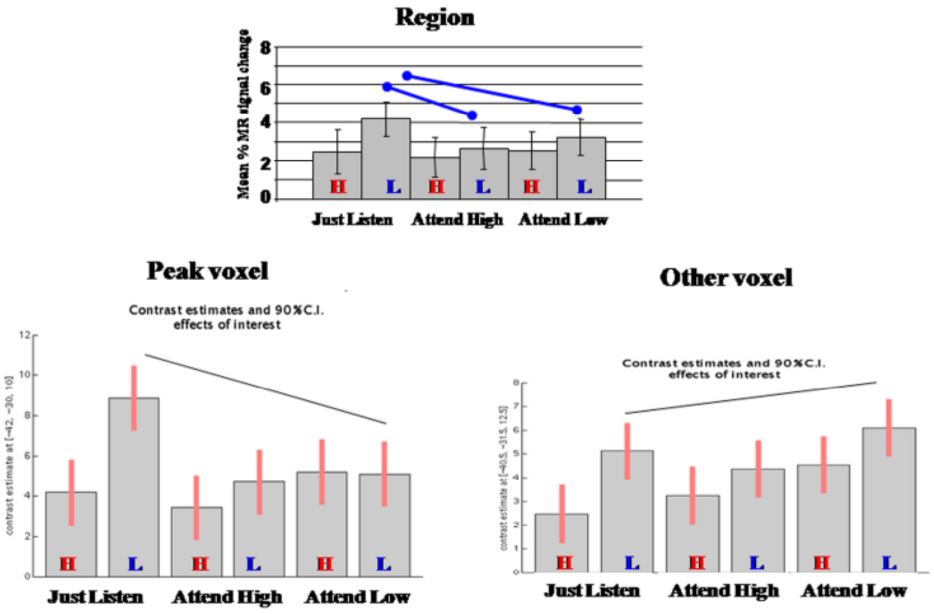


Figure 4.2 Graph for one of the low-frequency sensitive regions for participant #10 (pilot study 3), along with contrast estimates for two voxels; the peak voxel (left) and another voxel (right).

To investigate the distribution of auditory cortical responses across participants, incidence maps were plotted for both the frequency-specific response and the attentional modulation. The statistical threshold was $p<0.01$ as used previously (see Section 3.2.1). To examine the feature-specificity of such attentional modulation, the same attentional contrasts were masked by the frequency-sensitive regions, so that only attentional modulation within the appropriate frequency-sensitive regions could be viewed. This approach has the advantage of quantifying the amount of variability across participants because it is explicit how many participants contribute to each voxel.

To investigate the extent of the frequency-sensitive response and the general and specific attentional enhancement, the number of voxels across the auditory cortex of each participant was counted for the three different contrasts. To determine the extent of their overlap, the number of voxels common to each was calculated.

4.2 Methods

4.2.1 Participants

Six participants (#1 to #6, three females, mean age=24 years, age range 19-29 years) took part in Experiment 1. Note that none of the participants participated in any of the pilot studies. They were right handed and had normal hearing (≤25 dB for 250 to 8000 Hz). None of them had a history of audiological or neurological impairment. All participants were recruited via a poster at the University of Nottingham campus and were paid for their participation. Informed consent was obtained before MR scanning. Four more participants were tested behaviorally, but they were not scanned because they made a large number of errors despite extensive training.

4.2.2 Stimuli and task

Stimuli and task were the same as those in pilot study 3. Two minor changes were made to the stimulus paradigm compared with pilot study 3 in Chapter 3. The duration of each sound sequence was 16 s instead of 15.5 s as it was considered more optimal for the MR signal to have sound stimuli continuously presented throughout the TR period, instead of having 500 ms of silence at the end of the period. There were also 16 repeats of each condition, instead of 15. Finally, the timing of the irregularity in the slow stream was changed, from 60 ms to 65 ms. However, this was a minimal difference and did not affect performance, as the behavioral results show.

All the ‘attend’ conditions contained one or two timing irregularities (targets) in each stream, and participants had to attend to one frequency stream at any time. However, unlike the pilot studies, there were no irregularities in the ‘just listen’ conditions and all participants were informed that this was the case. This was done to discourage participants from searching for irregularities during passive listening, and thus covertly performing a target detection task.

4.2.3 Procedure

All participants completed two training sessions, in a sound-attenuating booth at the MRC Institute of Hearing Research. In the first session participants were given two or three practice runs on each ‘attend’ condition, each containing four sound sequences. When it was clear that they understood the task, they went through two runs which contained eight sound sequences for each of the four ‘attend’ conditions. Each run lasted about 9 mins. The conditions were randomized by a latin square design. The ‘attend’ conditions changed after every two sound sequences. The second session was the same as the first. The criteria for accepting participants to take part in Experiment 1, was to reach a d prime of 2 in each of the four ‘attend’ conditions.

Experiment 1 contained two scanning runs, each containing an equal number of all conditions (eight repeats of each condition). Unlike the pilot studies, the order of the passive and attend conditions was fully randomized by a latin square design and counterbalanced across participants. Visual instructions (‘high pitch’, ‘low pitch’) were presented throughout each ‘attend’ condition. During the four ‘attend’ conditions, participants were requested to press the button whenever they heard an irregularity in the stream to which they were instructed to attend. A fixation cross was presented throughout the ‘just listen’ and the silent baseline conditions, and participants were asked to just look at a fixation cross. Target discrimination performance was calculated using hit minus FA rates, as well as d prime. The procedure for scoring detection performance has been reported in Section 3.2.1. For this analysis the low-frequency targets that occurred at the same time as the scanner noise were excluded from the analysis, because it was observed that they were fully masked by the scanner noise and were impossible to detect.

It was considered important to have a difficult task that would require participants to attend to only one stream and ignore the other. To demonstrate the benefit of selective attention on performance, participants #4, #5 and #6 were tested in an additional divided attention version of the task, as well as the original selective attention version. This testing was done in a separate behavioral session; some days after Experiment 1 took place. In the divided attention conditions, participants were asked to respond to timing irregularities in both high- and low-frequency streams.

4.2.4 Imaging protocol

All participants were scanned in a single session which lasted up to one hour. All scanning steps were identical to pilot study 3, described in Section 3.2.1. In each of the two scanning runs, 114 T2*-weighted functional scans were acquired (TE=55 ms and flip angle=90°). The voxel resolution was 1.5x1.5x2.5 mm (matrix size 64x64) with saturation bands (Figure 3.6B, yellow rectangle) to reduce artifacts. Each functional image consisted of 18 slices, chosen to include the superior temporal gyrus (Figure 3.6B, red rectangle). Each run lasted approximately 17 mins. Sparse sampling was used to reduce acoustic masking and to reduce auditory cortical activation due to the scanner noise (TR=9 sec, see Figure 4.3). The delay between the start of the acquisition and the start of a sound sequence was increased from 1.5 s to 2 s to reduce the overlap with the stimulus and the scanner noise.

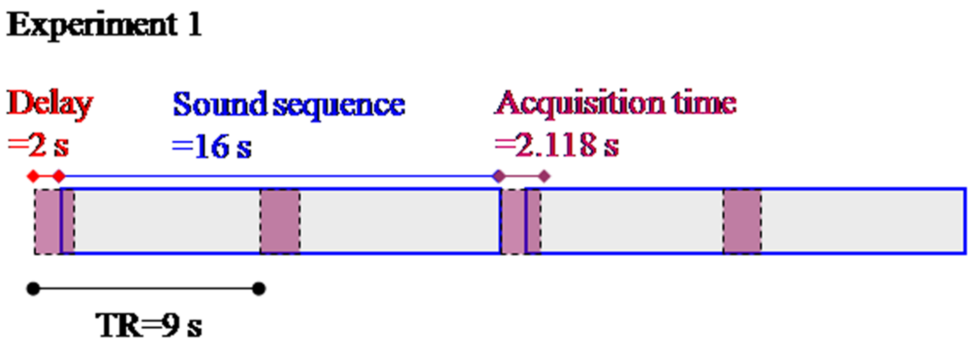


Figure 4.3 Sparse sampling protocol (Hall et al., 1999) used in Experiment 1. The grey rectangular bar represents the duration of the rhythm sequence (16 s). The red rectangular bars represent the acquisition of a single volume of data (2.118 s for all 18 slices of the functional scan). TR: repetition time.

4.2.5 Image analysis

Standard procedures for preprocessing were performed, including realignment, coregistration and normalization. The procedures are reported in detail in Section 3.2.1. Again, the normalized scans maintained their original resolution of 1.5x1.5x2.5 mm and were smoothed using a Gaussian kernel of 3 mm FWHM. Individual time series were modeled in the framework of the GLM. The design matrix contained six regressors, one for each experimental condition. Two regressors, accounting for the differences in the mean signal between the two runs, were also included in the design matrix. The high-pass filter, applied to avoid low-frequency artifacts due to physiological fluctuation, was 864 s for each run.

Magnitude of the response in the frequency-sensitive regions, measured using ROI analysis

To test the feature-specific attentional modulation hypothesis, ROI analysis was performed on the selected frequency-sensitive regions around the primary auditory cortex for participants #1, #2 and #3. The analysis was done for only three participants because the size of the data set was sufficiently powerful to support an exploratory investigation. The first step was therefore to locate the high- (regions 2, 3 and 4) and low-frequency sensitive regions (regions 1a and 1b) for each participant around HG. Suprathreshold frequency-sensitive responses were superimposed onto a probability atlas which contained three anatomical subdivisions of HG (Te 1.0, Te1.1, Te1.2 Morosan et al., 2001), using the SPM anatomy toolbox (Eickhoff et al., 2005). This atlas was used to identify the clusters of frequency-sensitive activation that overlapped with HG. Localization of the peaks of activation also considered the tonotopic scheme reported by Talavage et al. (2000). This procedure was described for the pilot studies in Section 3.2.1.

To make sure that all ROIs were present and were distinct from one another (especially regions 1a and 1b), different statistical thresholds were adopted for different contrasts and different participants. Note that the same statistical threshold was used for all of the high-frequency-sensitive regions or all of the low-frequency-sensitive regions for any individual participant (Table 1.4). The method by which the chosen threshold was determined was to start with a statistical threshold of $p<0.01$ and it increased by steps of 0.01 .

Within these ROIs, the MR signal time course was extracted for every voxel using the marsbar toolbox (http://marsbar.sourceforge.net/). The experimental conditions were re-coded according to whether the frequency that was attended corresponded to the best frequency (BF) of a particular frequency-sensitive region. Using terminology from auditory neurophysiology (e.g. Merzenich & Brugge, 1973), the best frequency of a region is the sound frequency to which that region responds to most. Off-frequency (OFF BF) sounds were the ones to which a region does not have the greatest response. Conditions were collapsed across best frequency because there was no hypothesis that the high- and the low- frequency sensitive areas would show a different pattern of attentional modulation. To investigate attentional modulation within these regions, the MR signal for all voxels within these regions was extracted and a univariate analysis of variance was performed in SPSS. To meet the normality requirements for parametric analysis, the data were log-transformed prior to the analysis.

Distribution of auditory responses

To investigate the distribution of feature-sensitive and attentional responses across the six participants, incidence maps were constructed by summing individual thresholded statistical maps ($p<0.01$) for the different contrasts. Firstly, the distribution of feature-sensitivity was explored using two incidence maps corresponding to the high- and low-frequency responses (contrasts A>B and B>A respectively, see Table 3.2). Two additional incidence maps explored how reliable the frequency-dependent responses were across the listening tasks by computing the incidence of the conjunction ($p<0.01$) of the three frequency-sensitive pairwise comparisons (high frequency; A>B, C>D and E>F and low- frequency; B> A, D>C and F>E, Table 3.2).

The distribution of two different types of enhancement (termed general and attention-specific enhancement respectively) were explored using incidence maps. The effect of general enhancement refers to the increase in activity when attending compared to passive listening for the same sound stimulus and does not control for task differences. For the high-frequency sound this was computed using the contrast ‘attend high-frequency targets in high-frequency majority stimulus> just listen high-frequency-majority stimulus’ (C>A). For the low-frequency sound this was computed using the contrast ‘attend low-frequency targets in the low-frequency-majority stimulus> just listen low-frequency-majority stimulus’ (F>B). To investigate the frequency specificity of this pattern, the effects of general enhancement were also masked by the corresponding frequency-sensitive regions. Note that the effect of general enhancement for each participant was masked by their own corresponding thresholded map of frequency sensitivity before being summed to form the group-level incidence maps. In other words, each mask was different for each participant. To investigate the stimulus-independent pattern of the general attentional enhancement across auditory cortex, a conjunction analysis ($p<0.01$) was performed between the two original nonmasked contrasts (C>A and F>B).

The effect of attention-specific enhancement refers to the increase in activity when attending to targets in the frequency-majority stream compared to attending to targets in the minority stream for the same sound stimulus. This contrast controls for the task-related factors. For the high-frequency-majority stimulus this was computed using the contrast, 'attend high-frequency targets in the high-frequency majority stimulus> attend low-frequency targets in the high-frequency majority stimulus' (C>E). For the low-frequency-majority stimulus this was computed using the contrast, 'attend low-frequency targets in the low-frequency majority stimulus> attend high-frequency targets in the low-frequency majority stimulus', (F>D).

Finally, two maps corresponded to the general suppression effect for high- and low-frequency sounds, defined by the contrasts 'just listen high-frequency-majority stimulus>attend low-frequency targets in the high-frequency majority stimulus' (A>E), and 'just listen low-frequency-majority stimulus>attend high-frequency targets in the low-frequency majority stimulus' respectively (B>D).

Extent of frequency-sensitive and attentional modulation responses and the extent of their overlap

To perform this analysis, the number of activated voxels for the contrasts of interest (showing frequency sensitivity, general attentional enhancement, specific attentional enhancement were counted from the thresholded statistical maps ($p<0.01$) and their overlap was calculated from both auditory cortices of each participant using a Matlab script. The statistical threshold for all contrasts was $p<0.01$. The auditory cortex, containing PP, PT and HG, was defined by an anatomical 'mask' (Figure 4.4).

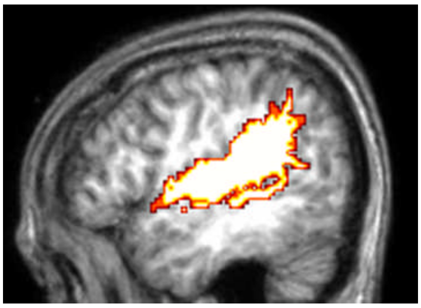


Figure 4.4 Anatomical mask of the auditory cortex, overlaid on the mean anatomical scan. This mask was used to isolate the region from which the number of activated voxels were calculated.

4.3 Results

4.3.1 Target detection results

The mean and individual performance for the six participants across the two training sessions are shown in Figure 4.5. The mean hit minus FA rate fell between 70 and 85% (Figure 4.5A). The mean d primes fell between 3 and 4.5 (Figure 4.5B). This confirms that participants have understood the task and can perform satisfactorily. Figure 4.6 shows the mean performance during Experiment 1. The mean hit minus FA rate fell between 28 and 65% (Figure 4.6A). The individual d primes fell between 2 and 3.1 (Figure 4.6B).

When comparing Figure 4.5 and Figure 4.6, performance appears to have declined in the scanner. This could be due to the presence of the scanner noise, despite the fact that low-frequency targets that overlapped with scanner noise were excluded from the analysis. In fact, for the condition ‘attend high-frequency targets in the low-frequency majority stimulus’ participants showed a rather low score (hit-FA rate 28%, d prime of 2) although the mean was skewed by participant #4. However, with the exception of this condition, all other conditions showed improvement compared with the results in pilot study

3 in which the means ranged between hit-FA rate of -20 to 40% and d prime of 1.5 and 2.5. A paired sample t-test between the results of the training and the fMRI session, showed this difference was significant (t(23)=5.030, p<0.001).

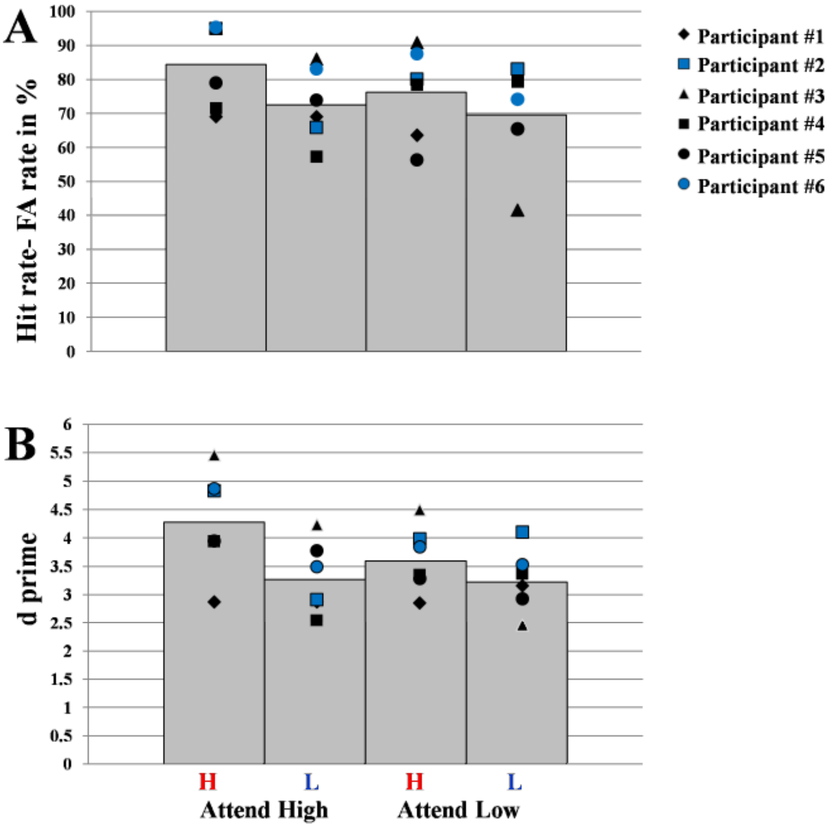


Figure 4.5 (A) Hit rate minus FA rate in % and (B) d prime of target discrimination performance for the four ‘attend’ conditions for all six participants in the sound-attenuated booth. Bars indicate average across all six participants, symbols indicate individual performance. H: high-frequency-majority stimulus. L: Low-frequency-majority stimulus.

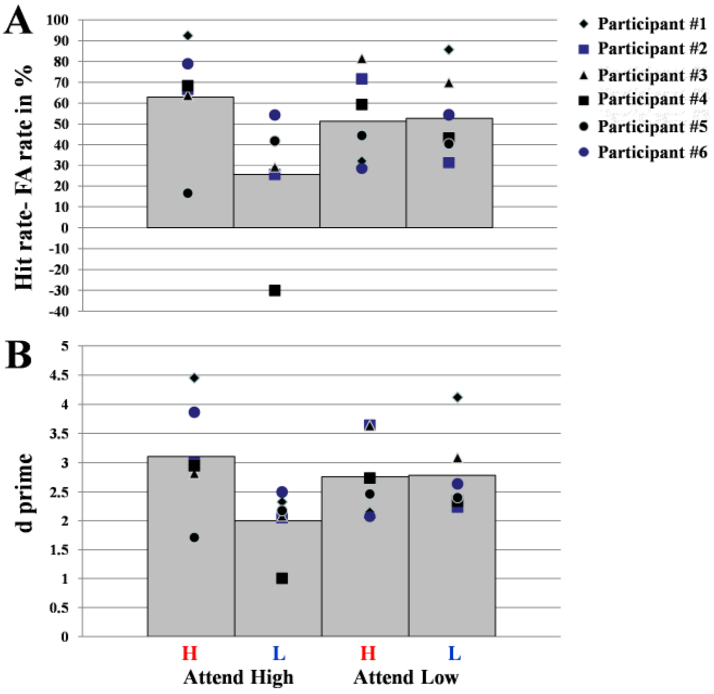


Figure 4.6 (A) Hit rate minus FA rate in % and (B) d prime of target detection performance for the four 'attend' conditions for all six participants during Experiment 1. Bars indicate average across all six participants, symbols indicate individual performance.

For those participants completing both divided and selective attention tasks, a performance cost was observed when dividing attention across both streams (Figure 4.7).

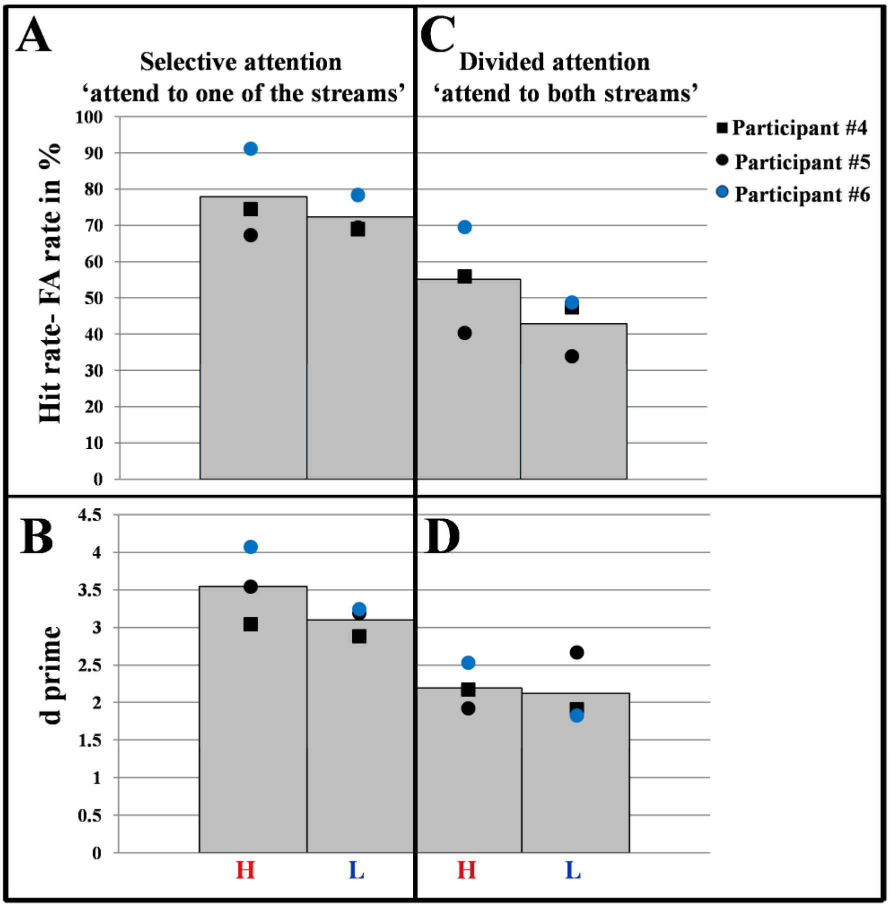


Figure 4.7 Target-detection performance for the selective attention task in (A) Hit rate minus FA rate in % and (B) d prime and for the divided attention task in (C) Hit rate minus FA rate in % and (D) d prime for participants #4, #5 and #6 during testing in a separate session after the fMRI experiment. The H column of the selective attention graph, is an average of the two high-frequency majority conditions, while the L Bar is the average of both low-frequency majority conditions ('attend high' and 'attend low').

4.3.2 Localization of high- and low-frequency-sensitive regions

All frequency-sensitive regions (regions 1a, 1b, 2, 3 and 4) were identified for participants #1, #2 and #3 (Table 4.1). However, for participant #3 there was one unified region 1 instead of two distinct regions 1a and 1b. Note that in some cases more than one cluster of activation was assigned to a particular frequency-sensitive region. For example, participant #1 had two regions 2 in the left hemisphere. Note that this analysis was conducted only in participants #1 to #3, as the preliminary results for the incidence maps showed widespread attentional effects spreading across the auditory cortex.

sb	region	t	Peak Coordinates			voxels	Subdivisions of HG	Probability	p	Side
			x	y	z					
#1 H>L	2	6.4	39	-21	5	70	Te.1.1	30	0.01	R
	“	6.1	-33	-26	10	44	“	20	“	L
	“	4.6	-40	-21	0	19	“	10	“	L
	4	4.8	50	-24	5	20	“	20	“	R
	“	4.5	-46	-30	2	15	“	20	“	L
	3	5.8	56	-16	5	14	Te 1.0	50	“	R
	“	3.4	-58	-20	5	15	OP4	10	“	L
#1 L>H	1a	4.7	46	-12	5	3	Te 1.0	60	0.001	R
	“	4.4	46	-16	10	1	“	40	“	R
	“	4.2	44	-20	12	3	OP1	70	“	R
	“	6.1	-48	-8	0	21	OP4	30	“	L
	1b	4.4	44	-27	12	2	Te 1.1	50	“	R
	“	8.6	-42	-24	8	65	Te 1.0	80	“	L
#2 H>L	2	2.8	-36	-21	2	7	Insula	N/A	0.01	L
	“	4.2	39	-30	18	33	OP1	50	“	R
	4	2.6	46	-27	10	4	Te 1.1	50	“	R
	3	2.4	56	-22	10	1	OP1	30	“	R
	“	2.4	60	-22	10	1	OP1	10	“	R
	“	3	-46	-33	12	17	Te 1.1	20	“	L
#2 L>H	1a	7.1	40	-22	10	40	Te 1.1	40	0.002	R
	“	4.5	51	-20	2	5	Te 1.0	60	“	R
	“	4.1	54	-14	-2	8	Te 1.0	20	“	R
	“	4	-44	-18	-2	8	Te 1.1	10	“	L
	1b	3.8	50	-22	10	5	Te 1.1	60	“	R
	“	5.4	-38	-27	10	41	Te 1.1	90	“	L
	“	4.7	-52	-12	2	27	Te 1.0	50	“	L
#3 H>L	2	6.1	33	-27	15	105	OP2	30	0.01	R
	“	4.1	-34	-26	5	26	HG	N/A	“	L
	4	5.4	69	-9	18	227	OP4	50	“	R
	“	3.9	-50	-16	8	18	Te 1.0	40	“	L
	3	4.1	57	-28	10	16	STG	N/A	“	R
	“	3	-44	-40	20	2	STG	N/A	“	L
	“	3.4	-56	-30	22	12	OP1	40	“	L
#3 L>H	1	7.5	50	-20	5	221	Te 1.0	70	0.002	R
	“	5.2	-46	-28	10	49	Te 1.1	30	“	L

Table 4.1 showing the regions of interest which were selected for participants #1, #2 and #3 in order to perform the ROI analysis. R: right. L: left. H>L: just listen high-frequency majority>just listen low frequency majority. L>H: ‘just listen low-frequency majority>just listen high-frequency majority. Sb: subject. OP: operculum. p<0.01: T=2.34. p<0.002: Te=2.9. p<0.001: T=3.7

4.3.3 Magnitude of the response (ROI analysis)

In the ANOVA design, the dependent variable was the log-transformed time courses. There were five independent variables: listening condition (six levels), run (two levels), hemisphere (two levels), participant (3 levels) and auditory region (six levels). The six auditory regions were low-frequency regions 1a, 1b and 1 (containing both 1a and 1b), and the high-frequency regions 2, 3 and 4. All two-way interactions were included in the model. The univariate ANOVA showed an effect of condition [F(6, 265529) =62.954, p<0.001] and an interaction between listening condition and auditory region [F(30, 265529)=3.249, p<0.001]. When plotting the means of all regions, it was clear that region 4 showed a different pattern of results (see Figure 4.8). Primarily region 4 showed a large response when passive listening to the high-frequency majority stimulus, but showed little response sensitivity to the other listening conditions.

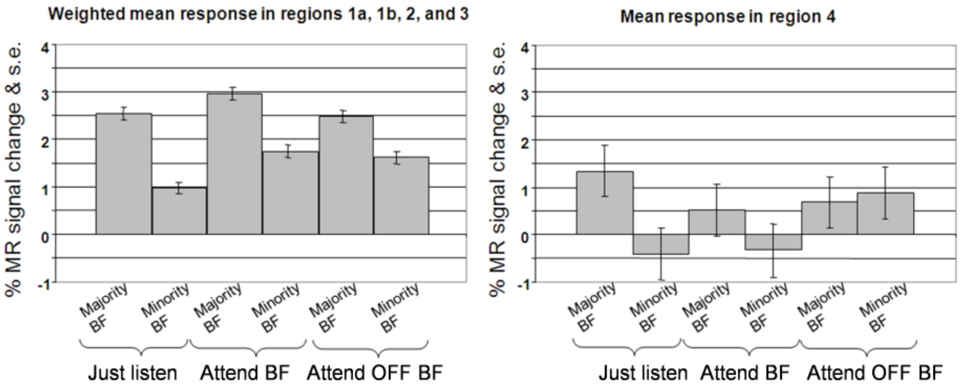


Figure 4.8 Graphs showing the mean response for regions 1a, 1b, 2 and 3 (left) and the response for regions 4 (right) for all conditions. The error bars denote the standard error of the mean.

For this reason, the same univariate ANOVA was computed again for regions 1a, 1b, 1, 2 and 3, excluding high-frequency-sensitive region 4. The results showed that there was a significant effect of listening condition [F(6, 200788)=72.842, p<0.001], while there was no interaction between condition and region. To investigate the effect of listening condition, two planned comparisons were computed. The first comparison looked at the effect of general attentional enhancement (not controlling for task differences) by comparing Attend BF-just listen BF and the second looked at attention-specific enhancement (controlling for task differences) by comparing Attend BF with Attend OFF BF. The Bonferroni correction was 0.05/2=0.025. The planned comparisons revealed a significant effect of attention-specific enhancement when participants were attending to the BF of that area, compared to when attending off BF (p < 0.003). In addition, there was a marginally significant effect of general attentional enhancement when attending to the BF of that area, compared to 'just listen' (p < 0.027) (Figure 4.9). However, this ROI analysis was not completed for all six participants, as the preliminary results for the incidence maps indicated that the effects of general and attention-specific enhancement were more widespread than expected and extended beyond the ROIs.

The examiners pointed out that individual voxel measurements cannot really be considered independent, because the signal time course in adjacent voxels are partially correlated with each other. This is a valid point. One possible way to avoid this problem might have been to average the signal across all voxels within a region. Although reducing the number of observations in the spatial domain there would still be sufficient observations for each condition in the temporal domain (32 scans for each condition).

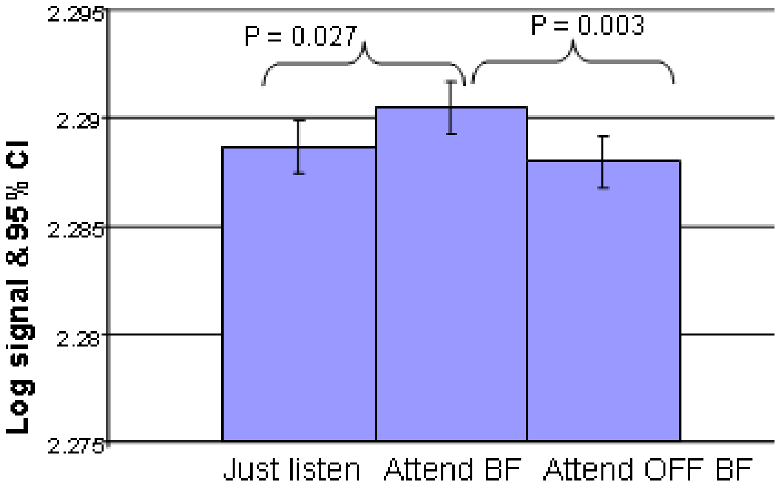


Figure 4.9 Response magnitude of the BF majority conditions.

4.3.4 Distribution of auditory cortical responses

Figure 4.10A shows the incidence maps of the frequency-sensitive responses for high- (red) and low-frequency sounds (blue) across all six participants (A>B and B>A respectively). The most consistent response was around primary auditory cortex. This confirms previous studies which have shown that narrowband noise bursts activate mainly the primary auditory cortex (Rauschecker et al., 1995). For the high-frequency-sensitive regions, three participants consistently showed a response in the antero-medial border of HG bilaterally, where region 2 has been reported. In fact, the response extended somewhat along the first transverse temporal sulcus bilaterally, although it is more consistent in the medial most part of HG. High-frequency region 3 can be also identified, although there is no overlap across participants.

The low-frequency-sensitive regions, on the other hand, were located in the more posterior part of HG, and in antero-lateral HG. Regions 1 and 6 can be identified bilaterally in medial and lateral HG respectively. Low-frequency-sensitive regions spread out more than high-frequency-sensitive regions across the auditory cortex.

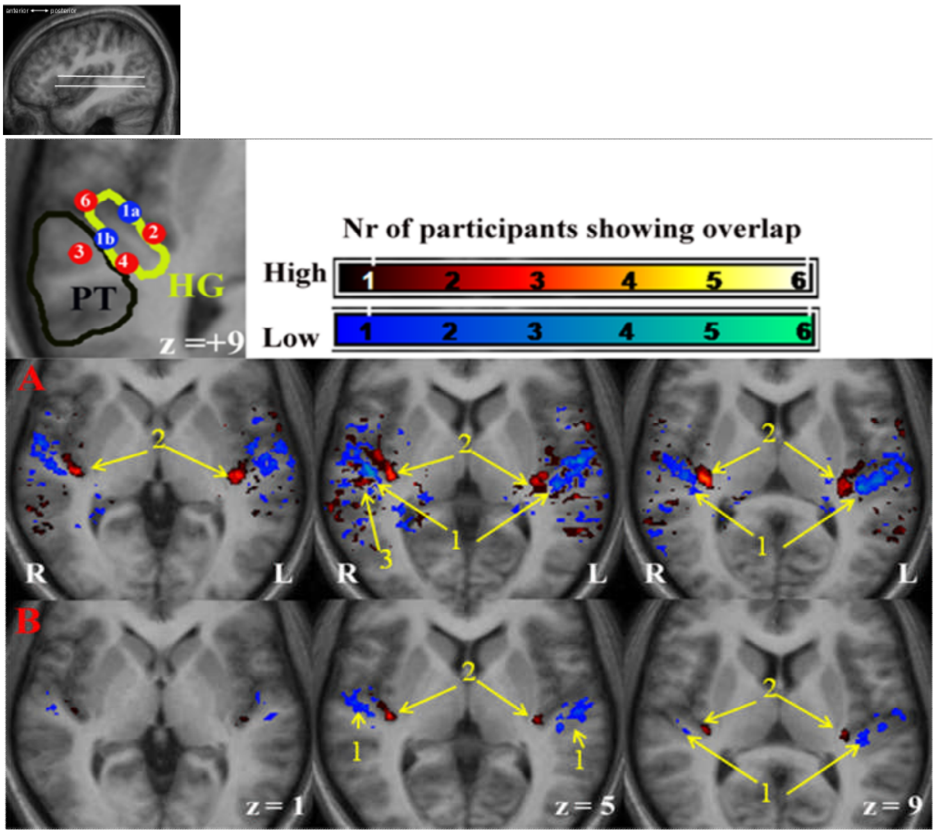


Figure 4.10 Incidence maps for (A) High- (red) and low-frequency-sensitive regions (blue), overlaid on the mean anatomical scan and presented on four axial slices across the auditory cortex, as shown by the white lines on the sagittal image on the top. (B) Conjunction analysis for high- (red) and for low-frequency sounds (blue). For the high-frequency sounds, each high-frequency majority was contrasted with the low-frequency majority that had the same listening instruction. The reverse contrasts were computed for the low-frequency sounds. The functional results were superimposed on the mean anatomical image, which was computed from the normalized anatomical scans of all 6 participants.

Figure 4.10B displays the distribution of the conjunction analysis which investigated whether the pattern of frequency sensitivity was the same across all three listening conditions for high- (A>B, C>D, E>F) and low-frequency regions (B>A, D>C, F>E). Although the responses were much more spatially restricted, the most consistent high-frequency response corresponded to region 2 bilaterally (3/6 participants). For the low-frequency sounds, a consistent response occurred in region 1, in HS (3/6 participants). This highly conservative test confirmed a reasonable consistency across conditions and across participants.

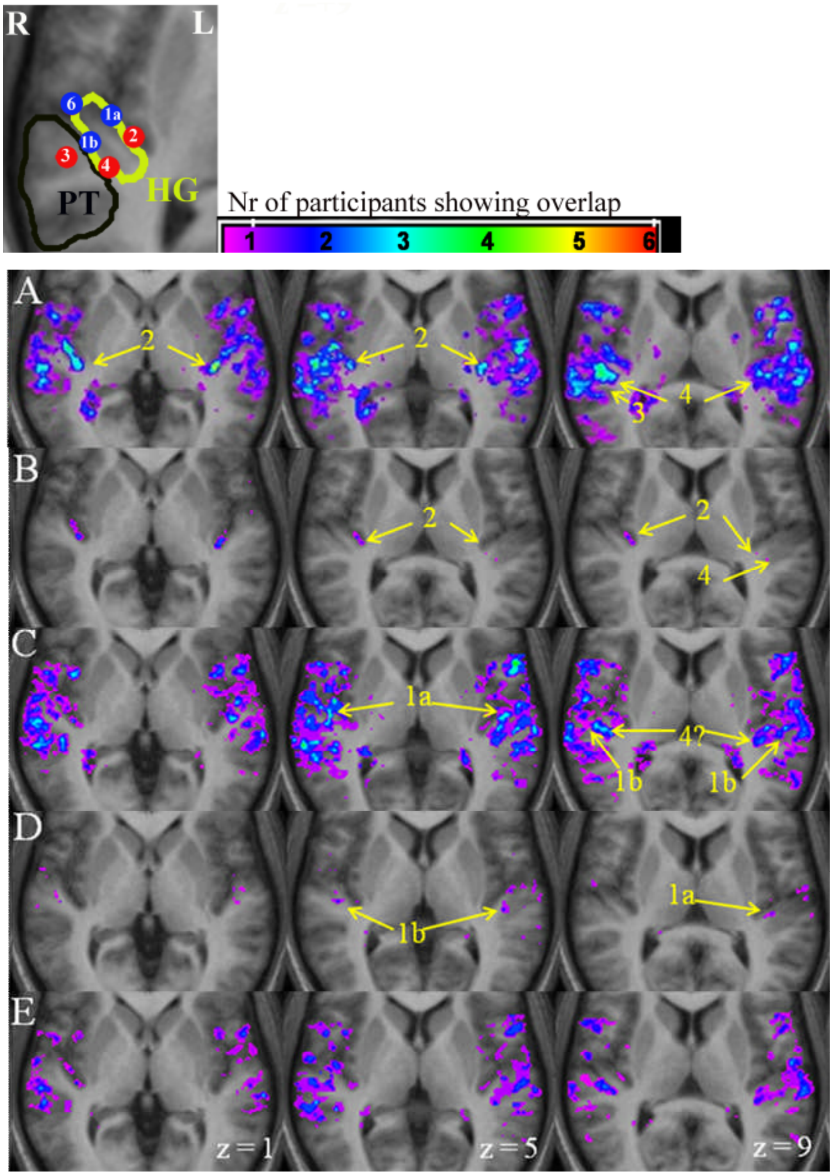


Figure 4.11 Incidence maps on the general effect of enhancement for (A) the high-frequency sounds, shown by the contrast 'attend high targets in high-frequency majority stimulus>just listen high-frequency majority'. (B) the high-frequency sounds masked by the high-frequency sensitive regions contrast ('just listen high-frequency majority>just listen low-frequency majority'). (C) the low-frequency sounds, shown by the 'attend low targets in low-frequency majority stimulus>just listen low-frequency majority' contrast. (D) the low-frequency sounds 'attend low targets in high-frequency majority stimulus>just listen high-frequency majority', Masked by the low-frequency sensitive regions contrast ('just listen low-frequency majority>just listen high-frequency majority'). (E) Conjunction between (A) and (C).

Figure 4.11A shows the distribution of the general effect of enhancement for the high-frequency sounds, i.e. an increase in response when attending to the high-frequency-majority stream compared to passive listening

(C>A). Widespread enhancement was observed across primary and non-primary auditory cortices, as well as in frontal cortex. High-frequency-sensitive region 2 was clearly identified bilaterally and was very consistent across participants (5/6). High-frequency-sensitive region 4 was also seen bilaterally and this was consistent in 4/6 participants on the right side. Interestingly, region 4 did not come up in the sensory activation incidence maps (Figure 4.10). Finally, region 3 was found in lateral PT on the right. Figure 4.11B shows the above contrast masked with the high-frequency-sensitive regions. The only overlap was for region 2 bilaterally, in medial first transverse temporal sulcus, (2/6 participants). Although regions 3 and 4 were also identified, there was no overlap (1/6 participant).

Figure 4.11C shows the distribution of the general effect of enhancement for the low-frequency sounds, i.e. an increase in response when attending to the low-frequency majority stream compared to passive listening (F>B). Again, an extended response was revealed across the auditory cortex and in the frontal cortex. Region 1a and 1b were identified, both in central HG (z=5 mm), but also in HS, extending medially, where region 4 was expected to be located. Figure 4.11D shows the above contrast masked by the low-frequency-sensitive regions. Although the response is very sparse, there was overlap for about 2 participants for this analysis where regions 1a and 1b were expected to be found.

To investigate whether the same voxels showed an effect of general enhancement for both high- and low-frequency sounds, a conjunction analysis was computed between the two general enhancement contrasts (C>A and

F>B) The conjunction results are shown in Figure 4.11E. There are some regions that show general enhancement for both high- and low-frequency majority sounds, including posterior PT and lateral HG, especially at the locus where high-frequency region 4 was expected to be found. This was some evidence that attention to either sound exerted a widespread effect that was not restricted to a particular frequency.

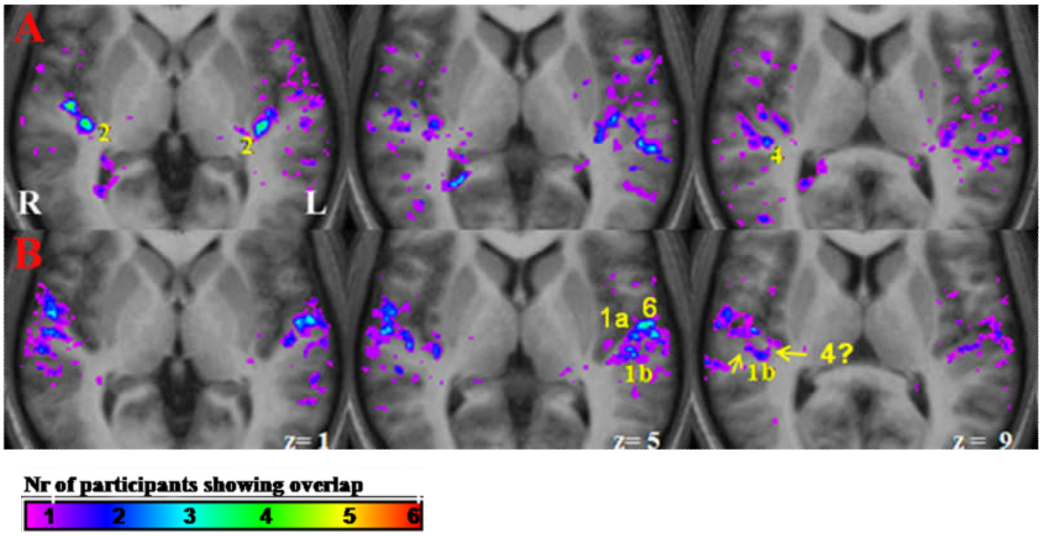


Figure 4.12 Incidence maps showing the effect of attention-specific enhancement for (A) high- and (B) low-frequency sounds. The contrasts performed were (A) 'attend high frequency targets in high-frequency majority stimulus> attend low frequency targets in high-frequency majority stimulus'(C>E) and (B) 'attend low frequency targets in low frequency majority stimulus> attend high frequency targets in low-frequency majority stimulus' (F>D).

Figure 4.12A and B show the incidence maps for the effect of attention-specific enhancement for the high- and for the low-frequency sounds respectively (C>E and F>D). For both contrasts attention-specific enhancement mainly occurred around primary auditory cortex. For the high-frequency sounds there was consistent overlap in medial first transverse temporal sulcus, where region 2 was expected to be found (3/6 participants), as well as in medial HS, where region 4 was expected to be found (2/6 participants). For the low-frequency sounds, there was a consistent response in lateral HG (region 6, Te 1.2, 3/6 participants). Beyond that, activation was not quite as consistent and focal, although there was some overlap where regions 1a and 1b was expected to be found. It is interesting to note that, as with general enhancement, there was response where high-frequency-sensitive region 4 was expected to be found. This spatial organization does bear some resemblance to the tonotopic scheme, and is therefore evidence that attention operates by enhancing response in frequency-sensitive regions.

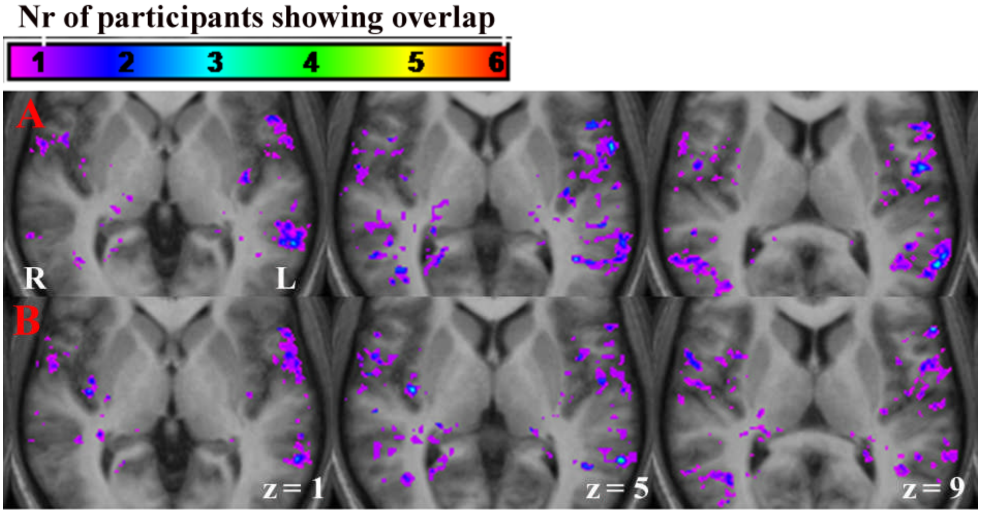


Figure 4.13 Incidence maps showing attentional suppression for six participants, for (A) high- and (B) low-frequency sounds. The contrasts performed were (A) 'just listen high-frequency majority stimulus> attend low-frequency targets in high frequency-majority stimulus' (A>E) and (B) 'just listen low-frequency-majority stimulus> attend high-frequency targets in low-frequency-majority stimulus' (B>D).

Figure 4.13A and B show the incidence maps for general effect of suppression for the high- and the low-frequency sounds (A>E and B>F) respectively. There was no reliable effect of suppression across the auditory cortex for either contrast. The only site that showed suppression was in a region that appears to correspond to the posterior part of temporo-parietal-occipital junction (TPO) bilaterally (Figure 4.14). The locus of this region was identical for suppression for the high and the low-frequency sounds. The overlap across participants was very high (5/6 participants).

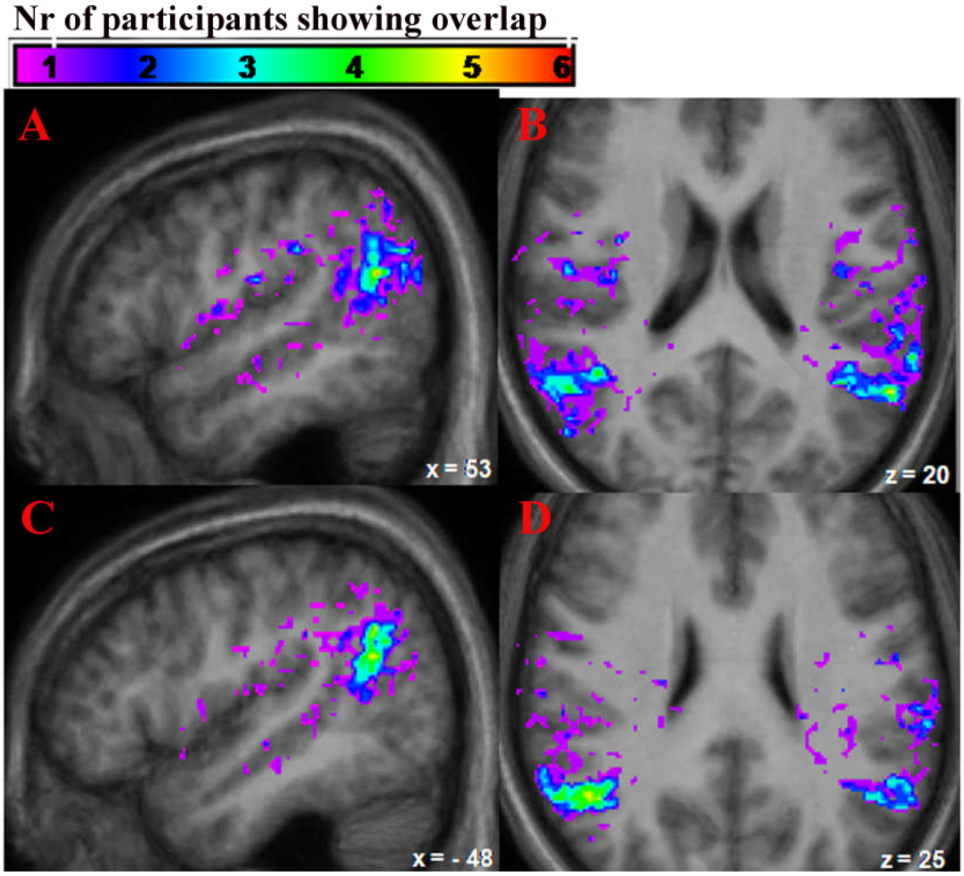


Figure 4.14 Incidence maps showing the most consistent effect of suppression for high- (A and B) and low-frequency sounds (C and D).

Plotting the beta values across conditions for the peak voxels of the activated regions in this region, showed that, typically, response for the two passive listening conditions was around 0, while response for all ‘attend’ conditions was below 0. Figure 4.15 shows a representative example of the peak voxel of the left TPO cluster of participant #5.

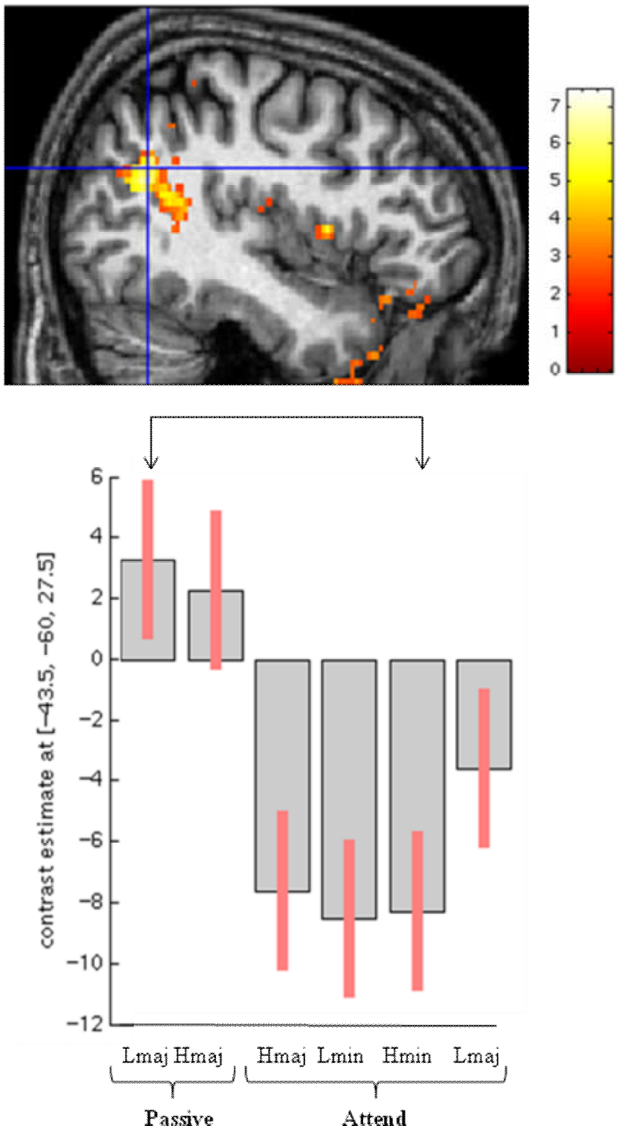


Figure 4.15 Representative example of suppression for the 'attend' conditions in TPO shown for the peak voxel of participant #5 in the left hemisphere. The crosshair on the figure indicates the location of this voxel. The graph shows the contrast estimates for all six conditions, only for the second fMRI session. The arrows indicate the two conditions that were contrasted: 'passive listening low majority stimulus> attend high frequency targets in low-frequency-majority stimulus'. The functional results are overlaid on the participant's anatomical scan. Hmaj: high majority. Lmaj: low majority. Hmin: high minority. Lmin: low minority.

4.3.5 Extent of the frequency-sensitive response and attentional modulation

This analysis quantified the extent of activity and its overlap evoked by the different contrasts (Figure 4.16) in primary and non-primary auditory cortex. The results were collapsed according to the BF of a region, because contrasts for both high- and low-frequency majority stimuli showed the same trend in terms of attentional modulation. For example, the condition showing general enhancement for the high-frequency sounds (C>A) and the condition showing general enhancement for the low-frequency sounds (F>B) were placed in the same category. First of all, the effect of general enhancement showed the most widespread of response (1915 voxels), followed by the attention-specific enhancement (747 voxels) and lastly the effect of frequency-sensitivity (583 voxels). In total, 98 out of the 583 voxels showing frequency-sensitive response also showed general enhancement (17%), while 113 voxels out of 583 shows controlled attention-specific enhancement (19%). Conversely only 5% (98/1915) of the voxels that showed an effect of general enhancement, and 15% (113/747) that showed attention-specific enhancement also showed frequency-specific response. This suggests that most attentional modulation occurred outside the frequency-sensitive regions, although it includes some of these frequency-sensitive regions.

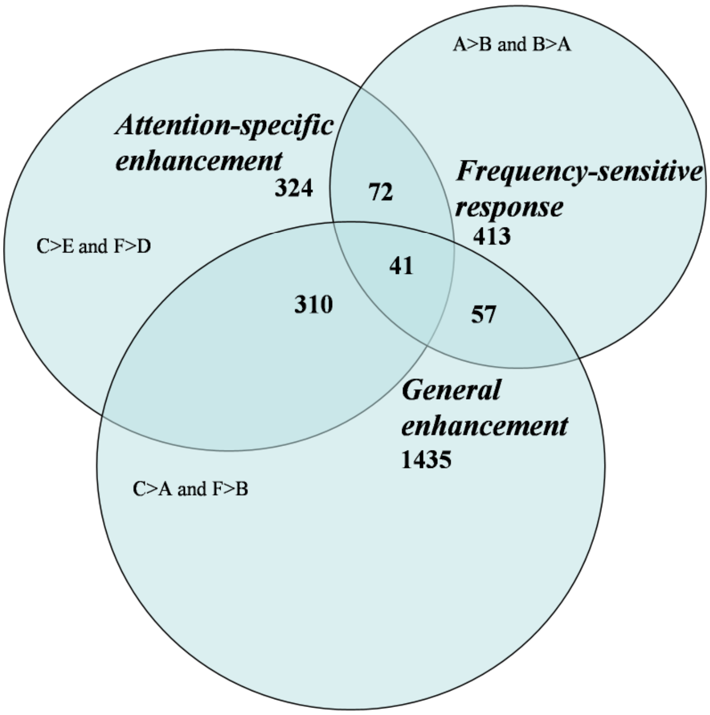


Figure 4.16 Number of voxels showing frequency-sensitivity and attentional enhancement, averaged across the six participants.

4.4 Discussion

In this study, the hypotheses of frequency-specific attentional enhancement and suppression were investigated in terms of magnitude, distribution and extent of frequency-sensitive responses and attentional modulation. The first aim of the study was to localize these regions within the auditory cortex of each participant. The responses for both high- and low-frequency sounds were restricted around primary auditory cortex. This was expected, as the sounds were narrow-band noise bursts, to which the primary auditory cortical neurons are sensitive to. All regions of interest were localized (1a, 1b, 1, 2, 3), except for region 4 in the incidence maps analysis. It is worth mentioning the results for high-frequency region 4. First of all, it did not come up in the incidence maps of frequency sensitivity (Figure 4.10), and had to be excluded from the ROI analysis, because the pattern of results was totally different from the other regions. Additionally, it came up in the both general (Figure 4.11C) and attention-specific (Figure 4.12B) enhancement contrasts for both high- and low-frequency sounds. This is in line with findings whereby medial portions of HG respond to both high- and low-frequency sounds, although slightly higher for the high (Schönwiesner et al., 2002). Interestingly, region 4 is not part of the scheme of Talavage et al. (2004) for tonotopic gradients across the auditory cortex. Perhaps it forms a functionally distinct region characterized by a preference for high frequencies.

The second and most important aim was to investigate whether there was a feature-specific attentional modulation in the frequency-sensitive regions, in terms of attentional enhancement and suppression. The ROI analysis indicated that there was a larger response when attending to the best frequency of a region, than when attending off best frequency. This is consistent with the hypothesis of frequency-specific, attention-specific enhancement. These results contrast those of Petkov et al. (2004), who reported attention-related modulation only in non-primary auditory cortex, and not organized in a frequency-specific way. However, the results reported here are more in keeping with the predictions made by the neurophysiological results reported by Fritz et al. (2007). Although the incidence maps of the effect of general enhancement did not show much response, the unmasked general effect, although widespread, seemed to be particularly consistent across participants in the frequency-sensitive regions, especially in region 2. Additionally, the incidence maps for attention-related enhancement, revealed a pattern of activation that very much resembles the tonotopic scheme.

A third aim was to investigate evidence for widespread attentional modulation across the auditory cortex, and not just in the frequency-sensitive regions. There was a widespread effect of general enhancement for both high- and low-frequency sounds, across primary and non-primary auditory cortex, as well as in frontal regions bilaterally. In contrast, in the attention-specific enhancement, there was no response in frontal regions and PT. This indicates that these regions showed probably task-related activation. Additionally, the analysis that looked at the number of voxels activated for the two enhancement contrasts and the sensory-response contrasts across the auditory cortex, showed that most attentional enhancement occurred outside the frequency-sensitive regions. Thus, this analysis revealed that, although there is enhancement in the regions that are sensitive to the attended frequency, this enhancement is not restricted to these regions, but extends along the primary and non-primary auditory cortex.

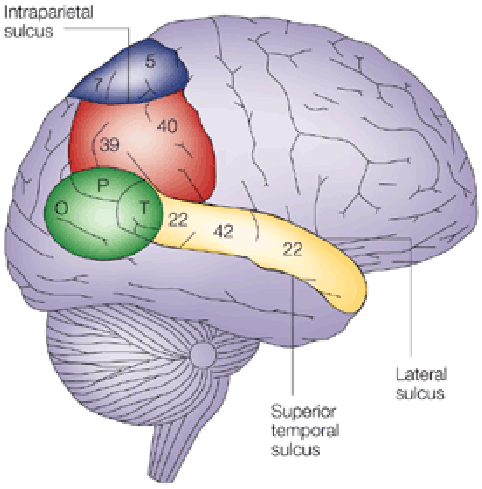


Figure 4.17 Schematic illustration of the approximate location of TPO. T=temporal, P=parietal, O=occipital. (Karnath, 2001).

In contrast with attentional enhancement, no consistent effect of suppression was found in the auditory cortex. The only consistent suppression was in an area that appears to correspond to posterior part of TPO. Further analysis showed that there was actually deactivation for all ‘attend’ conditions compared to the passive listening conditions. Note that there is evidence that lesions in TPO (including TPJ) are associated with spatial neglect (Samuelsson, Jensen, Ekholm, Naver, & Blomstrand, 1997; Vallar & Perani, 1986). Furthermore, a similar effect shown in the current study, has been shown in previous studies of visual selective attention on healthy volunteers, although more anteriorly, in right TPJ; i.e., deactivation in TPJ during performing a visual search task containing only distractors (Shulman, Astafiev, & Corbetta, 2004; Shulman et al., 2007; Shulman et al., 2003). In a study by Serences et al. (2005) participants attended to objects presented in the midline and had to detect a particular target, while presented with peripheral distracter. TPJ responded to these distractors only when they had a feature that was common with the target, and thus captured the attention of the participants. Furthermore, enhancement of response has been observed in TPJ when passively listening to stimulus transitions (Downar et al., 2001). Specifically, Downar et al. (2001). presented participants with visual, auditory and tactile stimuli, unrelated to each other, each of which was changing at a different point. For example, the visual stimulus would change color. The stronger activation for all types of transitions was observed in TPJ, which seemed to respond to salient stimulus changes. According to all the above evidence, TPJ appears to be activated during bottom-up capture of attention (Downar et al., 2001), and deactivated during top-down attention, when irrelevant stimuli are present (Shulman et al., 2007). Note that all the above studies showed these effects in the right TPJ. In summary, the deactivation of response in TPO for the ‘attend’ conditions gives some more evidence that this region is involved in filtering out unwanted information and extends this finding to auditory attention. This is tentative evidence for a parallel mechanism of filtering out unwanted information to vision, which has been shown to involve a more anterior region, TPJ.

In summary, Experiment 1 showed evidence for frequency-specific attentional enhancement in the primary auditory cortex, when attending to the best frequency of an area, compared to attending off best frequency. Furthermore, the general effect of attention (attend>passive) showed widespread enhancement across the auditory cortex.

It is interesting to note that it was mainly primary auditory cortex that showed both general and attention-specific enhancement in frequency-sensitive regions, while the non-primary auditory cortex showed more widespread general enhancement. This was expected, as the tonotopic organization of the non-primary auditory cortex is more diffuse. An interesting question is whether the lack of feature-specific modulation in non-primary auditory cortex was due to reliance on tone stimuli (Kaas, Hackett, & Tramo, 1999; Merzenich & Brugge, 1973). Perhaps if complex sounds were used, an effect of feature-specific attentional enhancement might be seen in non-primary auditory cortex. This study is discussed in Chapter 5.

Chapter 5: Experiment 2: Attentional modulation when selectively attending to complex sounds

5.1 Introduction

In this study, the hypothesis of feature-specific selective enhancement was tested when attending to spatial and non-spatial aspects of an auditory object. Feature-specific visual selective attention appears ubiquitous since localized increases in activity have been reported in many different visual regions when attention is directed to either spatial or non-spatial features that are processed within that region. For example, Corbetta et al. (1990, 1991) reported attentional modulations in distinct regions of extrastriate visual cortex that were each specialized for processing information related to the selected visual feature (color, motion and shape). Moreover, these attention-related increases in activity appear to be restricted to the cortical region that is specialized for processing the selected feature (i.e. they are not present when attention is directed to a different feature of the same visual object) (Schoenfeld et al., 2007).

In the auditory system, demonstrations of a general attentional enhancement are prevalent. Widespread increases in auditory cortical activity are demonstrated whenever the task requires attending to a sound compared with passive listening (Ahveninen et al., 2006; Grady et al., 1997; Hall et al., 2000; Johnson & Zatorre, 2005) or compared with attending to a competing visual stimulus (Degerman et al., 2006; Johnson & Zatorre, 2005, 2006; Petkov et al., 2004). Although these studies show widespread enhancement in non-primary auditory cortex, there is little evidence for its feature specificity. For example, fMRI studies have shown increased activity when listening required an active response compared to a passive state, but the location of the attention-related increase appears different from the region in which the acoustic properties of sound feature are believed to be analyzed (Hall et al., 2000; Petkov et al., 2004).

The results of two recent neuroimaging studies of the central auditory system claim to support feature-specific auditory selective attention. In a combined MEG and fMRI adaptation study, Ahveninen et al. (2006) used a discrimination task requiring judgments of phoneme identity or spatial location in order to tease apart the differential effects of attention in the anterior (‘what’) and the posterior (‘where’) processing streams. In posterior auditory cortical regions, activity was found to be significantly greater when attending to location than to phoneme identity, for both fMRI and MEG measurements. In anterior auditory areas, selectively attending to the phonetic content produced analogous effects but only for the MEG measurements. In an fMRI study, Krumbholz et al. (2007) used a task that required listeners to detect a reversal in the direction of a pitch sweep or of spatial motion. In the posterior motion-sensitive area, activity was significantly greater when attending to motion changes than to pitch changes. However, there was no corresponding effect in the pitch-sensitive area.

Two particular observations from these studies are worth highlighting. These observations offer informative recommendations for future investigation and they have explicitly been addressed these in the present study. First, in the study by Ahveninen et al. (2006), supplementary fMRI information (supporting

Figure 6) indicates that the response within both anterior and posterior auditory cortical regions was significantly increased by the two attention-demanding tasks relative to a passive listening (‘ignore’) baseline and this general enhancement was much greater than the reported effect of feature-specific auditory selective attention. Although not the main focus of the author’s interpretation, this result is more consistent with the interpretation that selective attention broadly enhances the auditory cortical response, and not in a feature-specific manner. Thus, one recommendation is for future investigations to systematically evaluate attentional effects not only relative to one another, but also relative to a baseline context such as passive listening (Johnson & Zatorre, 2005). The second observation relates to the lack of attentional modulation of non-spatial auditory features. In the study by Krumbholz et al. (2007), we speculate that the lack of an attentional effect in the pitch-sensitive region could be attributed to the relative ease of the pitch judgment compared to the difficulty of the spatial judgment. It is known that task difficulty affects attention modulation. For example, performing a difficult spatial orientation task enhances the size of the selective attention effect (Boudreau et al., 2006).

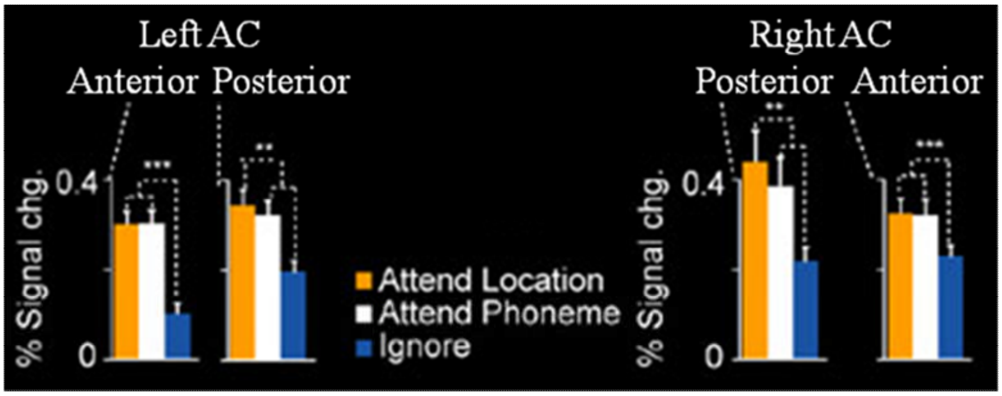


Figure 5.1 Attentional modulation in anterior (‘what’) and posterior (‘where’) auditory cortex for ‘attend location’ (orange), ‘attend phoneme’ (white) and ‘ignore sounds’ (blue). Adapted from Ahveninen et al. (2006).

The present study re-examines the neural correlates of feature-specific auditory selective attention using two sound features that are known to engage distinct auditory cortical regions. As discussed in Section 2.3.1, sounds that are modulated in the frequency domain at a slow rate (5 Hz or so) elicit activity around the lateral portion of HG and lateral PT (Binder et al., 2000; Griffiths, Buchel et al., 1998; Hall et al., 2002; H.C. Hart et al., 2004; Thivard, Belin, Zilbovicius, Poline, & Samson, 2000). Furthermore, sound sequences that contain discrete shifts in spatial location are known to elicit activity in posterior non-primary regions of PT (Baumgart et al., 1999; Deouell, Heller, Malach, D'Esposito, & Knight, 2007; H.C. Hart et al., 2004; Krumbholz, Schonwiesner et al., 2005; Lewis, Beauchamp, & DeYoe, 2000; Pavani et al., 2002; Warren et al., 2002). The effects of feature-specific auditory attention, i.e. enhancement and suppression, were analyzed within specific FM- and motion-sensitive regions that had been defined in our group of participants. In Experiment 2 the listening conditions also included active listening to either a motion target or an FM target and also passive listening where there was no target to attend to. This enabled me to assess both attention-related as well as general enhancement respectively (same logic as in Experiment 1). Finally, care was taken to equate the difficulty of the two tasks for each participant in successive behavioral screening sessions, by varying the parameters of the targets according to the participants' performance.

A secondary goal was to explore the way in which the listening task might influence the location and extent of the FM- and motion-related activation. A relatively unusual feature of the experimental design used by Krumbholz and her colleagues (2007) was that it took care to map out the pitch- and location-sensitive regions by presenting pairs of sound conditions in which attention was controlled by requiring subjects to attend to an irrelevant auditory feature in both conditions (but see also Petkov et al., 2004). This approach may be preferable to a passive listening paradigm which is behaviorally unconstrained and so the cognitive component cannot be fully excluded from the sensory component. Krumbholz et al. argued that controlling for attention better eliminated the effects of higher-order cognition on sensory activity and thus lead to more focal and feature-specific maps than passive listening. On the other hand, there is evidence that passive listening provides a good representation of the sensory response to sounds (Newman, Twieg, & Carpenter, 2001; Peck et al., 2004; Scott, Malone, & Semple, 2007; Shulman et al., 1997). Furthermore, there is evidence that when attending to one feature of an object, then other unattended features of the same object are also attended (O'Craven et al., 1999), and so asking participants not to do any task might more representative of sensory response than asking them to attend to a ‘distractor’ that exists in the same object.

Another secondary goal was to examine whether performance accuracy was compromised when attention was divided across the two feature dimensions. This was also done in Experiment 1, and showed that there was a cost when attending to both ‘high’ and ‘low’ sounds. However in Experiment1 participants had to divide their attention between two objects, while in this experiment both features are part of the same objects. Thus, due to the ‘same object advantage’, it is possible that attending to both features might show better performance than attending to one feature at a time (Duncan, 1984). For this reason, six participants completed an additional task in which they were instructed to ‘attend both FM and motion targets’.

5.2 Methods

5.2.1 Participants

Sixteen normally hearing (≤25 dB hearing level in octave steps from 250 to 8000 Hz) participants with normal, or corrected to normal vision took part in both the behavioral screening and the fMRI experiment. There were seven males and nine females with a mean age of 24.7 years (range 18-37). All but one participant were right-handed and three were non-native English speakers, but had no difficulty understanding the task instructions. None of the participants had any history of neurological problems. Seventeen additional participants took part in the behavioral screening, but were rejected for one of two reasons. A majority of the participants (N=15) failed to reach adequate performance on one of the target discrimination tasks (usually the motion task). Here adequate performance was defined as a score exceeding the 40% hits-FA rate. Two further participants completed the behavioral screening, but then could not fit comfortably in the SENSE receiver head coil. The experimental procedures were approved by the University of Nottingham Medical School Research Ethics Committee and all participants gave written informed consent.

5.2.2 Stimuli, tasks and apparatus

Sequences of 18 harmonic-complex tones (fundamental frequency = 400 Hz, harmonics 1-5, 400 ms duration, 10 ms onset and offset ramps, 50 ms inter-stimulus interval) were digitally created using Matlab (version 7.0.1). The total duration of each tone sequence was 8 s. Spatial (horizontal motion) and non-spatial (FM) acoustic features were manipulated to create three sound conditions, (i) moving FM, (ii) stationary FM and (iii) moving steady state. The parameters of the frequency modulation were a rate of 5 Hz and a depth of 12.5% of the fundamental frequency. For the steady-state sequences, the complex tones had no frequency modulation. The percept of a sound source that swept horizontally back and forth was generated using inter-aural time difference (ITD) cues in successive (150 μs) steps. For half of the ‘moving’ sequences, the sound source started on the left (ITD = -600 μs), while for the other it started on the right (ITD = +600 μs). The tones moved to the opposite side, and then went back to where they started, thus completing a whole cycle in 8 s. For the stationary sequences, all tones had ITD of 0 μs and were perceived at the midline.

In the fMRI experiment, the auditory stimuli were presented through custom-made MR-compatible headphones fixed inside Bilsom ear defenders. The intensity of the sounds was 96.1 for the left, and 94.1 dB SPL for the right ear. Listening instructions were presented, via a projector, onto a screen which hung from the ceiling about 226 cm away from the participant’s eyes. Participants could view the screen using two small mirrors that were mounted on the head coil. The span of the instructions was 17 cm across, subtending a visual angle of 4°, and the instructions were clearly visible for all participants.

For the ‘just listen’ and ‘silence’ conditions, a fixation cross was presented at the centre of the screen. Participants were instructed to fixate on the cross.

	Listening Instructions		
Stimulus features	Just listen	Attend FM	Attend Motion
Moving/ FM	A	B	C
Moving/ Steady state	D		E
Stationary/ FM	F	G	

Table 5.1 Experimental conditions

The three sound conditions were partly crossed with three listening tasks (i) ‘just listen’, (ii) ‘attend FM’ and (iii) ‘attend motion’, to create seven experimental conditions as shown in Table 5.1. Conditions will be referred to by their alphabetical label. Those tone sequences comprising the ‘just listen’ task (A, D, F) did not contain any stimulus targets. The task in the ‘attend FM’ condition required participants to respond to the tone that had a frequency-modulation depth that was lower than the reference (B, G). The size of the depth was pre-determined on the basis of the individual performance during behavioral screening and ranged from 8.5 to 11.5%. The task in the ‘attend motion’ condition required participants to respond to the tone that jumped in the opposite direction to the arc of the motion (C, E). Again, the step size of the jump varied across participants from 100 to 350 μs and was determined during the behavioral screening sessions. The moving-FM condition contained both spatial and non-spatial targets, either one or two of each. In other words, each sequence could contain one motion and one FM target, two motion and one FM target, or one motion and two FM targets. Targets in the other two sound conditions occurred in only one feature dimension. In other words, the sequence of stationary FM tones contained only (one or two) FM targets, while the sequence of moving steady-state tones contained only motion targets. Targets could appear in any position in the sequence, except the first two or last two tones. Participants were instructed to attend and respond only to the relevant target dimension and to press a button with their right thumb as soon as they detected the target. Participants' responses were recorded for off-line analysis.

Behavioral screening took place in a sound-attenuated room and the target detection training always used the 'moving FM' tone sequences presented through Sennheiser (type HD480II) headphones at a sound level of 86.6 dB SPL. Participants were instructed either to 'attend FM' or 'attend motion'. Instructions were presented on a 15-inch flatscreen and the order was randomized across tone sequences. Each instruction appeared at the beginning of the tone sequence and disappeared at the end. Participants' button press responses were recorded for off-line analysis. Each training session comprised of two runs, each containing 26 repeats of the two listening tasks. Each run lasted approximately seven minutes. The goal was to equate performance across the two tasks (i.e. no more than 5% difference in hits-FA rate), whilst also ensuring that performance fell between 60 and 90%. On subsequent behavioral sessions, subjects repeated the tasks, but this time using targets whose parameters had been adjusted to approach matched performance. Some participants needed only one session to identify the appropriate target parameters, while others needed to complete up to four sessions. Therefore, in the fMRI experiment, participants were presented with targets defined by different acoustic values for FM depth and motion jump to ensure a similar level across the two listening tasks. To examine whether performance accuracy was compromised when attention was divided across the two feature dimensions, six participants completed an additional task in which they were instructed to ‘attend both FM and motion’.

5.2.3 Imaging protocol

For each participant, a T1-weighted anatomical scan was acquired to facilitate processing of the functional image data. The anatomical scan was composed of 160 sagittal slices at 1 mm^3 resolution (matrix size: 128x128) and was completed in about 5 minutes. For one participant, a rapid 60-slice functional image was acquired instead of the anatomical scan. The anatomical scan was used to select the orientation of the functional images. Each scan consisted of 32 oblique axial slices at 3 mm^3 resolution (matrix size: 64x64, flip angle: 90°, TE: 36 ms, acquisition time: 1971 ms) to include the whole brain. The lower slice cut across the cerebellum and prefrontal cortex, while the upper slice cut across superior parietal cortex (leaving out a small part of parietal cortex, at the top). A sparse sampling protocol was used (repetition time TR=10 s; Hall et al., 1999) to avoid the temporal overlap between the intense scanner noise and the tone sequence. A total of 106 scans were collected in each experimental run. The experiment was divided into two 17-minute runs. Each of the sound conditions were repeated either 12 or 14 in each run, so that every condition was presented a total of 26 times. An additional condition in which no sound stimulus was presented was also included as a baseline. Conditions were presented in a pseudo-randomized order that changed every two trials. The order of conditions was counterbalanced for each participant.

5.2.4 Analysis of target detection data

Target-detection accuracy and reaction times were analyzed separately and a response was considered correct if its reaction time was ≤1200 ms post-stimulus onset. To enable direct comparison with the results reported by Krumbholz et al. (2007), accuracy scores were transformed into a percentage measure of the hit- minus false-alarm rate. The fMRI performance measures were subjected to repeated measures ANOVA to test for differences across stimulus conditions.

5.2.5 Image analysis

Image pre-processing and analysis was performed on a Sun Ultra 2 computer (Sun Microsystems) using SPM2 software (http://www.fil.ion.ucl.ac.uk/spm/) running in MATLAB v6.5 (Mathworks, Natick, MA, USA). The functional scans were motion corrected to account for head movements both within and between the two runs (Friston et al., 1995) using the last scan of the first run as a reference. Head movements did not exceed 3 mm (translation) and 3° (rotation). Individual anatomical scans were used to compute the transformation into standard brain space. The first step was to co-register the anatomical scan with the mean realigned functional image. The second step was to automatically segment the anatomical scan into its grey matter component which was then spatially transformed to match the SPM2 grey-matter template. Finally, the same transformation parameters were applied to both the anatomical and the functional scans for that individual. The normalized anatomical scan preserved its voxel resolution of 1 mm^3, while the functional scans were upsampled to 2 mm^3. EPIs were smoothed by a Gaussian kernel of 8 mm FWHM.

The individual data were modeled within the framework of the GLM. The data were modeled using fourteen regressors; seven describing each of the experimental conditions of interest, six describing the realignment parameters to account for any residual head motion and one to account for session differences between the two runs. The 'silent' baseline condition was implicitly modeled. Low-frequency artifacts were removed by high-pass filtering the time series. The high-pass filter cut-off varied from 580 to 780 s across participants, depending on the order of the experimental conditions.

A random effects analysis was conducted to account for both within and between participants variability and to look at the brain regions activated for the majority of the participants. A number of statistical contrasts (shown in Table 5.2) were performed on the individual data to enable the analysis of the main sensory effects of FM and motion and the main effects of auditory selective attention at the group level (Holmes and Friston, 1998). The contrasts are defined using the alphabetical labels given in Table 5.2.

	Conditions contrasted	Functional response
A>D	Just listen (stimulus: Moving/FM)> Just listen (Moving/Steady state)	FM sensory response (passive listening)
C>E	Attend Motion (Moving/FM)> Attend Motion (Moving/Steady state)	FM sensory response (controlled for attention)
A>F	Just listen (Moving/FM)> Just listen (Stationary/FM)	Motion sensory response (passive listening)
B>G	Attend FM (Moving/FM)> Attend FM (Stationary/FM)	Motion sensory response (controlled for attention)
B>A	Attend FM (Moving/FM)> Just listen (Moving/FM)	General enhancement when attending to FM
C>A	Attend Motion (Moving/FM)> Just listen (Moving/FM)	General enhancement when attending to Motion
B>C	Attend FM (Moving/FM)> Attend Motion (Moving/FM)	Attention-specific enhancement when attending to FM
C>B	Attend Motion (Moving/FM) > Attend FM (Moving/FM)	Attention-specific enhancement when attending to Motion

Table 5.2 Functional contrasts. Note that in the 'just listen' conditions, there are no FM or motion targets.

FM-sensitive regions were defined using two pairwise contrasts; one during passive listening (A>D) and one while attending to motion (an irrelevant feature) (C>E). Motion-sensitive regions were defined in the same manner using two other pairwise contrasts (A>F and B>G). The effects of FM and motion were identified using a conjunction approach that was performed at the group level using a repeated measures one-way ANOVA. Conjunctions identify those voxels that are activated in several different pairwise comparisons (Price & Friston, 1997). Hence, the voxels showing a response to the feature of interest irrespective of the listening instructions were identified ('just listen', or 'attend to the irrelevant feature'). Specifically, the FM-sensitive regions were defined by the conjunction of 'A>D' and 'C>E', while the motion-sensitive regions were defined by the conjunction of 'A>F' and 'B>G'. Previous findings enabled us to make specific anatomical predictions about the spatial location of the sensory activation and so results are reported at a voxel threshold of p<0.001 (Friston, Holmes, Poline, Price, & Frith, 1996). Applying correction for the whole brain would be inappropriately conservative, since the conjunction analysis is already a stringent analysis. The analysis that compared the controlled and passive mapping contrasts were computed at the same threshold.

Analysis of the feature-specific effects of selective attention first considered the magnitude of the response within those FM- and motion-sensitive regions that had been identified by the preceding conjunction analyses. This ROI analysis was particularly concerned with the effect of the listening instructions on the magnitude of the response for the same sound stimulus (namely conditions A, B and C). For each participant, the mean GLM parameter estimates were extracted for each condition within the two ROIs and subjected these to repeated-measures ANOVA with task and hemisphere as independent variables.

It is also informative to report whether or not the effects of selective attention extend beyond the feature-sensitive regions within the auditory cortex. Two different types of attention comparison are reported. First, the general effect of attention was investigated, by contrasting each attend condition with 'just listen' (i.e. B> A and C>A). Second, the differential effect of attention was investigated by directly contrasting the two 'attend' conditions (B>C and C>B). In the absence of a priori anatomical predictions these results are reported at a threshold of p<0.05, corrected for multiple comparisons using the false discovery rate (Genovese et al., 2002). Note that the responses of the above contrasts were superimposed on a probability atlas which contains maskers of several brain regions, using an SPM toolbox (Eickhoff et al., 2005). This method was used to identify the location of the clusters of activation across the brain.

5.3 Results

5.3.1 Target detection accuracy and reaction times

The non-significant effect of listening task on target detection performance during the fMRI experiment (F[3, 45]=1.9, p>0.05) confirmed that task difficulty was well matched across conditions (Figure 5.2, upper panel). On the other hand, analysis on reaction times (Figure 5.2, lower panel) showed a significant effect of condition [Wald(3)=19.78 , p<0.001]. The paired-sample t-tests showed that participants responded to motion (C, E) significantly slower than to FM targets (B, G) (p<0.015). In Krumbholz et al.'s (2007) study, the 'attend-motion' condition showed significantly better accuracy, but slower reaction times compared to the 'attend pitch conditions.

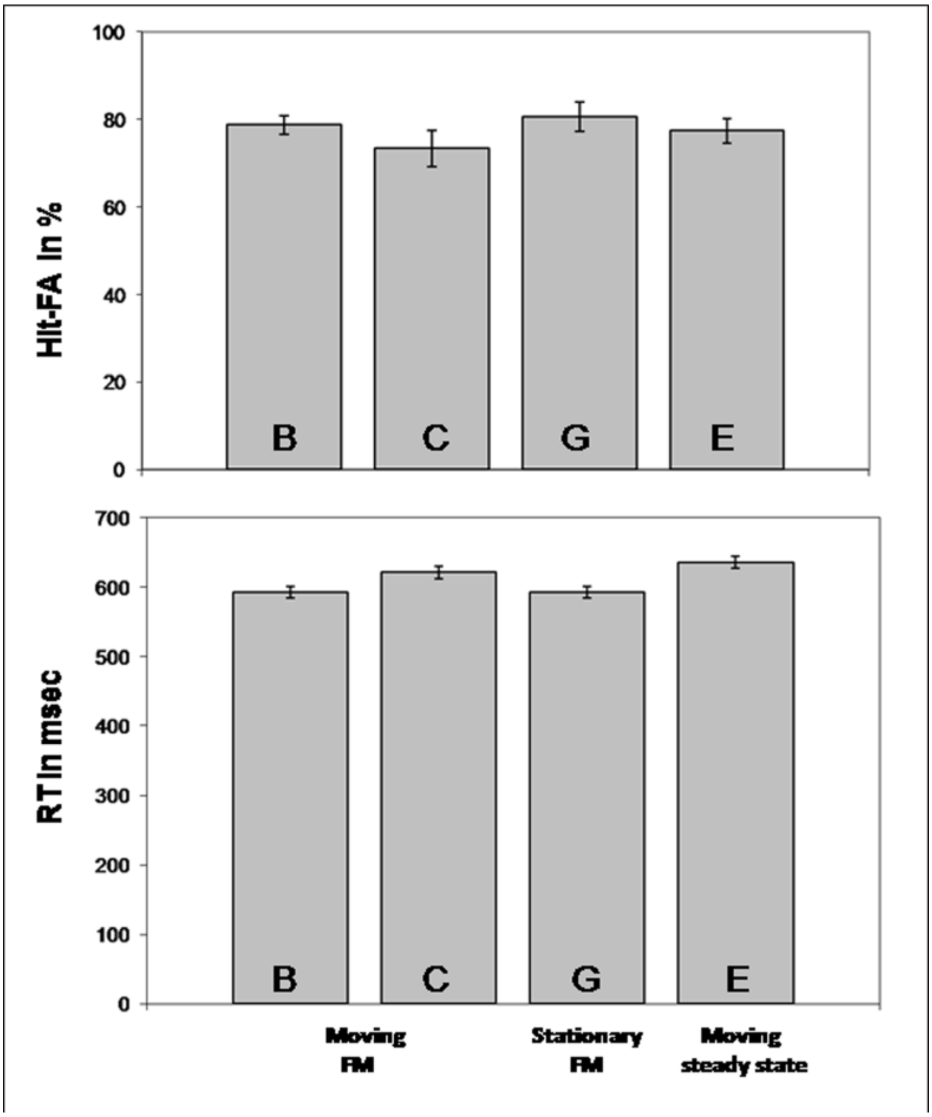


Figure 5.2. Hit-FA rate (upper panel) and RT (lower panel) for the four active conditions that participants performed in the scanner. Error bars represent standard errors for the 16 participants.

The reason for the slower reaction times in the ‘attend motion’ condition might be explained by the distributions of reaction times in each of the four conditions. As Figure 5.3 shows, in the ‘attend FM’ conditions, reaction times have a unimodal distribution, while in the ‘attend motion’ conditions, reaction times have bimodal distribution; that is, there seem to be two peaks, one at 400 ms and one at 800 ms. This indicates that sometimes participants might have waited for the sound after the target, to confirm that it was indeed a motion target; note that the sound after the target jumped forward to the normal direction.

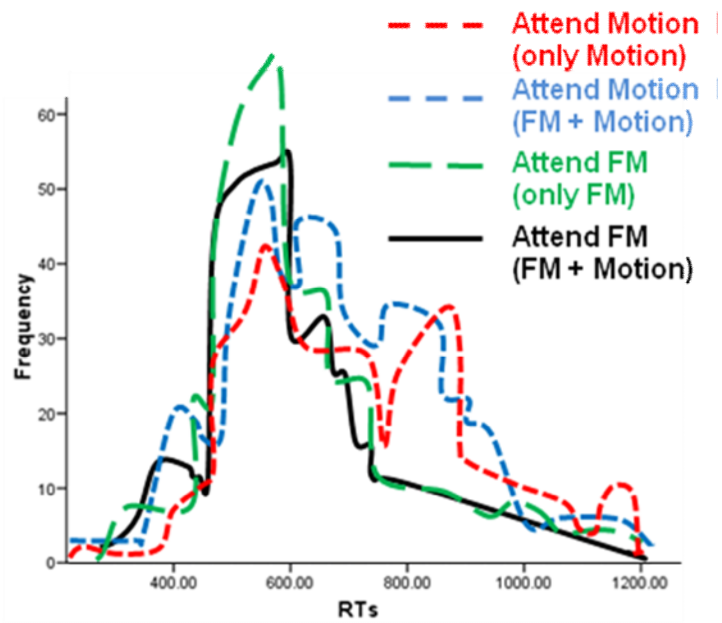


Figure 5.3 Shape of reaction times distributions for all four active conditions in the scanner.

Surprisingly, participants performed better when attending to both features (79.7% hit rate minus false alarm rate) than when attending to one of the features (74.62%). However, this difference was small.

Participant	Attend FM	Attend Motion	Attend both
1	52.95	64.39	64.16
2	78.97	75.71	100.00
3	80.00	68.61	79.67
4	71.44	83.33	59.68
5	88.25	76.92	88.15
6	79.49	77.73	86.56
Average	75.18	74.45	79.70

Table 5.3 Results for the behavioral session whereby selective and divided attention to FM and motion were compared.

5.3.2 Sensory effects of FM and motion

As expected, FM- and motion-dependent responses were found bilaterally in focal regions of the auditory cortex (Figure 5.4A). Specifically, FM sensitivity was localized in lateral and central parts of HG (within the black outline) and in antero-lateral PT (blue outline). On the left side, FM-dependent activity included 351 voxels with a peak at x -54, y -10, z 0 mm and on the right, 181 voxels with a peak at x 60, y -6, z -4 mm. Motion sensitivity was found in a distinct region of HG and PT, relatively more postero-medial to the FM response. Suprathreshold activity included 376 voxels on the left side with a peak at x -52, y -22, z 4 mm and, 60 voxels on the right with a peak at x 52, y -22, z 8 mm.

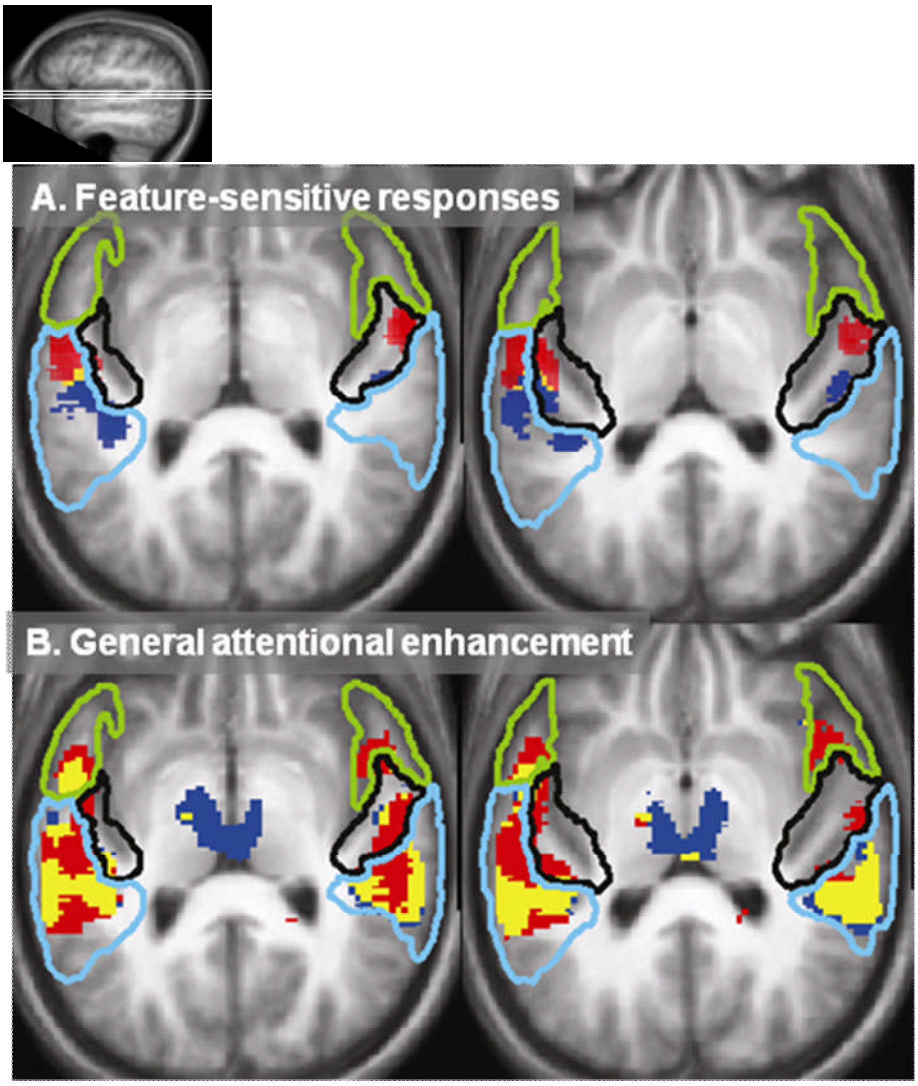


Figure 5.4 A: Sensory response for FM (red) and motion (blue) across the brain, as revealed by conjunction analysis. Yellow=overlap. B: effect of general enhancement. Red: 'Attend FM>just listen', Blue: 'Attend Motion>just listen'. (p<0.05, corrected). Yellow: overlap. Green outline: PP; black outline: G; blue outline; PT.

5.3.3 Feature-specific effects of auditory selective attention

The ROI analysis revealed that listening instructions had a significant influence on the magnitude of the response, both in the FM-sensitive region (F[1.6, 49.3]=2.0, $p<0.01$] and in the motion-sensitive region (F[1.5, 46.3]=3.9, $p<0.05$). The results are plotted in Figure 5.5. Post-hoc testing revealed that the instructions to attend to a target sound feature always significantly increased the response compared to passive listening (p<0.017, with Bonferroni correction for multiple comparisons). Within the two regions of interest, there was no evidence for any differential responses to the two attend conditions. The motion-sensitive region responded both when attending to motion targets and when attending to FM targets and vice-versa. Hence, attentional enhancement was not feature specific. Furthermore, the results do not support the existence of attentional suppression for the unattended feature. For example, in the FM-sensitive regions, there was no reduction in the magnitude of the response for the 'attend motion' condition compared to the 'just listen' condition, in fact there was a significant increase in the opposite direction. Note that the effects of suppression were not investigated further.

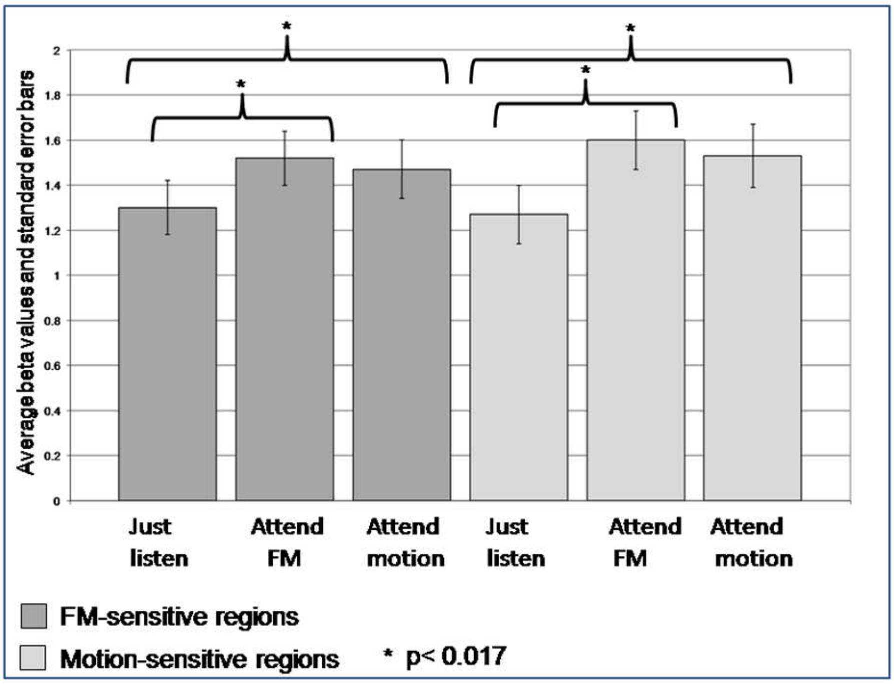


Figure 5.5 The effect of attending to FM and motion stimulus targets within regions of the auditory cortex that are sensitive to processing FM and motion features.

5.3.4 General effects of active listening

Effect of general enhancement. A group analysis of the whole-brain image data identified which regions showed significantly greater activity in the 'attend' conditions compared to passive listening (i.e. B>A and C>A, Table 5.2). The distribution of attention-dependent activation across the auditory cortex is illustrated in Figure 5.4B. Compared to the regions that respond to the FM and motion features in the sound stimulus (Figure 5.4A), attending to those same features engages widespread regions of the non-primary auditory cortex. Notably, PT responded to attending to both FM and motion, whereas attending to FM additionally extended anteriorly along the PP. Active listening also engaged a number of other brain regions (Figure 5.6) including superior, middle frontal gyri (1) inferior parietal lobe (2), inferior frontal gyrus (3), prefrontal cortex (4), primary and non-primary visual cortex (5). Response was bilateral, but more on the right hemisphere. Most of these areas showed an overlap for both contrasts. Details of the activation are provided in Table 5.4.

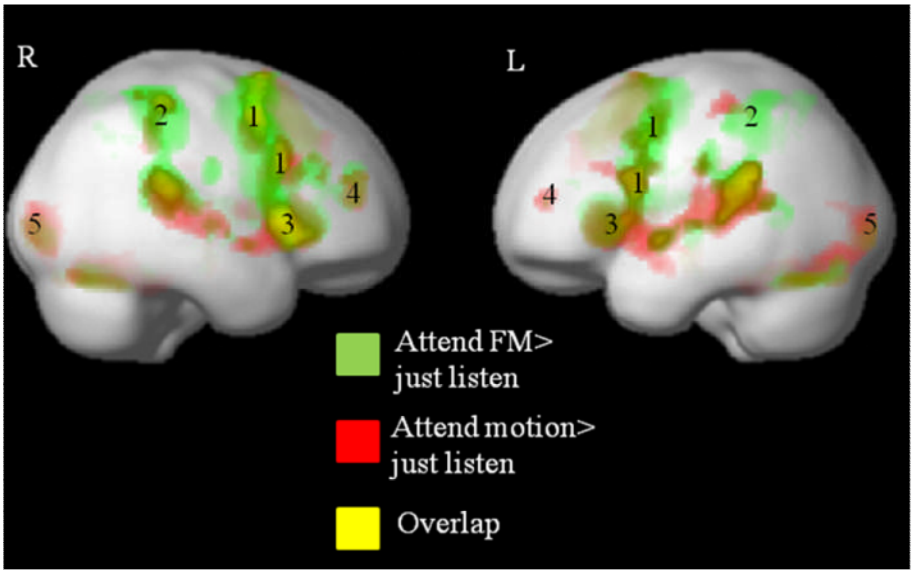


Figure 5.6 General effects of attention for attend FM (green) and attend Motion (red), superimposed on a rendered view of the average anatomical, which shows activation on the surface. Brightness of colour indicates depth of activation.

Peak MNI coordinate (mm)			Cluster size (# voxel)	Z value	Side	Putative anatomical region	Localization confidence
x	y	z					
Attend FM>just listen							
52	-2	-6	3738	5.10	L	PP, Te 1.2, Te 1.0, PT, Posterior STG, Insula	N/A
28	70	26	2748	4.71	R	Cerebellum	N/A
54	14	0	2600	4.87	R	Inferior, Middle, Superior Frontal gyrus , Insula	N/A
8	18	38	1826	4.84	R	Middle cingulate cortex, Superior and middle frontal gyrus	N/A
66	32	6	1041	4.57	R	Posterior edge of PT	N/A
16	90	-6	511	4.01	L	Around calcarine sulcus	60
52	34	54	215	3.51	R	Inferior and Superior parietal lobule Supramarginal gyrus	N/A
34	50	12	152	3.42	R	Anterior Middle frontal gyrus	N/A
48	36	61	73	3.38	L	Post-central gyrus	N/A
36	48	10	56	3.59	L	Anterior Middle frontal gyrus	N/A
46	34	20	10	2.73	R	Middle frontal gyrus	N/A
Attend Motion>Just listen							
34	28	-4	10,332	5.84	R/L	Insula, Superior and Middle Frontal gyrus, Middle Cingulate gyrus,	N/A
62	42	20	2147	5.29	L	PT, Supramarginal gyrus Inferior parietal gyrus, Postcentral gyrus	N/A
66	28	16	2125	5.30	R	PT, Supramarginal gyrus Inferior parietal lobule	N/A
10	16	-2	983		R/L	Thalamus	
16	68	22	488	4.78	R	Cerebellum	N/A
32	62	26	477	4.85	L	Cerebellum	N/A
14	88	-6	163	3.27	R	Calcarine gyrus	70
56	-2	-6	109	4.37	L	PP, Te 1.2	N/A
54	50	10	105	3.21	L	Posterolateral Middle temporal gyrus	N/A
12	66	60	83	3.20	L	Precuneus, Superior parietal lobule	N/A
62	16	24	67	3.64	R	Operculum, Inferior postcentral gyrus	N/A
10	98	-2	65	3.32	L	Calcarine gyrus	90
52	14	0	29	3.00	R	Posterior edge of central HG (Te1.0)	50
20	62	40	21	2.91	R	Superior occipital gyrus	N/A
56	2	-8	8		R	Te 1.2, Temporal pole	40
60	-8	-4	6	3.34	R	Underneath Lateral HG	N/A

Table 5.4 General attentional enhancement ($p<0.05$, false discovery rate).

Effect of targets. As noted earlier, there were no targets in the 'just listen' conditions. It is possible that this was partly the reason for the differential response in the auditory cortex when contrasting 'attend' with 'just listen' conditions. For example, the deviant target would evoke a mismatch response similar to that measured by EEG (the miss-match negativity, MMN). To investigate this, post-hoc analysis on the results of three of the subjects was conducted. These participants were chosen because their results showed bilateral activation for both contrasts. Each target was assigned a weight, according to each place within the sequence. It was assumed that the hemodynamic response to the sounds peaked at 4.5 s before the scan. So, if a target was on the ninth sound before the end, then the score was 1. If there was also another target in sound 5, this sound was scored 0.75. Then, the total score was 1+1.75=2.75 for that sound sequence. Then the GLM model was fitted for each of the three participants, including this additional regressor. For all participants, there was no response around the superior temporal gyrus (i.e. the auditory cortex). This is consistent with the findings of Deouell et al. (2007), who showed that although spatial variation evoked activation in PT, the occasional oddball target did not result in any increase of response.

Attention-specific enhancement Although the ROI analysis failed to find evidence for feature-specific attention, these results alone do not suggest the complete absence of any differential effects of active listening. Therefore, we contrasted the two attend conditions with each other (i.e. B>C and C>B). Attending to motion produced a significantly greater response than attending to FM in several brain regions (C>B) (Figure 5.7, Table 5.5), including superior and middle frontal gyrus (1), inferior parietal (2), precuneus (3), supramarginal gyrus (4), middle temporal gyrus (5). Response was bilateral, but more extended and stronger on the right hemisphere. There was no suprathreshold activation for the converse contrast (B>C), similarly to previous studies (Degerman et al., 2006; Krumbholz et al., 2007).

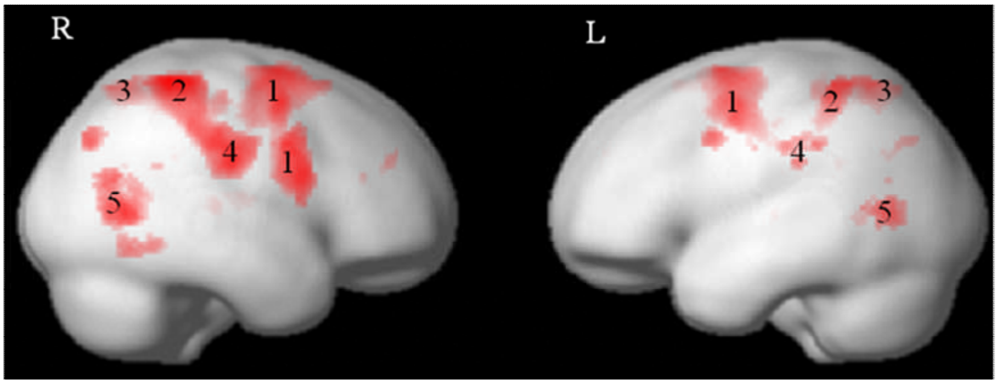


Figure 5.7 Differential effect of attention: Attend Motion> Attend FM (C>B).

Peak MNI coordinate (mm)			Cluster size (#voxels)	Z value	Side	Putative anatomical region	Localisation confidence
x	y	z					
Attend Motion>Attend FM							
30	-8	58	1829	4.72	R	Superior and Middle and Inferior frontal gyrus	N/A
38	44	60	1603	5.23	R	Precentral gyrus, Supramarginal gyrus, Postcentral gyrus	80
18	0	62	1139	5.23	L	Precentral gyrus, Central gyrus, Superior and Middle frontal gyrus	50
10	60	56	553	4.11	R	Precuneus	N/A
52	62	6	387	4.42	R	Middle temporal gyrus	N/A
40	44	54	294	4.26	L	Postcentral gyrus	50
56	68	2	85	4.05	L	Middle temporal gyrus	N/A
54	62	10	73	3.92	R	Inferior temporal gyrus	N/A
38	76	36	65	3.81	R	Middle occipital gyrus	N/A
56	28	32	58	4.04	L	Supramarginal gyrus	N/A
8	16	10	47	3.77	R	Right Thalamus	N/A
54	6	38	38	3.25	L	Precentral gyrus	50
18	26	8	21	3.47	R	Right thalamus	N/A
28	70	34	16	3.20	L	Middle occipital gyrus	N/A
24	46	28	16	3.65	R	Middle frontal gyrus	N/A
60	30	26	13	3.17	L	Operculum	50

Table 5.5 Attention-specific enhancement for motion (p<0.05, false discovery rate).

Comparison between the two mapping contrasts Finally, to examine the secondary goal regarding the influence of the listening task on the location and extent of the FM- and motion-related response, the same contrasts for the mapping of activation for FM (A>D, C>E) and motion (A>F, B>G) were performed, but without performing a conjunction this time.

As Figure 5.8 shows, there was more extended response in the auditory cortex for the ‘controlled’ localizer contrasts than for the passive ones.

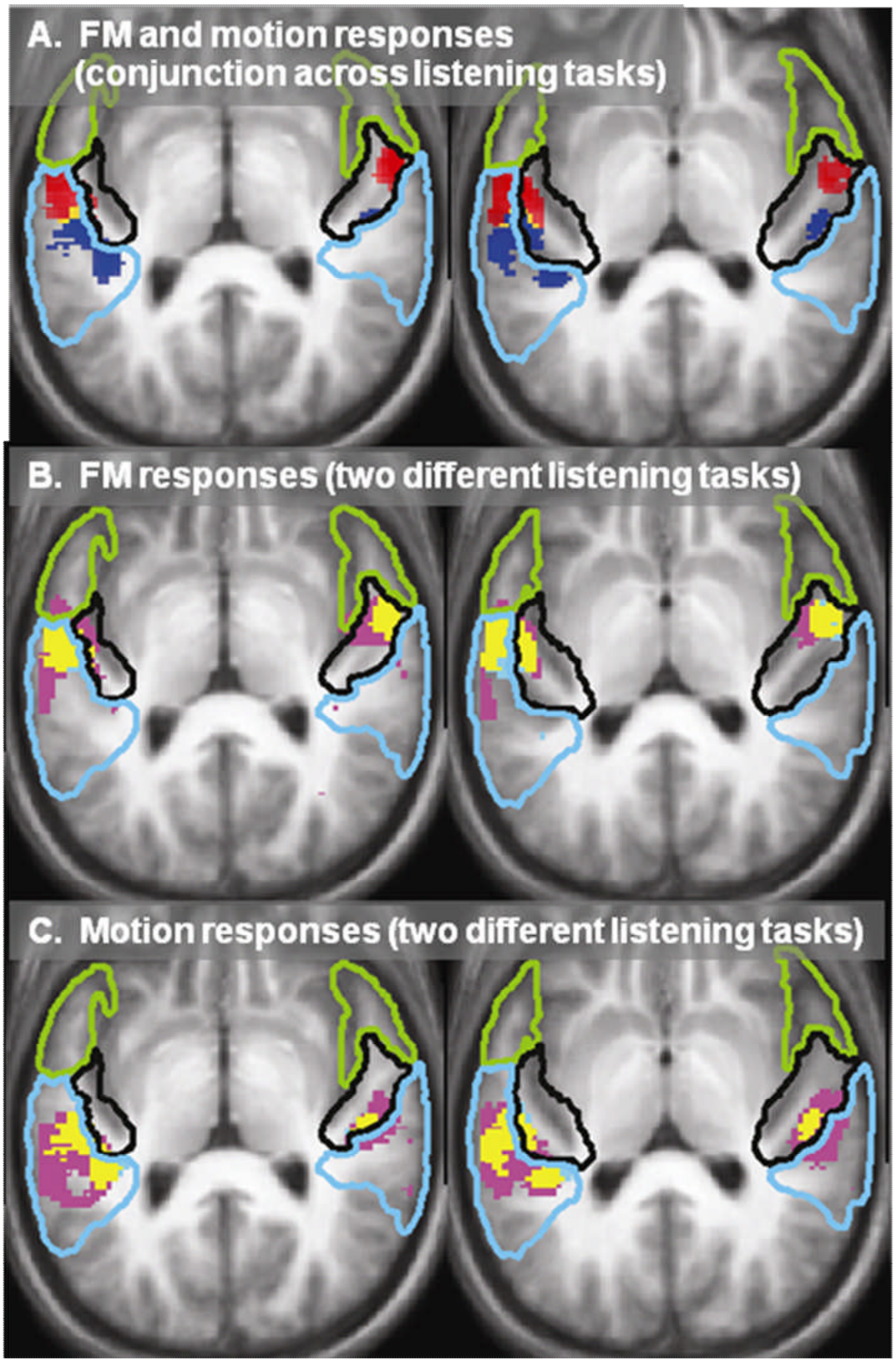


Figure 5.8 Response across the auditory cortex for the FM and motion mapping contrasts. A: conjunction. B: FM mapping for controlled (pink) and passive listening (blue) contrasts. (p<0.001, uncorrected) Yellow: overlap. C: Motion mapping for controlled and passive listening contrasts. Green line: PP; black line: HG; blue line; PT.

What about outside the auditory cortex? As Figure 5.9 shows, the controlled regions showed response outside the auditory cortex, in frontal regions, while the passive listening contrast showed response only within the auditory cortex.

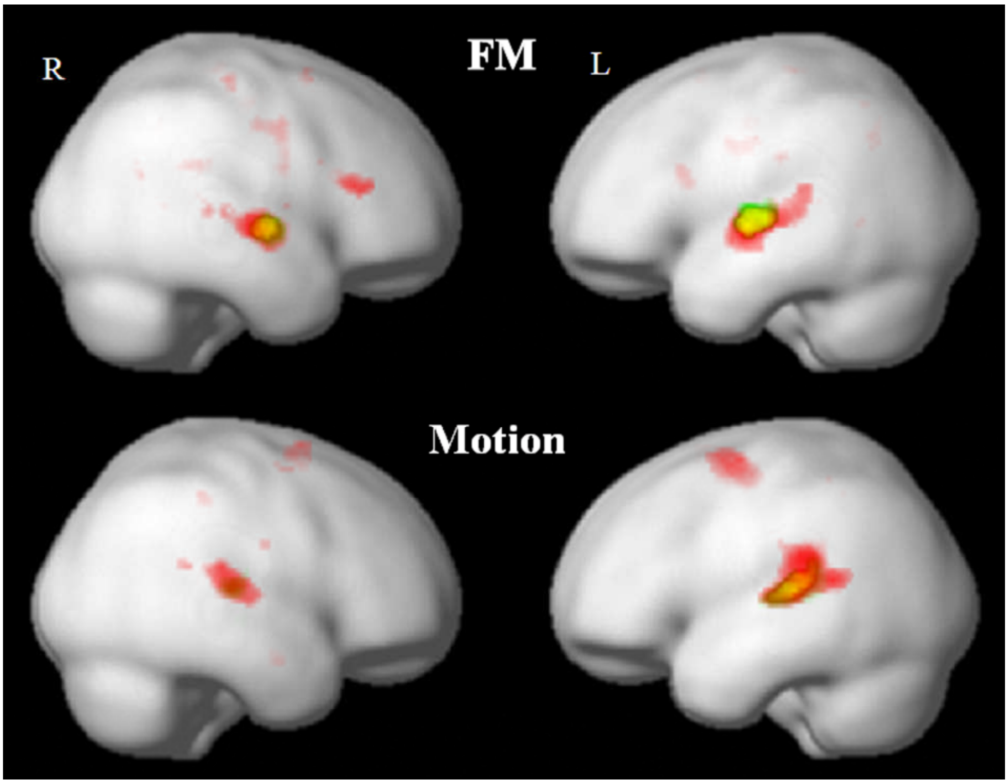


Figure 5.9 Response across the brain for the mapping contrasts. FM (upper panel) and motion (lower panel) mapping for controlled (red) and passive listening (green) contrasts (p<0.001, uncorrected). Yellow: overlap.

5.4 Discussion

In Experiment 3 the hypothesis of feature-specific attentional modulation was investigated when attending to spatial (auditory motion) and non-spatial (FM) features of an auditory object. The results revealed no evidence for this. Instead, there was evidence for widespread general enhancement, across the auditory cortex, when attending to either feature.

5.4.1 Feature-specific modulation in the auditory cortex

The present results do not support the claim for feature-specific attentional modulation in the auditory cortex, neither for enhancement or suppression. Within feature-sensitive regions there was a general increase in the response when the listener was engaged in an active listening task compared to passive listening, but the magnitude of the response was not influenced by the nature of the feature that was attended. This general attentional enhancement spread beyond the feature-sensitive regions and was observed in large portions of the non-primary auditory cortex. However, there was a lot of overlap of the two contrasts in PT. This perhaps can be explained by the notion that PT is a ‘computational hub’, that is involved in the analysis of sounds that are spectrally and temporally complex (Griffiths & Warren, 2002). As a consequence of the important role for PT in analyzing complex sounds it is perhaps not surprising that activity in PT is significantly modulated by attending to these sounds. An alternative explanation could be that PT plays a role in auditory working memory or perceptual decision making.

It is possible that there was no attention-related enhancement, because both features belonged to the same object (O'Craven et al., 1999). The fact that the divided attention task showed slightly better performance than the selective attention task reinforces this interpretation. On the other hand, Krumbholz et al. (2007) compared a condition whereby the same auditory object contained both features, with a condition whereby two separate objects contained each feature. The results for both showed exactly the same thing; enhancement in the motion region, but no enhancement in the pitch region. Therefore it is difficult to conclude as to the reason of the absence of attention-related enhancement in the auditory cortex. In any case, my results agree with previous behavioral and neuroimaging studies, which have shown that one cannot attend separately to one or the other feature (Mondor, Zatorre, & Terrio, 1998; Zatorre, Mondor, & Evans, 1999). Attention appears to operate at the level of the auditory object.

5.4.2 Sources of attentional modulation across the brain

Posner & Fan (2004) suggested that top-down attentional modulation of response in the sensory regions is implemented through signals sent via anatomical projections from a network of higher order regions in frontal and parietal cortices to the sensory cortices. These signals are thought to modulate response in the sensory cortices. It is important to note that the main hypothesis of the study concerned the auditory cortex. The paradigm used is not optimal for investigating the ‘control’ attentional regions across the brain. However, since the whole brain was scanned, it is informative to look at the results in this way and attempt to interpret the function of the different regions in relation to previous studies.

Response for the general enhancement contrasts for both FM and motion overlapped considerably in frontal and parietal regions. This is in accordance with the view that attending to either spatial or non-spatial features involves similar networks (Zatorre et al., 1999). One of these regions was inferior frontal gyrus bilaterally (region 3 in Figure 5.7). Note that this region was not activated for the attention-related enhancement contrasts, which indicates that this region is involved in performing the task, or in the presence of targets, since passive listening in either study did not contain targets. In fact

Corbetta and Shulman (2002) provide evidence that this region is part of the bottom-up network that is activated when attention is redirected to unattended visual targets. Therefore, another possibility is that this region was engaged in capture of attention to a target that participants were not supposed to respond to (for example, FM targets for the 'attend motion' conditions).

Premotor cortex (region 1), especially on the right hemisphere, responded as a general enhancement for both FM and Motion as well as for attention-specific enhancement for the motion (C>B). Note that premotor cortex has been activated in previous fMRI studies, both for general enhancement (Degerman et al., 2006), but also for attention-related enhancement, when attending to motion (Krumbholz et al., 2007). Furthermore, there is evidence that the posterior part of region is involved in spatial selective attention and working memory for visual stimuli (Simon et al., 2002). This study provides additional evidence that this region is involved in selectively attending to both spatial and non-spatial auditory stimuli, but more so in the former.

Prefrontal activation for the general enhancement for motion suggests that it is possibly involved in working memory for the targets, since this region appears to be the most important region for working memory for visual stimuli (Knudsen, 2007). Furthermore, lesions to the dorsolateral prefrontal cortex appear to impair working memory, as well as eliminate the ability to make judgements about what is relevant (Knudsen, 2007).

In the current study, inferior parietal lobe was activated bilaterally when attending to motion, for both general and attention-related enhancement. There is also evidence that inferior parietal lobe is central for auditory working memory for spatial location (Alain, He, & Grady, 2008). Since the task has a working memory component that cannot be separated from the selective attention component, activation in this region provides support for Alain's claims. An alternative explanation is that right IPL is part of the attentional network for top-down selective attention for visual stimuli (Corbetta et al., 2000; Hopfinger et al., 2000).

Precuneus showed an attention-related response when attending to motion. This is a region in postero-medial parietal cortex, thought to be involved in several higher order brain functions (for a review see Cavanna & Trimble, 2006). One of these functions concerns preparing to perform (Astafiev et al., 2003) or actually performing a spatial task (Kawashima, Roland, & O'Sullivan, 1995). More interestingly, there is evidence that these regions are activated when attentively tracking the spatial movement of visual objects compared with passively viewing, without actually performing any movement (Culham et al., 1998). Along with bilateral activation in the visual motion region V5/MT, precuneus could have been activated because participants used a visual imagery strategy to track the spatial location of the sounds to perform the spatial task. In fact two of the participants stated that they used this strategy. Alternatively, V5/MT activation could reflect an involvement of V5 in auditory motion processing. Previous studies have found a response in V5 when listeners had to perform an easy task while listening to auditory motion (compared to stationary sounds) (Poirier et al., 2005; Poirier et al., 2006). Additionally, a TMS study showed that, when deactivating the occipital cortex, participants were unable to perform sound lateralisation (Lewald et al., 2004). This indicates that occipital areas could be involved in processing non-visual stimuli. The results here indicate that V5 is not involved in sensory processing of auditory motion, but rather in *attention* to auditory motion.

5.4.3 Comparing controlled and passive localizers

Although not the primary focus of the current study, the experimental design permitted to explore the way in which the listening task might influence the location and extent of the FM- and motion-related sensory activation. Different studies attempted to control the 'localizer' contrasts by either directed the participants' response to visual stimuli (Petkov et al., 2004) or to an irrelevant feature of the same stimulus (Krumbholz et al., 2007), while others use passive listening (Ahveninen et al., 2006). In this study, the two latter methods were compared. The results demonstrated that when control was exerted over the listening context by requiring participants to attend to the irrelevant feature, the feature-related response was rather widespread and extended beyond the auditory cortex. For example, when participants attended to the motion targets, the effect of FM included lateral prefrontal cortex, right pre- and postcentral gyri, and left superior parietal lobe. It is interesting to that note that Krumbholz et al. (2007) had shown a response in right precuneus (for the motion) and bilateral prefrontal areas (for pitch). In contrast, when the listening context was rather unconstrained (passive listening), the feature-related response was restricted to the auditory cortex. In fact, the response for the passive listening conditions was more focal, while the 'controlled for attention' conditions showed a more extended response across the auditory cortex. The conclusion from this pattern of results is that the ‘passive listening’ contrasts are the most conservative localizers. This finding contradicts the rationale proposed by Krumbholz et al. (2007) that the ‘attend’ contrasts are better for localizing feature-specific maps and is in line with evidence that passive listening gives a good representation of sensory activation, despite not being controlled (Scott et al., 2007). The ‘controlled’ for attention mapping contrasts appear to involve additional higher processes, despite the fact that the ‘attention to the irrelevant feature’ should have effectively cancelled out (Krumbholz et al., 2007).

5.5 Conclusion

No evidence for feature-specific attentional enhancement or suppression was shown in the auditory cortex, when attending to either motion or FM of the same auditory object. Selective attention to complex stimuli, no matter if attending to spatial or non-spatial properties, seemed to widely enhance activity compared to passive baseline. The results of this study appear to give support to the notion that attention can be directed to auditory objects, but not to different features within an object.

Chapter 6: Experiment 3: Acuity of auditory spatial attention

6.1 Introduction

As discussed in Section 1.5, visual perceptual acuity deteriorates the further away from the fovea (Morrone et al., 1989). Similarly, the acuity of visual attention as reflected in performance for static visual tasks also deteriorates (Golla et al., 2004; Yeshurun & Carrasco, 1999). For example, Downing and Pinker (1985) found that the attentional tuning curve is sharper at the midline and shallower in the periphery. In their cueing experiment, they normalized the reaction time costs and benefits against a baseline reaction times, to factor out differences due to perceptual acuity. Thus, their results isolate the effect of attention.

In audition, discrimination thresholds also increase the further away from the midline (Mills, 1958). This increased acuity at the midline is tentative evidence for an auditory equivalent of the visual fovea. Discrimination data would therefore predict a gradient of spatial attention tuning for targets displaced from the centre. However visual attention studies show that attentional tuning either increases or decreases with eccentricity, depending on the task (see Section 1.5). A similar result has been demonstrated in audition. Specifically, Mondor and Zatorre (1995) required participants to initially attend to midline, far right (90º) or far left (-90º), and they were cued to shift their attention to locations of 45, 90, 135 or 180º away from the initial focus. The results showed that the attentional curves for all conditions, constructed by the time required to shift attention for all conditions, had V-shapes and were centered at the midline. However, there was no effect of eccentricity when, in other experiments, they required participants to perform a target identification task and not an explicit localization task. This provides some evidence that there is a benefit for attentional shift at the midline, at least for explicit localization tasks.

There are studies that show that attentional tuning is narrower in the centre than in the periphery both in terms of evoked brain potentials and behavioral results (Roder et al., 1999; Teder-Salejarvi, Hillyard, Roder, & Neville, 1999; Teder & Naatanen, 1994). Specifically, Teder-Salejarvi et al. (1999) used free field stimuli, presented through four loudspeakers in the midline (‘attend centre’ condition), and four in the far right (‘attend right’ condition). Within each condition, standard tones and rare targets were presented from all four loudspeakers, but participants were instructed to attend to only one loudspeaker. Responses to the other three loudspeakers were considered as false alarms. Both behavioral and ERP data showed that attentional tuning in the midline was much narrower (3-5º) than in the periphery (9-12º). In the current study, I seek to confirm this result by using only ITD as a spatial cue, as well as a different paradigm to derive the attentional tuning curves.

To investigate the tuning characteristics of auditory spatial attention, the separation in terms of ITD between the targets and distractors in a sound sequence has been manipulated. The rhythmic masking release paradigm has been used to explore the width of attentional tuning curve across the azimuthal plane. The paradigm is based on the phenomenon of rhythmic masking release in which the grouping of a stream of sounds is influenced by the spatial separation of the distractors and the sound targets (Sach & Bailey, 2004; Turgeon & Bregman, 1997). In their study, Sach and Bailey (2004) used the stream-segregation task to measure the participants' ability to ignore irrelevant information that was presented at different spatial locations. Participants had to decide which of two target rhythms was presented in each trial. The target rhythm was interleaved with distractor tones of various ITDs. Analysis of the discrimination performance enabled a spatial tuning curve around the target to be plotted, reflecting how performance improved with increased spatial distance between the target and the distractors (see Figure 6.9). The stream segregation task requires attention because the two streams are perceptually ambiguous. Listeners have to focus on the target and ignore (i.e. suppress) the distractors in order to perform the task. Thus, the resulting tuning curves can be considered as reflecting attentional tuning curves. An attentional tuning curve bears some similarity to an attentional filter, as derived from probe signal studies (see Section 1.4).

The hypothesis tested in Experiment 3 is that, the width of the attentional band will be narrower when the target and distractors are centered on the midline, wider in near lateral position, and even wider in the far lateral position. Similarly, ITD discrimination thresholds are expected to be lower at the midline, and gradually fall in the near and far lateral spatial positions. To compare the attentional tuning curves between the midline and lateral positions, the position of the target was manipulated across conditions using ITD information across the two ears. The target tones appeared in one of three positions; centre (0º), near lateral (25º) or far lateral (50º). To provide an estimate of the size of spatial receptive fields, ITD discrimination thresholds for pairs of sounds were also computed for the same set of participants using reference tones at the same three spatial positions (0º, 25º and 50º).

6.2 Methods

6.2.1 Participants

Fourteen participants (seven females, mean age= 27 years, age range 18- 45 years) took part in the experiment. They all had normal hearing (≤25 dB for 250 to 8 kHz). Three participants were members of staff and the rest were volunteers who were paid for their participation. All participants signed an informed written consent form. All participants were naïve psychophysical participants, but all discriminated the rhythms with the greatest spatial separation from the maskers on at least 10 of 12 practice trials.

6.2.2 Stimuli, tasks and apparatus

Stream Segregation task

The stimuli were created in Matlab version 6.5.1 (using in-house routines for synthesizing sounds), running on an IBM-compatible PC soundcard. The sounds were 200 Hz harmonic-complex tones, synthesized at a sampling rate of 44.1 kHz, with 10 harmonic components (1-10) and were presented at 73.5 dB SPL. Each tone was 50 ms long, including 5 ms onset and offset ramps.

Each trial lasted for 8 seconds. The stimulus consisted of two streams, one containing a target sequence, and the other containing the distractors. The target at any one time was one of the two rhythmic sequences (Figure 1.6). Although both sequences contained short (300 ms) and long (600 ms) interstimulus intervals (Figure 1.6), the pattern of these intervals was different. Participants had to decide which of the two target rhythms was presented in each trial in a two alternative forced-choice (2AFC) paradigm. The two streams (target rhythm vs distractor) were separated in space using ITD cues only.

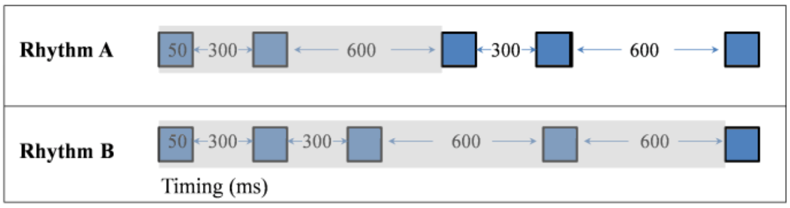


Figure 6.1 Schematic representation of rhythms A and B. The blue symbols represent each tone of the rhythm. The grey rectangular boxes represent one cycle of the stimulus. After that, the same rhythm is repeated. The numbers indicate timing in ms.

The three target location conditions were presented in three separate blocks (Figure 6.2). In one of the conditions, the target had ITD of 0 µs and was perceived in the midline. In a second condition, the target had ITD of 225 µs and was perceived at an angle of 25°, on the right side. In a third condition, the target rhythm had ITD of 430 µs and was perceived at an angle of 50° on the right side.

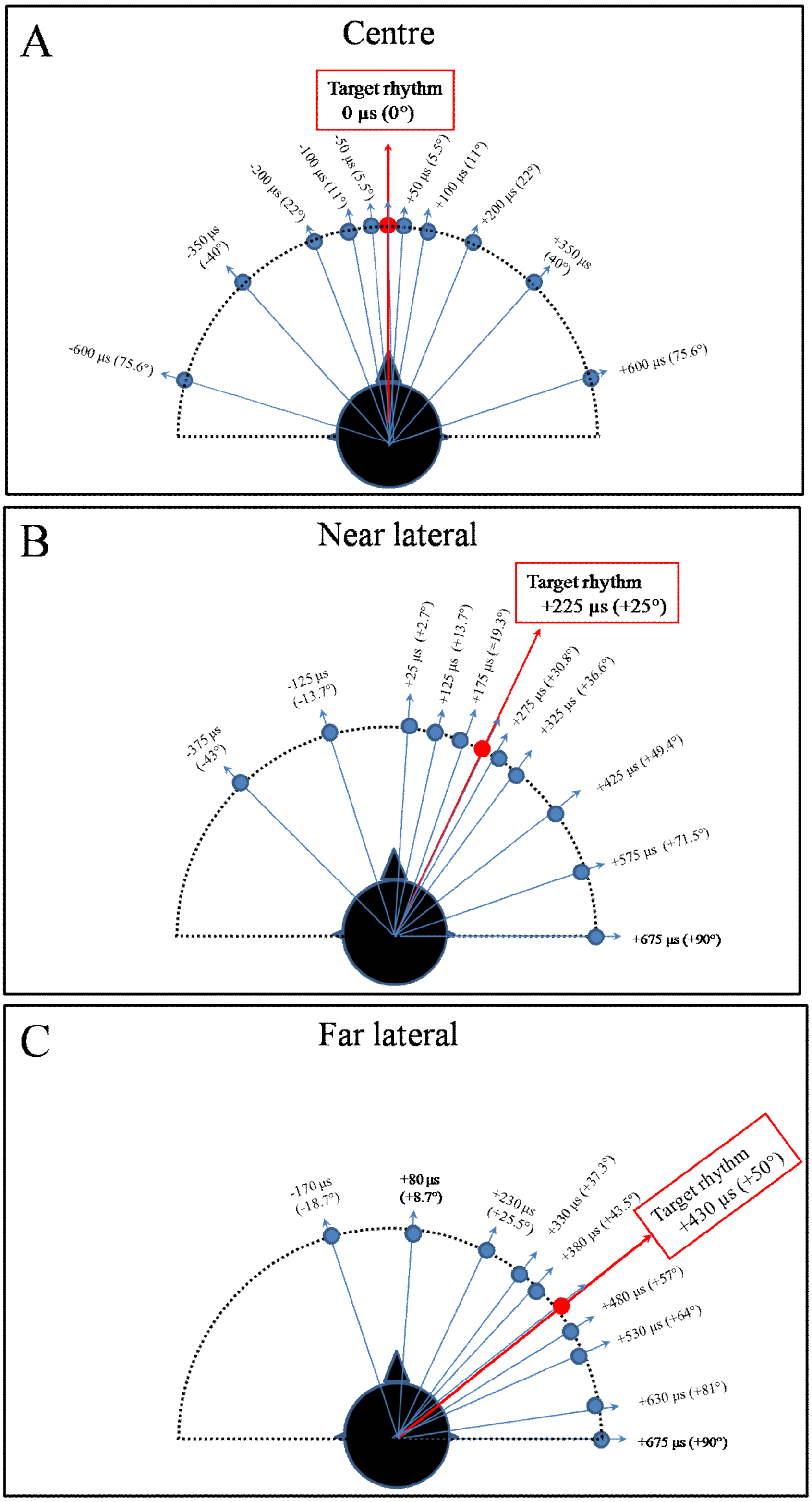


Figure 6.2 The three experimental conditions of the stream-segregation task.

As mentioned earlier, the tones of the target rhythm were interleaved with distractor tones of various ITDs (Figure 6.3). Note that within any one trial, all distractors had a fixed ITD value, i.e. not adaptive. The distractor tones were presented in the silent intervals between the tones of the target rhythm. Two distractors were inserted in the 300-ms interval, while four distractors were inserted in the 600-ms interval. The timing of the distractor tones was randomized. As Figure 6.3 shows, a minimum of 10 ms was set for separating two successive sounds. The rest of the interval was split into two or four equal parts, for the short and the long ITD respectively. For the 300-ms interval, every window was 135-ms long, and for the 600-ms interval it was 137.5-ms long. A distractor could appear within this interval at any time within these constraints. Each sequence began with a distractor so that participants would not be able to identify the target rhythm relative to the first sound they heard. The sequence continued until the participant responded, or terminated at 8 s. When the target rhythm and distractors had identical or very similar ITDs (for example difference of 5.5º or 11º between distractor and target), performance was expected to be around chance. Accuracy was expected to improve as the ITD difference between the targets and distractors became greater.

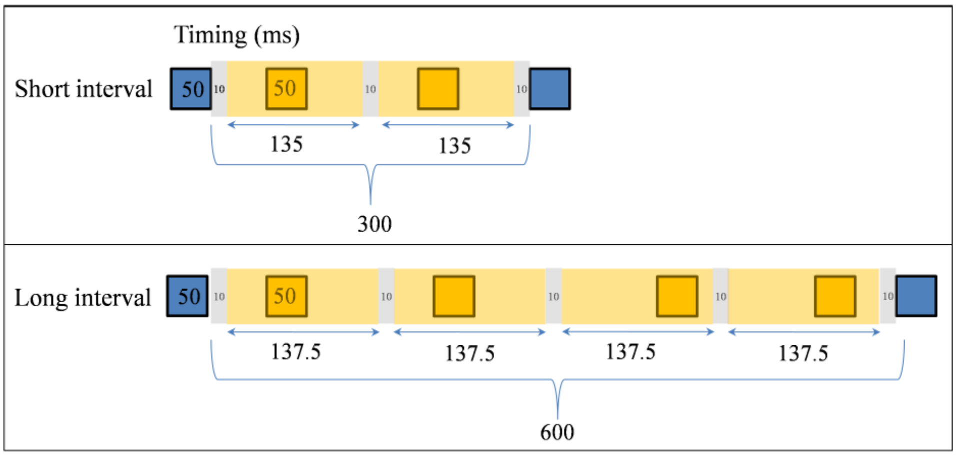


Figure 6.3 Schematic representation of the timings of the interleaved distractors in the short and the long intervals.

Spatial (ITD) discrimination task

ITD discrimination thresholds were measured with an adaptive 2AFC 3-down 1-up staircase method targeting the 79% point on the psychometric function. Stimuli were again 200-Hz harmonic-complex tones with 10 harmonic components, just like the tones in the stream-segregation task. Each sound had 100-ms duration, including 10-ms onset and offset ramps and were presented at 75.6 dB SPL. The interstimulus interval between the two tones was 200 ms. The screen showed a beach and a girl climbing a rope, indicating how far along the testing was before the end of the block. Participants heard two tones, which differed only in terms of azimuthal position (conveyed again by ITD). There were three conditions. In one of them, the standard tone was heard in the midline (ITD=0 µs), and the test tone approached from the left. In a second condition, the standard tone was heard slightly on the right side at an angle of 25º (ITD=225 µs), and the test-tone approached from the left. In a third condition, the target tone was heard at an angle of 50º (ITD=430 µs).

Each condition was tested in a separate block. The order of the three conditions and the sign of the ITD were counterbalanced across participants, similarly to the masking release task.

An example of how the ITD of the test tone varied adaptively is shown in Figure 6.4, for the central condition. In this example, only one staircase is included. The initial ITD difference between the standard and the test tone was always 500 µs. For the first trials of the staircase, there was a 1-down 1-up rule, whereby with each correct response the difference between the two tones reduced by 50 µs (5.5º) until the participant made an incorrect response (reversal point). Subsequently, the 3-down 1-up rule was followed with a step size of 25 µs for two reversals and then with a step size of 12.5 µs. The finishing rule involved the completion of 50 trials of a staircase.

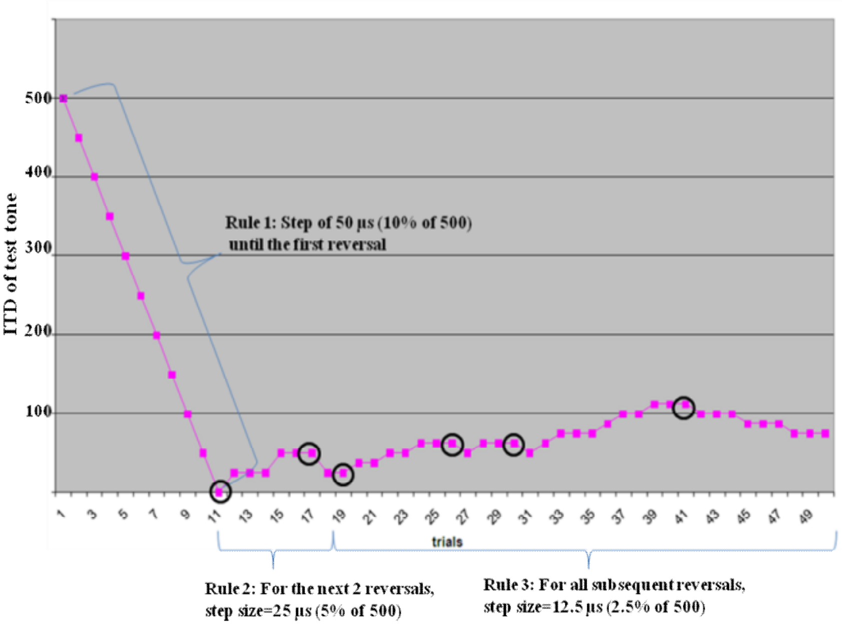


Figure 6.4 Example of a staircase procedure for participant #4. This example illustrates the step sizes and how they changed along the staircase. The black circles indicate the reversal points. This one of the three interleaved staircases for the centre condition (i.e., the test tone had ITD of 0 µs).

6.2.3 Procedure

Individual testing took place in a sound-proofed room, in two sessions. In the first session, which lasted nearly an hour, participants were tested on the stream-segregation task. In the second session, which lasted around 45 minutes, they were tested on an ITD-discrimination task to determine ITD thresholds.

For the stream-segregation task, a bespoke in-house computer program was used to present the stimuli and record the responses. Participants were seated in front of a computer screen in which the instructions were displayed. During the trials, the screen was blank. Participants were instructed to press the green button if they heard rhythm A, and the red button if they heard rhythm B. The order of the central, near lateral and far lateral conditions was counterbalanced across participants. The sign of the ITD was also counterbalanced across participants, so that half of the participants were presented with lateral targets in the left hemifield and half in the right hemifield.

Before they started each experimental block, they completed two practice blocks. The first one consisted of 12 trials, in which the ITD between target and distractor was 600 μs. This was the easiest block. If they responded correctly to 10 out of 12 trials, they proceeded to the second practice block. If not, the first practice block was repeated. The second practice block consisted of 12 trials (for the first block) and 6 trials (for the rest of the blocks). These practice trials contained both easy and difficult conditions with an ITD between target and distractor ranging from 50 to 600 μs. Participants were given feedback during the practice, but not during the experiment.

The ITD-discrimination test involved three blocks, one for each condition. Before each experiment block, participants completed a practice block, which consisted of 20 trials. Stimuli were created and presented via a bespoke in-house computer program that enables adaptive testing. Each experimental block included three interleaved staircases of 50 trials, resulting to 150 trials for each block. Participants were instructed to press the green button if the first tone was more on the left, and the yellow button if the second tone was more to the left.

For both tasks, the orientation of the buttons was perpendicular to the participant, with the green button closer to the participant, followed by the yellow and then the red. At the start of each session, participants heard a 200-Hz tone (identical to the tones used in the experiment) with ITD of 0 µs, and were asked to report where they localized it in the azimuth. This was done to make sure that they could perceive a sound with an ITD of 0 µs at the midline and that the headphones were positioned properly. Stimuli were presented via Sennheiser headphones (HD 480-13 UKII/1700R).

6.2.4 Data analysis

ITD discrimination thresholds were calculated using an average of the last four reversals for each of the three interleaved staircases of one condition. For the stream-segregation task, three different measures of tuning were computed; one on the ITD difference at 79% of the half-width attentional curve, and the one on the area of the whole tuning curves (not half-width) for the three different conditions, for both reaction times and accuracy. The first analysis served to compare the perceptual acuity of ITD measured by the staircase procedure, with the attentional acuity, measured by the stream segregation task. The half-width was taken because ITD discrimination thresholds were measured only in one side of the attentional focus, and it was not considered appropriate to assume a symmetrical result for the other side. The second and third analyses were informative in terms of the overall shape of the attentional tuning curves for accuracy and reaction times respectively, and not just the width at one point of the curve, as the first analysis.

To compute the 79% the results of the stream-segregation task were averaged across participants for each condition and were plotted as a function of ITD difference between the target and the distractors. From the resulting tuning curves of each condition, the value of ITD difference 79% was determined for each participant, and each condition, by visual inspection of the curves, which were constructed by connecting each data point with a straight line (Figure 6.6). Statistical analyses were performed, first to test whether the tuning curve was significantly narrower for the discrimination than the stream segregation tasks, and then to test each task separately for differences between the three experimental conditions.

The area of the attentional tuning curve for both accuracy and reaction times was calculated. This was done by fitting polynomials to the data, using an estimated weighted averages smoothing method for curve fitting and calculating the area of the curve for each participant. An example of the polynomial fitting is shown in Figure 6.5 (centre condition) for both reaction times and accuracy. This analysis enabled an investigation of the tuning curve width in its entirety, instead of looking only at the half-width at 79%. To measure reaction times, the time between the onset of a sequence until participant's button-press was calculated for each condition. Only the correct trials were included in the reaction times analysis. Different degrees of polynomials were used for different subjects, conditions and for the accuracy or reaction times curves. They ranged from 5th to 8th degree. The decision was made for each curve individually, by fitting polynomials of different degrees and deciding which one fitted the data best. Also note that the analysis was computed not for the whole tuning curve, but only for the 200 µs below and 200 µs above the location of the target. This was decided because it reflected the question of interest, as beyond difference of 200 µs performance was at ceiling.

While the examiners did not require re-analysis of the data they requested some clarification about other families of curves that may have given a more reasonable fit to the data than high-order polynomials which appear to overfit the data. I agree that there are different curves that could have been used that could have provided a better fit. The examiners suggested that a Gaussian curve or difference of Gaussians would have given a more appropriate. However, a Gaussian curve would always be symmetrical at the centre, so a skewed Gaussian would have been best, since the data are not always symmetrical across the centre.

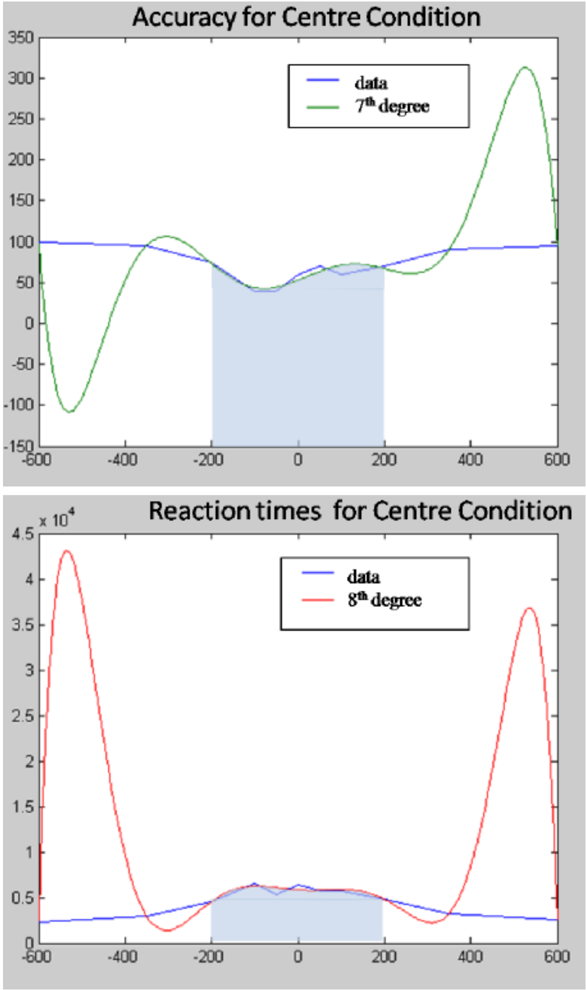


Figure 6.5 Examples of polynomial fitting for participant #7, for the condition in the stream-segregation task in which the target was in the centre. Note that the area calculated is shaded in blue.

6.3 Results

6.3.1 79% accuracy for ITD discrimination and stream segregation

Because neither ITD discrimination nor stream segregation data were normally distributed, a Friedman test was conducted. This is the non-parametric equivalent of repeated-measures ANOVA. For the same reason a Wilcoxon test was conducted for post-hoc analysis. This is the non-parametric equivalent of a paired-sample t-test. Table 6.1 shows the mean ITD difference for the 79% accuracy analysis, while Figure 6.6 shows the average results for attentional tuning curves, along with the discrimination thresholds for each condition. Two results are of interest. First the discrimination task produced much lower ITD thresholds than the stream-segregation task. Second, the near lateral condition showed better acuity than both central and far lateral conditions for both the discrimination and steam-segregation tasks.

Task	Mean (SE) / ms		
	Centre	Near Lateral	Far Lateral
Stream Segregation	171 µs,19° (17 µs, 1.9°)	148 µs, 16.3° (12 µs, 1.4°)	191 µs, 21.1° (21 µs, 2.3°)
Discrimination	89 µs, 9.8° (31 µs, 3.4°)	88 µs, 9.6° (16 µs, 1.7°)	160 µs, 17.6° (22 µs, 2.4°)

Table 6.1 showing the half-width of the tuning curve at 79% for the three experimental conditions for the stream segregation task, and the discrimination thresholds at 79%. Note that these results correspond to one hemifield.

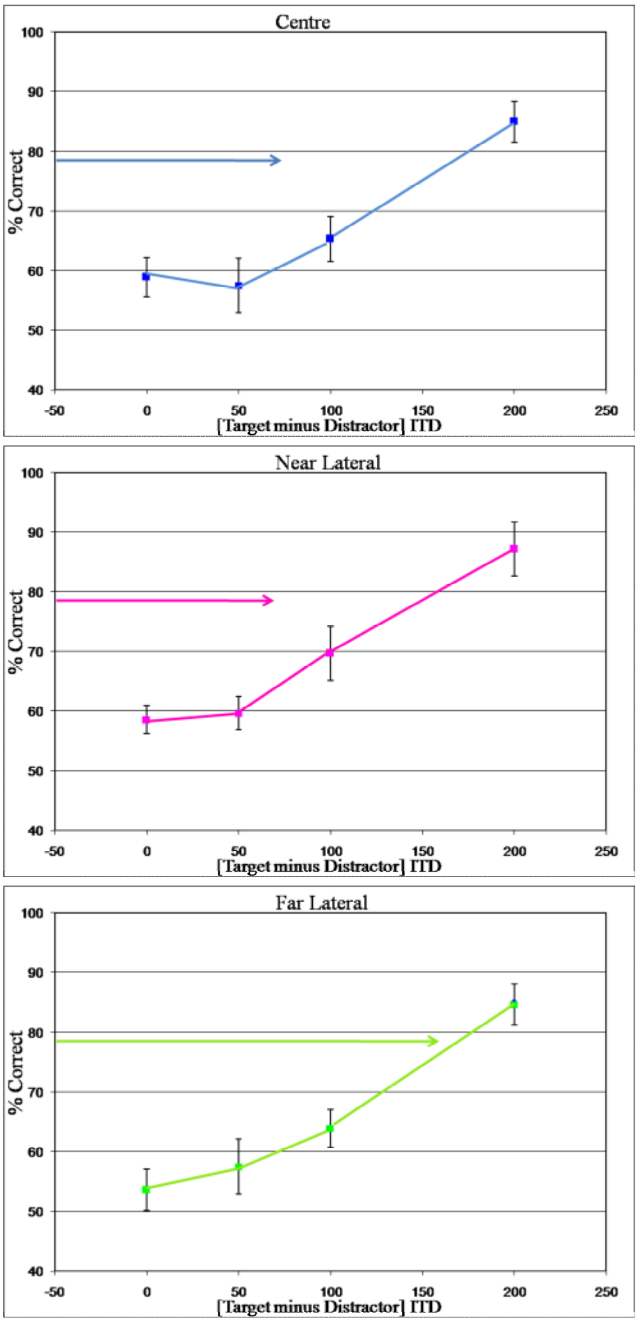


Figure 6.6 Graph showing mean accuracy in % correct, for the three conditions of the spatial stream-segregation and the ITD spatial discrimination tasks. The curves denote the results from the stream-segregation task, while the arrows denote the results for the discrimination task at 79% point of the psychometric function. The error bars indicate the standard error of the mean.

Statistical tests were performed to validate these observations. First, a Wilcoxon test showed the discrimination threshold at 79% was generally lower than the width of spatial attention at the same performance level at the same hemifield [z=-3.332, $p<0.001$]. This indicates that we can discriminate ITDs in a simple 2AFC task which are much smaller than the ITDs required to be able to discriminate rhythms.

In the spatial stream-segregation task there was no significant effect of target stream location ($p>0.05$). In the ITD discrimination task, a Friedman test revealed a significant effect of target stream location [$\chi^2=16.71$, $p<0.001$]. Wilcoxon tests on pairs of conditions showed that performance was significantly better at the centre (z=-3.045, $p<0.002$, with Bonferroni adjustment) and the near (z=-3.296, $p<0.001$) than the far lateral condition. Therefore, as predicted by the literature, perceptual acuity declined as the target moved from the centre to the side.

To explore the degree of association between perceptual and attentional performance across the group, a Pearson correlation was performed. The correlation was not significant ($p>0.05$). This analysis is illustrated in the scatter plot at Figure 6.7. This result implies that good discrimination performance was not predictive of good performance in the attentional task. Examining the results of the best participant in terms of ITD discrimination (3º, 5.2º, 8.2º), the results for the half width at 79% for stream segregation were 14.3º (130 µs), 17º (154 µs) and 15.4 º (140 µs) for the central, near lateral far lateral conditions respectively. Again, for this listener there was an effect of eccentricity on discrimination, but not for attentional tuning and even the best performance in the stream-segregation task did not reach the worst performance in the discrimination task.

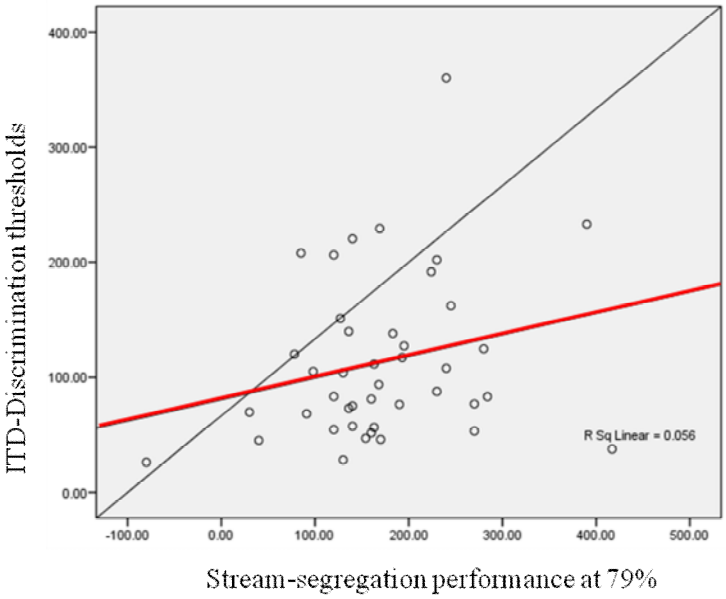


Figure 6.7 Scatter-plot showing the correlation between ITD discrimination threshold (y-axis) and stream-segregation performance at 79% (x-axis). The black line indicates the regressor line that would fit the data if they were perfectly correlated, while the red line indicates the regression line for the current data.

As noted earlier, the discrimination threshold was measured only on one side of the target. To compare the perceptual thresholds with attentional acuity, I used the data only from the corresponding side to the discrimination, i.e., half width of the attentional curve. However, taking the half width leaves the other side of the curve out thus reducing the statistical power. As I will discuss later, the attentional curve is not always symmetrical across the centre. To incorporate both sides of the attentional curve and to reduce statistical bias, the average half-width was calculated by averaging across the points of the of the two sides of the attentional curve (for example, averaging across the trials with ITD difference of 250 µs and -250 µs, when the target is at midline). The 79% point was this time determined by fitting a second order polynomial. The results show exactly the same pattern as in the half width analysis: the discrimination task had overall narrower tuning width than the average attentional half-width; there was an insignificant trend for narrower tuning for the near lateral condition than the two other conditions (especially compared with the far lateral). Finally, the correlation between attentional and perceptual data was not significant. This confirms the original analysis, even by using a different way to determine the 79% point. Nevertheless this analysis had caveats; one being that in the far lateral condition the ITD difference was not symmetrical across the two sides. For example, there was 245 ms for one side, but not for the other, and so only the value from that side was taken. The ideal would have been to have measured discrimination thresholds both sides from the target, and then compare them with the full attentional curve.

6.3.2 Area of the attentional tuning curve for the stream segregation task

The average values derived from this analysis are shown in Figure 6.8. The statistical analysis showed a significant effect of accuracy [F(2, 26)=4.00, p<0.03]. Participants performed better on the near-lateral condition than on the far-lateral condition [t(13)=-3.11, p<0.008, with Bonferroni adjustment]. Accuracy for the central position was no different from the two lateral positions. There was no significant effect of target position on the reaction-time data (possibly due to the greater inter-subject variability). There was a trend towards slower reaction times the furthest away from the midline.

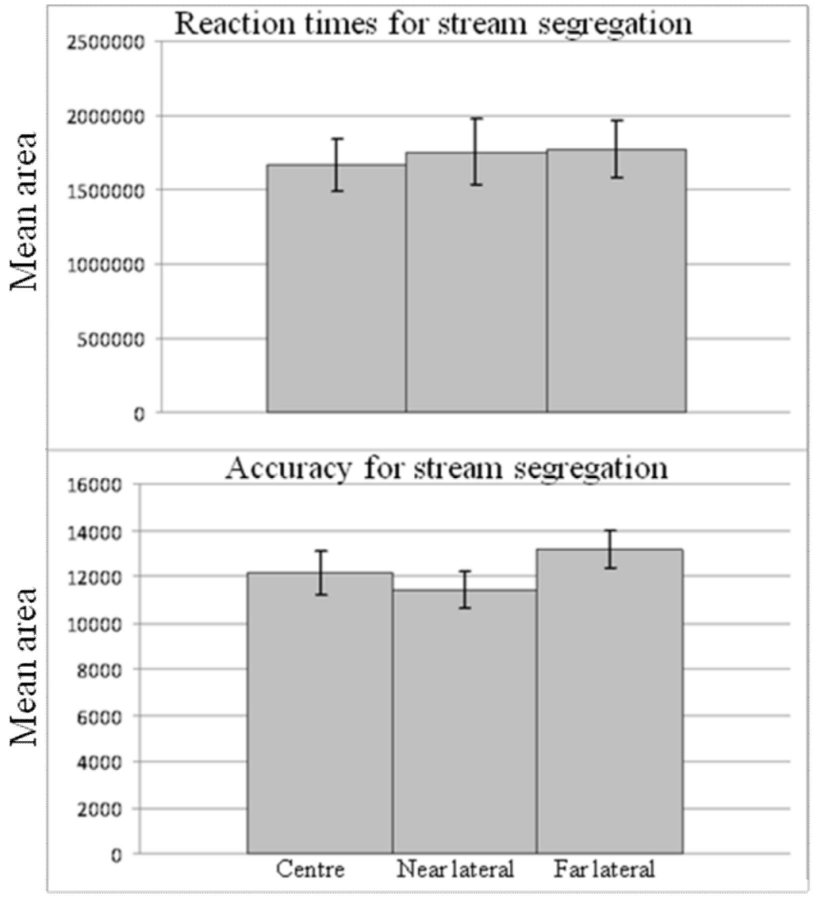


Figure 6.8 Mean area of the curve between -200 and 200 μs, for accuracy and reaction times in the stream segregation task. Errors denote standard error of the mean.

6.4 Discussion

In this study, the distribution of spatial attention in the azimuth was investigated using the masking release paradigm. The resulting attentional tuning curves and the ITD thresholds at the same locations were measured to determine whether spatial attention and perception are optimal in the midline and decline gradually as the targets depart from the midline. While the ITD discrimination thresholds showed a partial support for this hypothesis with the centre and near lateral conditions having a narrower width of tuning than the far lateral condition, the attentional data do not support this hypothesis. The width of attentional tuning was not significantly narrower at the centre than at the near and far lateral positions. Rather, there was some advantage for the near lateral condition, compared to the far lateral and (a non-significant trend) to the centre condition. This was a rather surprising result. Note that the central condition for stream segregation has the highest variability compared to the other two conditions, which might be partly responsible for the null result.

It is true that the ITD thresholds, as well as the variability in thresholds between participants, were very high compared to previous psychophysical studies (Mills, 1958). Mills (1958) found that participants could detect differences in timing up to 10 µs (1.1°). Nevertheless, participants in psychophysical studies such as the one by Mills are often trained listeners. Furthermore, as Sach and Bailey (2004) note, it is probably fairer to compare the tuning curves of the masking release paradigm with the discrimination of concurrently presented sounds (Perrott, 1984), since target rhythm and distracters were interleaved and the listening task therefore required source segregation. Using concurrent sounds stimuli, Perrott's measurements showed wider minimum audible angles (ranging from 5º to 46º) than in the study of Mills (1958). Another reason for the high ITD thresholds might be that participants in the current study were naïve listeners, while in the previous studies they were expert participants that received a lot of training and had been tested multiple times. Note, however, there was no effect of training from the first block (mean=103 µs, standard deviation (sd)= 74 µs) to the third (mean=115 µs, sd= 102 µs).

Additionally, it is possible that this experiment did not have sufficient experimental power to determine significant differences between the three conditions. Power depends on the number of participants, the size of the effect and the amount of variability in the results. The results showed a lot of variability, which was greatest for the central condition in the stream segregation task. This problem could have been addressed by testing more participants (to reduce the size of the error estimates for the mean).

When comparing the perceptual and attentional curves at 79%, the perceptual curves were significantly narrower than the attentional curves. This is in agreement with frequency discrimination thresholds being much smaller (Sek & Moore, 1995) compared to the width of the attentional tuning curve derived by the probe-signal paradigm (Greenberg & Larkin, 1968). On the other hand, as discussed in Section 1.4, the auditory filter derived from notched-noise measurements has very similar bandwidth with the one derived from the probe-signal paradigm (Dai et al., 1991). This indicates that the auditory filter can vary according to the paradigm that was used to measure it. In fact, the two tasks compared here were conceptually very different and most possibly reflecting very different aspects of ITD coding (perceptual discrimination versus stream segregation). This is partly supported by the fact that there was no correlation in performance between the two tasks. Additionally, discriminating the location of two discrete sounds is an easier task than the masking release task that involves a dynamic auditory object (and hence stream segregation). Perhaps the better performance for the former should have been expected. In fact there is considerable evidence that ITD alone is a rather weak cue for auditory streaming (Culling & Summerfield, 1995; Darwin & Hukin, 1997, 1999; Sach & Bailey, 2004). The fact that studies that presented stimuli in free field showed that tuning is narrower in the midline (Roder et al., 1999; Teder-Salejarvi et al., 1999; Teder & Naatanen, 1994) confirms this notion.

The attentional widths are quite similar to the study of Sach and Bailey (2004). In their study, the ITD of the target was always in the midline, similarly to the centre condition here. Specifically, most of the half-widths of the curves were about 160 µs (~17º); three of them were 300 µs (~33º) and one at 530 µs (64º). In the current study, the centre condition for the stream segregation task ranged from 40 to 417 µs (4º to 48º), with 10 participants scoring between 120 and 270 µs (13º to 30º). The current results appear to be much more variable than those of Sach and Bailey.

It is possible that the nature of the task does not allow very narrow tuning curves, because it is very difficult. In fact in the studies by Teder and colleagues (1999, 1994) discussed earlier, much narrower tuning was found; 3-5º in the midline, and 9-12º in the far lateral condition. This much narrower tuning might be due to the fact that they used free-field stimuli which elicit better performance on spatial attention tasks than headphone presentations (Roberts, Summerfield, & Hall, 2009). However, other differences between the two tasks might account for the difference in the tuning between the studies. For example, the task in the study by Teder & Naatanen (1994) was to attend to the tones (target) in speech (distractors), while in the current study, both targets and distractors were identical in terms of all acoustical characteristics apart from ITD. Thus presumably the task in the current study is more difficult. Nevertheless, a follow-up for Experiment 3 would be to repeat the same experiment in free field with stimuli presented through loudspeakers. It is possible that then selective attention would decline with eccentricity, just as it did in other free-field studies (Roder et al., 1999; Teder-Salejarvi et al., 1999; Teder & Naatanen, 1994).

Symmetry of tuning curves. Initially it was assumed that spatial acuity is symmetric across the centre. This is why the ITD threshold was measured only in one side of the target. However, as noted earlier, it should not have been assumed, but rather thresholds should have been measured in both sides of the target. An important question is whether the tuning curve of spatial *attention* is symmetric to the centre. A subset of the results, from 6 participants, was examined in terms of the polynomial fits. For a polynomial curve to be considered symmetrical, it was expected to have a peak at the target ITD or the distractors that are very near to the target (i.e. 50 μs), and to have a symmetrical slope in both sides. Observation of the curves showed that 2, 4 and 4 out of 6 were symmetrical for the accuracy for centre, near lateral and far lateral, and 5, 5, 3 out of 6 for the same conditions for reaction times. This result indicates variability in the symmetry of the attentional curve. Thus symmetry should not be assumed. Examining the results of Sach and Bailey (2004), for seven out of ten participants the curves were symmetrical. Both my results and their results, suggest that the spatial attentional curve is not always symmetrical across the midline, and thus one should measure both sides of the curve and not assume symmetry.

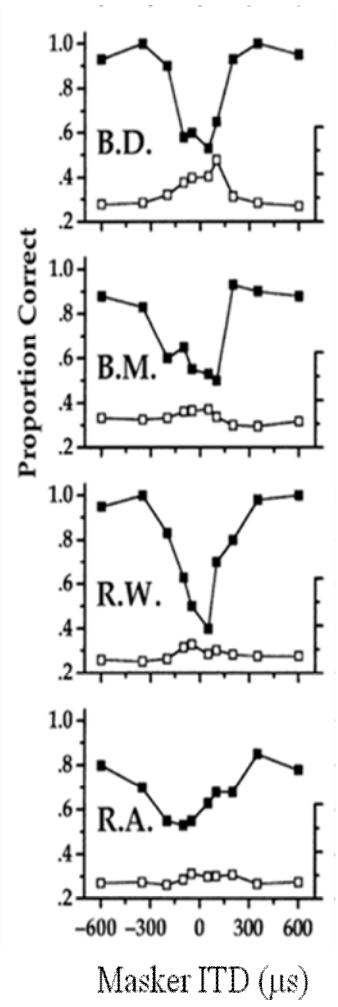


Figure 6.9 Tuning curves of 4 participants in the study by Sach and Bailey (2004).

In summary, the results of Experiment 3 indicate that while ITD discrimination performance declined from the centre, stream segregation performance did not. It was optimal at the near lateral position of 25º, and declined at the far lateral position. Therefore, auditory spatial attention is not distributed in the same manner as perceptual acuity (as measured by discrimination thresholds) and does not appear to be optimal at midline. Certainly for ITD cues, a rather weak cue for stream segregation (Darwin & Hukin, 1997), there is no gradient of spatial auditory attention.

Chapter 7: Summary of findings and discussion

The studies reported in this thesis examined a number of possible mechanisms of auditory selective attention. In this final chapter, I will summarize and discuss the most significant results and indicate future research that could follow up from these.

7.1 Research aims

The studies presented in this thesis aimed at examining the effects of auditory selective attention to sound features and auditory streams or objects defined by particular features. First, the hypotheses of feature-specific attentional enhancement and suppression as neural mechanisms of selective attention were investigated in two separate fMRI studies. Second, the hypothesis of a spatial gradient affecting the acuity of spatial attention was investigated in a behavioral study. These hypotheses were derived from visual attention research. The aim of this thesis was to establish whether these same mechanisms can be applied to auditory selective attention.

7.2 Summary of findings

7.2.1 Chapter 4: Experiment 1- An fMRI study on attending to two auditory streams or objects defined by sound frequency

The aim of this study was twofold: first, to map response for high- and low-frequency sounds in the auditory cortex. Second, to investigate whether there was an enhancement of the response in a frequency-sensitive region when attending to the best frequency of this region and a suppression when attending away from the best frequency. In this study, the attention of participants was directed to one of two concurrent auditory objects (streams) defined by their sound frequency. First, evidence supported feature-specific enhancement in the primary auditory cortex, when conducting the comparisons that showed ‘attention-specific’ enhancement. Attention-specific enhancement was tested by comparing the condition in which participants attended to the best frequency of the region, when that frequency was the majority, with the condition in which participants attended to the minority frequency (which was not the best frequency of the area). Additionally, the comparison to identify ‘general attentional enhancement’ showed widespread enhancement across the primary and non-primary auditory cortex. Although this enhancement included frequency-sensitive regions, it was not restricted to them.

There was no evidence for suppression of the unattended stream in the auditory cortex. Instead, there was general attentional suppression of response was found bilaterally in posterior part of TPO. This effect was tested by comparing the condition ‘passively listen majority stimulus’ with ‘attend minority stream of the same stimulus’.

7.2.3 Chapter 5: Experiment 2- An fMRI study on attending to FM or auditory motion of the same auditory object

Experiment 2 had the same aims as Experiment 1. However, in this paradigm the features attended were higher-order features, namely FM and auditory motion. Participants were instructed to direct attention to a single auditory object, but had to attend a different feature of that object at one time or another. The results showed no evidence for feature-specific attentional enhancement in the FM- and motion-sensitive regions.

Investigation of attentional enhancement across the brain was also conducted. The ‘attention-specific enhancement’ comparisons showed no significant response in the auditory cortex. Interestingly, there was activation in V5/MT visual motion region bilaterally for the contrast ‘attend motion>attend FM’. The ‘general enhancement’ comparisons showed a widespread response across the auditory cortex, which was not restricted to feature-specific regions of FM and motion. It is interesting to note that both Experiments 1 and 2 showed evidence for widespread enhancement across the primary and non-primary auditory cortex during attentionally demanding tasks, compared to idle passive listening. However, this type of enhancement was not confined in feature-sensitive regions.

The ROI analysis revealed no evidence for suppression of the unattended feature in feature-specific regions of the auditory cortex. However, suppression was not examined across the brain. Given more time for further analysis, it would have been interesting to investigate the effects of suppression across the brain and examine whether there is an effect of suppression in posterior TPO, as in Experiment 1.

Beyond the main questions, a secondary aim was to compare passive listening with conditions that controlled for attention, for the FM and motion localizer contrasts. The results showed that the ‘controlled for attention’ contrast showed a more widespread response in the auditory cortex than the passive contrast. Additionally, there was a greater response in frontal regions, such as premotor cortex. This indicates that passive listening is a suitable condition for using as a functional localizer, as albeit ‘uncontrolled’, it appeared to involve only the relevant sensory areas, and no higher-order areas. Attempting to control for attention appeared to involve attentional control regions differentially across the two ‘attend’ conditions as they did not simply ‘subtract’ away.

7.2.4 Chapter 6: Experiment 3- Eccentricity effect on auditory spatial attention using ITDs

Experiment 3 was a behavioral study that was conducted to test whether spatial attention to sounds defined by ITD declined as attention was directed further away from the midline. The results showed that this was not strongly the case because performance for the central condition was not the ‘best’. The only fairly consistent result was that spatial attention directed to the near-lateral condition facilitated performance more than for the far-lateral condition. Although this is partial evidence for a decline with eccentricity, it was expected that the central condition would be the condition that would show the best performance. This was the case for the ITD thresholds, which measure sensory auditory acuity for sounds that differ in ITD. In that case, performance in both central and near-lateral conditions was better than the far-lateral. Generally, the results provide very weak support for the eccentricity view of auditory spatial attention.

7.3 Discussion

7.3.1 Feature-specific attentional enhancement but not suppression in the auditory brain

The existence of both enhancement in neurons sensitive to the attended information, along with suppression of response in neurons sensitive to the unattended information (Treue & Trujillo, 1999), indicate that the attentional top-down bias activates both inhibitory and excitatory circuits. The combined effect of enhancement and suppression serves to enhance the relative responsiveness of the neurons sensitive to the attended attribute (Knudsen, 2007).

What is the functional role of feature-specific enhancement? Possibly to enhance the SNR ratio of the attended information, thus making neurons more sensitive to their preferred stimulus (Reynolds et al., 1999). Does this enhancement affect behavior? There is some evidence from primate electrophysiology that this feature-specific enhancement is related to improved behavioral performance (Boudreau et al., 2006; Chen et al., 2008; Spitzer et al., 1988). A more difficult task showed a greater response in V4 neurons as well as better performance, than a less difficult task. However, one study failed to show this relationship (Cook & Maunsell, 2002). In this thesis, the experimental paradigms were not set up to specifically examine this question. For example, in Experiment 2, great care was taken to equate performance across conditions and across participants and of course performance was not monitored in the baseline passive listening conditions.

What is the functional role of attentional suppression? Again, it is possibly to increase SNR by decreasing the 'noise', the irrelevant information. There is some evidence associating the neural suppression of unattended information with behavioral performance for the attended stimuli. For example, as discussed in Section 1.8.3, Sylvester et al. (2008) showed that there was more suppression in the region processing the unattended information when participants responded correctly to the primary task, than when they did not. Additionally, there is evidence that suppression for unattended stimuli is influenced by the perceptual load of the primary task (Rees et al., 1997). The study by Rees et al. (1997) presented in Section 1.2, showed that the higher the attentional load of the primary task, the more suppressed the activity in V5 (the area that processed the distractor). Furthermore, there is evidence that the extent to which a distractor is processed depends on the nature of the primary task. An ERP study (Sussman, Bregman, Wang, & Khan, 2005) showed that when there are three auditory streams and participants attended to a visual distractor, there was an MMN wave for distractors present in all three streams, which indicates that the unattended auditory deviants were detected. On the other hand, when attending to one of the auditory streams the deviants of the other streams showed no MMN. So, perhaps when having to direct attention to one auditory stream and ignore the others, the attentional network suppresses response for the unattended auditory streams, while it does not suppress it when ignoring all sounds and attending to the visual stimulus.

In Experiment 1, instead of finding suppression within auditory cortex, general response suppression was found in a region situated in posterior TPO bilaterally. As discussed in Section 4.4, this region might be involved in filtering out unattended information or bottom-up capture of attention from irrelevant targets. However, the design of Experiment 1 does not allow differentiating between automatic attentional processes (such as capture) and voluntary attentional processes (such as selective attention) and so this interpretation is only tentative.

There are several different ways to specify a pairwise contrast to identify feature-specific suppression. The one performed in Experiment 1 was the ‘passive listen majority> attend minority stream’. However a contrast that would control for the differential attentional demands of both conditions in the contrast would be ‘attend minority stream>attend majority stream’. For completeness, it might have been informative to have computed this contrast. However, given the time constraints it was thought wiser to just choose one comparison for suppression so that it could be fully investigated. In a way, attentional enhancement was more central to the main hypothesis examined in Experiment 1. The comparison used for suppression was initially used for the ROI analysis, and so it made sense to investigate the same when constructing the incidence maps.

7.3.2 Attending to one object in the receptive field results in attentional enhancement

The biased-competition model emphasizes that the attended and unattended objects should be within the same receptive field of a neuron (Desimone & Duncan, 1995; Luck et al., 1997). When only one object is presented in the receptive field and one outside, there is no effect of feature-specific attentional enhancement; two stimuli are needed within the receptive field, to provide competition between stimuli sufficient to require attention (Luck et al., 1997). On the other hand, there are physiological studies that show attentional enhancement when attending to the only stimulus that is presented within the receptive field (Spitzer et al., 1988; Treue & Maunsell, 1996; Treue & Trujillo, 1999). In Experiment 1, the two auditory objects were defined by sound frequency, but they were far apart in frequency so that they would not be represented in one neuron's receptive field. Therefore the results support the notion that it is possible to show attentional modulation even when one auditory object is presented within and one outside the receptive field of the auditory cortex. Experiment 2 does not show feature-specific feature specific enhancement, when a single auditory object is present. In a way, this agrees with the general statement of the biased competition model, that some kind of competition for resources is needed for attention to be necessary. On the other hand, both Experiments 1 and 2 showed evidence for general enhancement (attend>passive), widespread across the auditory cortex. This suggests that competition between objects in not a necessary requirement for general enhancement of neural coding of sensory input.

7.3.3 Attentional selection as a very dynamic and flexible function

A recent review by Gilbert & Sigman (2007) has claimed that selective attention is flexible and its neural effects are variable. The effects of attention depend a lot on the context, for example, the nature of the task, the difficulty of the task, and the stimuli used. Gilbert & Sigman (2007) do not agree with the concept of a hierarchical sensory organization, whereby in early sensory regions simple stimuli are processed and in later sensory regions more complex features are processed. According to their view, attention is the result of ‘reverberation’ between different feedback and feedforward interactions. Each cortical region is a dynamic processor that changes ‘program’ according to factors such as expectations and task requirements. The results by Fritz et al. (2005, 2003) seem to support this view, since the responses of AI neurons changed according to the task at hand. Thus, the results of the experiments presented here should not be taken as a proof that a mechanism exists or not. Rather, they should be taken as evidence that, for the specific task, there was (or there was not) evidence for the involvement of an attentional mechanism. For example, the fact that there was no evidence for suppression in auditory cortex does not mean that if a different experimental design and task was used, there would still be no evidence for it. The results reported here also lend support for a more flexible model than that proposed by the biased competition. While attentional enhancement may be greatest when there is stimulus competition, this is not a necessary requirement for demonstrating attentional enhancement.

A similar philosophical position could be taken with Experiment 3. Although it did not support the eccentricity hypothesis, the results can only account for this particular stimuli and task, as other studies that presented stimuli in free field have showed that attentional performance declines with eccentricity (Teder-Salejarvi & Hillyard, 1998; Teder-Salejarvi et al., 1999). Here space was defined by ITD, a weak cue for stream segregation. In fact, a study by Roberts et al. (2009) has shown more reliable effects of spatial orienting of attention when sounds were presented in free field than over headphones using only ITD as a spatial cue. In fact, some of these results from Robert’s spatial cueing task showed that ITD alone was not a salient cue.

7.3.4 Attending to objects versus attending to features of the same object

In Experiment 1, the high- and low-frequency sounds constituted two different objects. In Experiment 2, the two features (FM and motion) were part of the same object. The stimuli in Experiment 1 could perhaps have created a competition between targets and concurrent distracters, whereas the stimuli in Experiment 2 created no such competition as a result. In Experiment 2 participants were perhaps able to attend to both features of the object, as O’Craven et al. (1999) suggested. In fact, according to O’Craven et al. (1999), this is the central evidence for object-based attention. Scholl (2001) in his review suggests that this ‘same object advantage’ effect breaks down under conditions of high attentional load. So, perhaps, the stimuli and task in Experiment 2 did not create enough load to generate the need for selective attention to each feature. This difference in task demands was reflected in performance for ‘attend both’ than for ‘attend one object/stream’. In Experiment 1, performance was worse for ‘attend both’ than for ‘attend one object/stream’. On the contrary, for Experiment 2, performance was better for ‘attend both’ than for ‘attend one feature’. The results from Experiments 1 and 2 suggest that it is more possible to see the effects of attention when two competing objects are present than when two features belong to a single auditory object. Scholl (2001) also suggests that not all features are equal. Some features are more ‘tightly coupled’ with the object representation than others (Kubovy & Van Valkenburg, 2001). It is possible that this is the case for the features (FM and motion) in Experiment 2. Krumbholz et al. (2007) reported a perceptual asymmetry in the effects of selective attention. For example, there was a significant differential response for ‘attend motion>attend pitch’ but not for the converse contrast, possibly supporting that pitch is more tightly coupled with the object representation than was its motion.

In a review of behavioral studies of auditory object-based attention, Shinn-Cunningham (2008) has proposed an interesting and unifying framework to explain how top-down attention influences the formation of objects through grouping of certain auditory features and resolution of across object competition (Figure 7.1). She proposes an important role for selective attention in resolving across object competition. The paradigm in used in Experiment 1 created a competition across objects. The paradigm used in Experiment 2, on the other hand, did not. Rather, the different features both belonged to the same object and thus were probably grouped together. As a consequence of grouping, even unattended features would gain some benefit from selective attention because they would be grouped to the attended features. Our results suggest that across-objects competition is required for selective attention to exert strong influence on perceptual and neural responses.

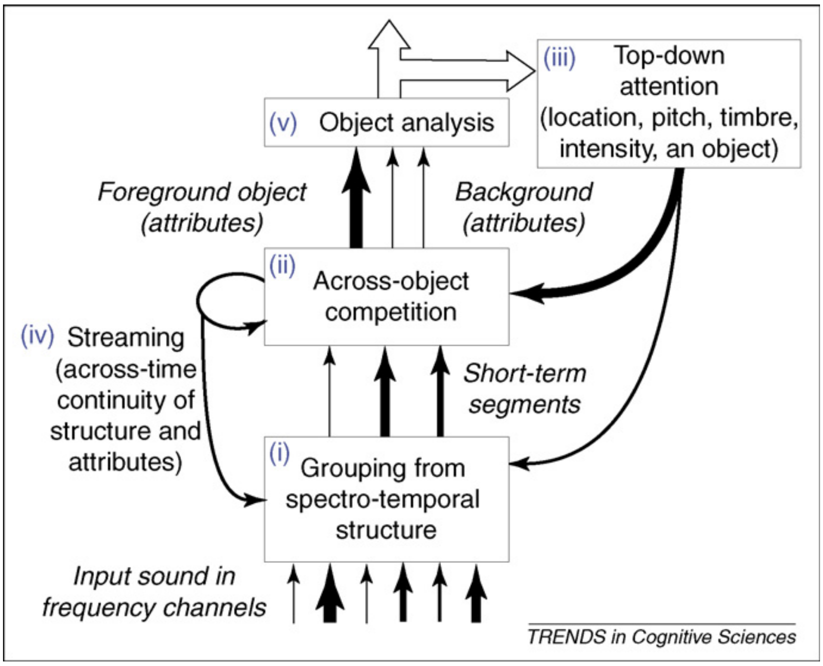


Figure 7.1 Schematic diagram showing the relationship of auditory object formation with sensory input and top-down attention. The widths of the arrows indicate the strength of a signal/connection. (Shinn-Cunningham, 2008).

7.4 Directions for further research

One of the most interesting topics for selective attention is the relationship between the neural mechanisms of selective attention (such as feature-specific attentional enhancement or suppression of neural responses) and behavioral performance. One question might be whether the size of enhancement and suppression is dependent upon the task load. Although there is some evidence from the visual literature, there are no studies explicitly studying this relationship in audition. Studies manipulating task difficulty (such as in Ress et al., 2000), or attentional load (such as in Rees et al., 1997) could provide useful paradigms for such a purpose.

Experiments 1 and 2 targeted the auditory cortex, and thus were not optimal to disentangle the function of those higher-order areas. More studies are needed that investigate the relationship between control regions and sensory regions. The functional dynamics of the attentional network could be examined using simultaneous EEG and fMRI. While fMRI could be used to localize specific control and sensory regions that are involved in performing an attentional task, EEG could examine the timing of these responses specifically in terms of their relative onsets. For example, results could show that activity in a source region preceded activity in a sensory region or that there was greater phase synchronization across source and sensory regions in the different frequency bands of the EEG signal. Both of these issues were examined in a electrophysiological study in the monkey visual system, as discussed in Section 1.7.4 (Saalmann et al., 2007).

The interesting effect of the involvement of V5 in auditory motion processing could be examined further by studies that use both visual and auditory moving stimuli, as well as by manipulating selective attention, to investigate more thoroughly to what extent, and under what circumstances V5 is involved in auditory motion processing.

7.5 Conclusions

There are several mechanisms by which selective attention can operate. Here, I have shown that feature-specific enhancement can occur in frequency-sensitive regions of the primary auditory cortex. However, there was no evidence for feature-specific enhancement in the non-primary auditory cortex for FM and motion. Experiment 2 supports the view that auditory selective attention cannot be directed to separate features of one object. This is evidence for object-based attention.

References

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated “what” and “where” pathway in human auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 14608-14613.

Alain, C., He, Y., & Grady, C. (2008). The contribution of the inferior parietal lobe to auditory spatial working memory. *Journal of Cognitive Neuroscience*, *20*(2), 285-295.

Anderson, L. A., Wallace, M. N., & Palmer, A. R. (2007). Identification of subdivisions in the medial geniculate body of the guinea pig. *Hearing Research, 228*(1-2), 156-167.

Archie, K. A., & Mel, B. W. (2000). *Dendritic compartmentalization could underlie competition and attentional biasing of simultaneous visual stimuli.* Paper presented at the NIPS.

Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *NeuroImage, 22*(1), 401-408.

Ashburner, J., & Friston, K. J. (1997). Multimodal image coregistration and partitioning - a unified framework. *NeuroImage, 6*(3), 209-217.

Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping, 7*(4), 254-266.

Astafiev, S. V., Shulman, G. L., Stanley, C. M., Snyder, A. Z., Van Essen, D. C., & Corbetta, M. (2003). Functional organization of human intraparietal and frontal cortex for attending, looking, and pointing. *Journal of Neuroscience, 23*, 4689-4899.

Barrett, D. J. K., & Hall, D. A. (2006). Response preferences for "what" and "where" in human non-primary auditory cortex. *NeuroImage, 32*(2), 968-977.

Bartlett, E. L., & Wang, X. (2007). Neural representations of temporally-modulated signals in the auditory thalamus of awake primates. *Journal of Neurophysiology, 97*, 1005-1017.

Battelli, L., Cavanagh, P., Intriligator, J., Tramo, M. J., Hénaff, M.-A., Michèl, F., et al. (2001). Unilateral Right Parietal Damage Leads to Bilateral Deficit for High-Level Motion. *Neuron, 32*(6), 985-995.

Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The 'when' pathway of the right parietal lobe. *Trends in Cognitive Sciences, 11*(5), 204-210.

Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H.-J., & Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature, 400*(6746), 724-726.

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *NeuroImage, 14*(2), 310-321.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science, 308*(5721), 529-534.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10,* 512-528.

Bisiach, E., Cornacchia, L., Sterzi, R., & Vallar, G. (1984). Disorders of perceived auditory lateralization after lesions of the right hemisphere. *Brain, 107*, 37-52.

Boudreau, C. E., Williford, T. H., & Maunsell, J. H. R. (2006). Effects of task difficulty and target likelihood in area V4 of macaque monkeys. *Journal of Neurophysiology*, *96*(5), 2377-2387.

Bourk, T. R., Mielcarz, J. P., & Norris, B. E. (1981). Tonotopic organization of the anteroventral cochlear nucleus of the cat. *Hearing Research*, *4*(3-4), 215-241.

Braddick, O. J., O’Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception*, *30*, 64-72.

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K. P., et al. (2001). Polymodal Motion Processing in Posterior Parietal and Premotor Cortex: A Human fMRI Study Strongly Implies Equivalencies between Humans and Monkeys. *Neuron*, *29*(1), 287-296.

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience, 3,* 243-249.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience, 13,* 87-100.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *Journal of Neuroscience, 25*(29), 6797-6806.

Butler, R. A., & Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *The Journal of the Acoustical Society of America, 61*(5), 1264-1269.

Carrasco, M., Giordano, A. M., & McElree, B. (2006). Attention speeds processing across eccentricity: Feature and conjunction searches. *Vision Research, 46*(13), 2028-2040.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(3), 564-583.

Chen, Y., Martinez-Conde, S., Macknik, S. L., Bereshpolova, Y., Swadlow, H. A., & Alonso, J. M. (2008). Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nature Neuroscience, 11*(8), 974-982.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25*(5), 975-979.

Clare, S., Francis, S., Morris, P. G., & Bowtell, R. (2001). Single-shot T*2 measurement to establish optimum echo time for fMRI: studies of the visual, motor, and auditory cortices at 3.0 T. *Magnetic Resonance in Medicine, 45*, 930–933.

Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., & Suetens, P. (1995). Automated multi-modality registration based on information theory. In G. Marchal, Y. Bizais, C. Barillot & R. Di Paolo (Eds.), *Proceedings of information processing in medical imaging* (pp. 263-274): Kluwer Academic.

Cook, E. P., & Maunsell, J. H. R. (2002). Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. *Journal of Neuroscience, 22*(5), 1994-2004.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron, 21*(4), 761-773.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience, 3*(3), 292-297.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science, 248*(4962), 1556-1559.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *Journal of Neuroscience, 11*(8), 2383-2402.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215.

Corchs, S., & Deco, G. (2002). Large-scale neural model for visual attention: Integration of experimental single-cell and fMRI data. *Cerebal Cortex, 12*(4), 339-348.

Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. H. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology, 80*, 2657–2670.

Culling, J. F., & Summerfield, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America, 98*, 785-797.

Dai, H., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *The Journal of the Acoustical Society of America, 89*(6), 2837-2842.

Darwin, C. J., & Hukin, R. W. (1997). Perceptual segregation of a harmonic from a vowel by interaural time difference and frequency proximity. *Journal of the Acoustical Society of America, 102*, 2316-2324.

Darwin, C. J., & Hukin, R. W. (1999). Auditory objects of attention: The role of interaural time differences. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 617-629.

Deco, G., Pollatos, O., & Zihl, J. (2002). The time course of selective visual attention: theory and experiments. *Vision Research, 42*, 2925–2945.

Degerman, A., Rinne, T., Salmi, J., Salonen, O., & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. *Brain Research, 1077*, 123-134.

Deouell, L. Y., Heller, A. S., Malach, R., D'Esposito, M., & Knight, R. T. (2007). Cerebral Responses to Change in Spatial Location of Unattended Sounds. *Neuron, 55*(6), 985-996.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193-222.

Deutsch, J. A., & Deutsch, D. (1963). Attention, some theoretical considerations. *Psychological Review, 70*, 80-90.

Donner, T., Kettermann, A., Diesch, E., Ostendorf, F., Villringer, A., & Brandt, S. A. (2000). Involvement of the human frontal eye field and multiple parietal areas in covert visual selection during conjunction search. *European Journal of Neuroscience, 12*(9), 3407-3414.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2001). The effect of task relevance on the cortical response to changes in visual and auditory stimuli: an event-related fMRI study. *Neuroimage, 14*, 1256-1267.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M. Posner & O. S. M. Marin (Eds.), *Attention and performance* (pp. 171-187). London: Erlbaum.

Driver, J., & Frith, C. D. (2000). Shifting baselines in attention research. *Nature Review Neuroscience, 1*(2), 147-148.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501-517.

Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping, 7*(2), 89-97.

Eggermont, J. J. (1991). Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hearing Research, 56*(1-2), 153-167.

Eggermont, J. J. (1994). Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hearing Research*, *74*(1-2), 51-66.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., & Amunts, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25*(4), 1325-1335.

Eramudugolla, R., Irvine, D. R. F., & Mattingley, J. B. (2007). Association between auditory and visual symptoms of unilateral spatial neglect. *Neuropsychologia, 45*(11), 2631-2637.

Ffytche, D. H., Skidmore, B. D., & Zeki, S. (1995). *Motion-from-hue activates area V5 of human visual cortex.* Paper presented at the Proceedings of Royal Society of London

Formisano, E., Kim, D. S., Di Salle, F., van de Moortele, P. F., Ugurbil, K., & Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron, 40*(4), 859-869.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science, 291*(5508), 1560-1563.

Fries, P., Womelsdorf, T., Oostenveld, R., & Desimone, R. (2008). The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *Journal of Neuroscience, 28*(18), 4823-4835.

Friston, K. J., Ashburner, J., Frith, C., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 2,* 165-189.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: Levels of inference and power. *NeuroImage, 4*(3), 223-235.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research, 229,* 186-203.

Fritz, J. B., Elhilali, M., & Shamma, S. A. (2005). Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *Journal of Neuroscience, 25*(33), 7623-7635.

Fritz, J. B., Elhilali, M., & Shamma, S. A. (2007). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of Neurophysiology, 98*(4), 2337-2346.

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

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science, 173*(3997), 652-654.

Gaese, B. H., & Ostwald, J. (1995). Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *European Journal of Neuroscience, 7*(3), 438-450.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage, 15*(4), 870-878.

Gilbert, C. D., & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron, 54*(5), 677-696.

Giraud, A. L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., et al. (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology, 84,* 1588-1598.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*(1-2), 103-138.

Golla, H., Ignashchenkova, A., Haarmeier, T., & Thier, P. (2004). Improvement of visual acuity by spatial cueing: a comparative study in human and non-human primates. *Vision Research, 44*(13), 1589-1600.

Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., & Rauschecker, J. (1997). Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport, 8*(11), 2511-2516.

Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method. *The Journal of the Acoustical Society of America, 44*(6), 1513-1523.

Griffiths, T. D., Buchel, C., Frackowiak, R. S. J., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience, 1,* 422-427.

Griffiths, T. D., & Green, G. G. R. (1999). Cortical Activation during Perception of a Rotating Wide-Field Acoustic Stimulus. *NeuroImage, 10*(1), 84-90.

Griffiths, T. D., Green, G. G. R., Rees, A., & Rees, G. (2000). Human brain areas involved in the analysis of auditory movement. *Human Brain Mapping, 9*(2), 72-80.

Griffiths, T. D., Rees, G., Rees, A., Green, G. G. R., Witton, C., Rowe, D., et al. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nature Neuroscience, 1,* 74–78.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences, 25*(7), 348-353.

Hackett, T. A. (2003). The comparative anatomy of primate auditory cortex. In A. A. Ghazanfar (Ed.), *Primate audition: Ethology and neurobiology*: CRC Press.

Haenny, P. E., Maunsell, J. H. R., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Experimental Brain Research, 69*(2), 245-259.

Hall, D. A. (2005). Sensitivity to spectral and temporal properties of sound in human non-primary auditory cortex. In L. E. Associates (Ed.), *The auditory cortex. A synthesis of human and animal research* (pp. 51-76).

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, Q., Elliott, M. R., et al. (1999). ‘‘Sparse’’ temporal sampling in auditory fMRI. *Human Brain Mapping 7*, 213–223.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Summerfield, A. Q., Palmer, A. R., Elliott, M. R., et al. (2000). Modulation and task effects in auditory processing measured using fMRI. *Human Brain Mapping, 10*(3), 107-119.

Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex, 12*, 140-149.

Hall, D. A., & Paltoglou, A. E. (2008). FMRI of the central auditory system. In M. Filippi (Ed.), *New Jersey*: Humana Press

Hall, D. A., & Paltoglou, A. E. (2009). FMRI of the central auditory system. In M. Filippi (Ed.), *New Jersey*: Humana Press

Hall, D. A., & Plack, C. J. (2008). Pitch processing sites in the human auditory brain. *Cerebral Cortex,* doi:10.1093/cercor/bhn1108.

Harms, M. P., Guinan, J. J. J., Sigalovsky, I. S., & Melcher, J. R. (2005). Short-term sound temporal envelope characteristics determine multisecond time patterns of activity in human auditory cortex as shown by fMRI. *J Neurophysiology, 93*(1), 210-222.

Harms, M. P., & Melcher, J. R. (2002). Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. *Journal of Neurophysiology, 88,* 1433-1450.

Harper, N. S., & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature, 430*(7000), 682-686.

Hart, H. C., Hall, D. A., & Palmer, A. R. (2003). 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), 104-112.

Hart, H. C., Palmer, A. R., & Hall, D. A. (2003). Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. *Cerebral Cortex, 13*(7), 773-781.

Hart, H. C., Palmer, A. R., & Hall, D. A. (2004). Different areas of human non-primary auditory cortex are activated by sounds with spatial and nonspatial properties. *Human Brain Mapping, 21*(3), 178-190.

Henkelman, R. M., & Bronskill, M. J. (1987). Artifacts in magnetic resonance imaging. In J. C. Gore (Ed.), *Reviews of magnetic resonance in medicine, volume 2* (pp. 1-127). New York: Pergamon Press.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 353*(1373), 1257-1270.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience, 3*(3), 284-291.

Howard, M. A., Volkov, I. O., Abbas, P. J., Damasio, H., Ollendieck, M. C., & Granner, M. A. (1996). A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Research, 724*, 260-264.

Huettel, S. A., Song, A. W., & McCArthy, G. (2004). *Functional magnetic resonance imaging*: Sinauer Associates, USA.

Humphreys, G. (1981). On varying the span of visual attention: evidence for two modes of spatial attention. *Quarterly Journal of Experimental Psychology, 33*(A), 17-31.

Jancke, L., Mirzazade, S., & Shah, N. J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci. Lett., 266*, 125-128.

Jiang, H., Lepore, F., Poirier, P., & Guillemot, J.-P. (2000). Responses of cells to stationary and moving sound stimuli in the anterior ectosylvian cortex of cats. *Hearing Research, 139*(1-2), 69-85.

Johnson, J. A., & Zatorre, R. J. (2005). Attention to simultaneous unrelated auditory and visual events: behavioural and neural correlates. *Cerebral Cortex, 15*, 1609-1620.

Johnson, J. A., & Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *NeuroImage, 31*, 1673-1681.

Johnsrude, I. S., Giraud, A.-L., & Frackowiak, R. S. J. (2002). Functional imaging of the auditory system: The use of positron emission tomography. *Audiology & Neuro-Otology, 7*, 251-276.

Joris, P. X. (1996). Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *Journal of Neurophysiology, 76*(4), 2137-2156.

Joris, P. X., & Yin, T. C. T. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America, 91*(1), 215-232.

Joris, P. X., & Yin, T. C. T. (1995). Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *Journal of Neurophysiology, 73*(3), 1043-1062.

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences, 97*(22), 11793-11799.

Kaas, J. H., Hackett, T. A., & Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Current Opinion in Neurobiology, 9*, 164-170.

Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology, 16*(11), 1096-1102.

Kandel, E. R., Schwartz, J. H., & Jessel, T. M. (1991). *Principles of neural science* (3rd ed.). New York: Elsevier Science Publications Co.

Karnath, H.-O. (2001). New insights into the functions of the superior temporal cortex. *Nature Reviews Neuroscience, 2*(8), 568-576.

Karnath, H.-O., Ferber, S., & Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature, 411*(6840), 950-953.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). *Increased activity in human visual cortex during directed attention in the absence of visual stimulation*.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience, 23*(1), 315-341.

Kauramäki, J., Jääskeläinen, I. P., & Sams, M. (2007). Selective attention increases both gain and feature selectivity of the human auditory cortex. *PLoS ONE, 2*(9), e909.

Kawashima, R., Roland, P. E., & O’Sullivan, B. T. (1995). Functional anatomy of reaching and visuomotor learning: a positron emission tomography study. *Cerebral Cortex, 5* 111–122.

Keilholz, S. D., Silva, A. C., Raman, M., Merkle, H., & Koretsky, A. P. (2004). Functional MRI of the rodent somatosensory pathway using multislice echo planar imaging. *Magnetic Resonance in Medicine, 52*(1), 89-99.

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience, 30*(1), 57-78.

Kosaki, H., Hashikawa, T., He, J., & Jones, E. G. (1997). Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *J. Comp. Neurol., 386,* 304-316.

Krumbholz, K., Eickhoff, S. B., & Fink, G. R. (2007). Feature- and object-based attentional modulation in the human auditory 'where' pathway. *Journal of Cognitive Neuroscience, 19*(10), 1721-1733.

Krumbholz, K., Schonwiesner, M., Rubsamen, R., Zilles, K., Fink, G. R., & von Cramon, D. Y. (2005). Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *European Journal of Neuroscience*, *21*(1), 230-238.

Krumbholz, K., Schönwiesner, M., Von Cramon, D. Y., Rübsamen, R., Shah, N. J., Zilles, K., et al. (2005). Representation of interaural temporal information from left and right auditory space in the human planum temporale and Inferior parietal lobe. *Cerebral Cortex, 15*, 317--324.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition, 80*(1-2), 97-126.

Langers, D. R. M., Backes, W. H., & van Dijk, P. (2007). Representation of lateralization and tonotopy in primary versus secondary human auditory cortex. *NeuroImage, 34,* 264–273.

Langner, G., & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat I. Neuronal mechanisms. *Journal of Neurophysiology*, 1799-1822.

Lauter, J. L., Herscovitch, P., Formby, C., & Raichle, M. E. (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hearing Research, 20*(3), 199-205.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology. Human Perception and Performance*, *21*, 451-468.

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75 - 82.

Lavie, N., & Cox, S. (1997). On the efficiency of attentional selection: Efficient visual search results in inefficient rejection of distraction. *Psychological Science, 8,* 395–398.

Levi, D. M., McGraw, P. V., & Klein, S. A. (2000). Vernier and contrast discrimination in central and peripheral vision. *Vision Research, 40*(8), 973-988.

Lewald, J., Foltys, H., & Töpper, R. (2002). Role of the posterior parietal cortex in spatial hearing. *The Journal of Neuroscience, 22*(RC207).

Lewald, J., Meister, I. G., Weidemann, J., & Töpper, R. (2004). Involvement of the Superior Temporal Cortex and the Occipital Cortex in Spatial Hearing: Evidence from Repetitive Transcranial Magnetic Stimulation. *Journal of Cognitive Neuroscience, 16*(5), 828-838.

Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A Comparison of Visual and Auditory Motion Processing in Human Cerebral Cortex. *Cereb. Cortex, 10*(9), 873-888.

Liang, L., Lu, T., & Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the auditory cortex of awake primates. *Journal of Neurophysiology, 87*, 2237-2261.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92*(3), 204-214.

Liégeois-Chauvel, C., Musolino, A., & Chauvel, P. (1991). Localization of the primary auditory area in man. *Brain, 114*, 139–153.

Liu, L.-F., Palmer, A. R., & Wallace, M. N. (2006). Phase-locked responses to pure tones in the inferior colliculus. *Journal of Neurophysiology, 95*(3), 1926-1935.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology, 66*, 735-769.

Lomber, S. G., & Malhotra, S. (2008). Double dissociation of 'what' and 'where' processing in auditory cortex. *Nature Neuroscience, 11*(5), 609-616.

Lu, T., Liang, L., & Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neuroscience, 4*(11), 1131-1138.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology, 77*(1), 24-42.

Lütkenhöner, B., Krumbholz, K., & Seither-Preisler, A. (2003). Studies of tonotopy based on wave N100 of the auditory evoked field are problematic. *NeuroImage, 19*(3), 935-949.

Malmierca, M. S., Izquierdo, M. A., Cristaudo, S., Hernandez, O., Perez-Gonzalez, D., Covey, E., et al. (2008). A discontinuous tonotopic organization in the inferior colliculus of the rat. *Journal of Neuroscience, 28*(18), 4767-4776.

Malone, B. J., Scott, B. H., & Semple, M. N. (2002). Context-dependent adaptive coding of interaural phase disparity in the auditory cortex of awake macaques. *Journal of Neuroscience, 22*(11), 4625-4638.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49*(5), 1127-1147.

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience, 19*(1), 431-441.

McAlpine, D. (2005). Creating a sense of auditory space. *Journal of Physiology, 566*(1), 21-28.

McKee, S. P., & Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Research, 24*(1), 25-32.

McMains, S. A., Fehd, H. M., Emmanouil, T.-A., & Kastner, S. (2007). Mechanisms of feature- and space-based attention: Response modulation and baseline increases. *Journal of Neurophysiology, 98*(4), 2110-2121.

Merzenich, M. M., & Brugge, J. F. (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Research, 50*, 275-296.

Middlebrooks, J. C., Harrington, I. A., Macpherson, E. A., & Stecker, G. C. (2008). Sound localization and the auditory cortex. In *The Senses: A Comprehensive Reference* (pp. 781-805). New York: Academic Press.

Mills, A. W. (1958). On the Minimum Audible Angle. *Journal of the Acoustical Society of America, 30*(4), 237-246.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences, 6*, 414-417.

Mishra, J., Fellous, J. M., & Sejnowski, T. J. (2006). Selective attention through phase relationship of excitatory and inhibitory input synchrony in a model cortical neuron. *Neural Networks, 19*(9), 1329-1346.

Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. . *Journal of Experimental Psychology: Human Perception & Performance, 21*, 387-409.

Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception & Performance, 24*(1), 66-79.

Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). San Diego: Academic Press.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science 229* 782-784.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *NeuroImage, 13*, 684–701.

Morrone, M. C., Burr, D. C., & Spinelli, D. (1989). Discrimination of spatial phase in central and peripheral vision. *Vision Research, 29*(4), 433-445.

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology, 70*(3), 909-919.

Newman, S. D., Twieg, D. B., & Carpenter, P. A. (2001). Baseline conditions and subtractive logic in neuroimaging. *Human Brain Mapping, 14*(4), 228-235.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401*(6753), 584-587.

Okamoto, H., Stracke, H., Wolters, C. H., Schmael, F., & Pantev, C. (2007). Attention improves population-level frequency tuning in human auditory cortex. *Journal of Neuroscience, 27*(39), 10383-10390.

Orban, G. A., Dupont, P., De Bruyn, B., Vogels, R., Vandenberghe, R., & Mortelmans, L. (1995). A motion area in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*(4), 993-997.

Palmer, A. R., Bullock, D. C., & Chambers, J. D. (1998). A high-output, high quality sound system for use in auditory fMRI. *NeuroImage, 7*, S359.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., et al. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology, 94*(1), 26-40.

Pantev, C., Eulitz, C., Elbert, T., & Hoke, M. (1994). The auditory evoked sustained field: origin and frequency dependence. *Electroencephalography and Clinical Neurophysiology, 90*(82-90).

Pantev, C., Hoke, M., Lehnertz, K., Lutkenhoner, B., Anogianakis, G., & Wittkowski, W. (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalography and Clinical Neurophysiology, 69*, 160-170.

Pantev, C., Roberts, L. E., Elbert, T., Ross, B., & Wienbruch, C. (1996). Tonotopic organization of the sources of human auditory steady state responses. *Hearing Research, 101,* 62–74.

Patterson, R. D., & Moore, B. C. J. (1986). Auditory filters and excitation patterns as representations of frequency resolution. In B. C. J. Moore (Ed.), *Frequency selectivity in hearing*: Academic, New York.

Pavani, F., Husain, M., Ládavas, E., & Driver, J. (2004). Auditory Deficits in Visuospatial Neglect Patients. *Cortex, 40*(2), 347-365.

Pavani, F., Macaluso, E., Warren, J. D., Driver, J., & Griffiths, T. D. (2002). A Common Cortical Substrate Activated by Horizontal and Vertical Sound Movement in the Human Brain. *Current Biology, 12*(18), 1584-1590.

Peck, K. K., Wierenga, C. E., Moore, A. B., Maher, L. M., Gopinath, K., Gaiefsky, M., et al. (2004). Comparison of baseline conditions to investigate syntactic production using functional magnetic resonance imaging. *NeuroImage, 23*(1), 104-110.

Penhune, V. B., Zatorre, R. J., Macdonald, J. D., & Evans, A., C. (1996). Interhemispheric anatomical differences in human primary auditorycortex:probabilistic mapping and volume measurement from magnetic-resonance scans. . *Cerebral Cortex 6*, 661–672.

Perrott, D. R. (1984). Concurrent minimum audible angle: A reexamination of the concept of auditory spatial acuity. *Journal of the Acoustical Society of America, 75*, 1201-1206.

Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience, 7*(6), 658-663.

Petkov, C. I., Kayser, C., Augath, M., & Logothetis, N. K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. *Publich Library of Science Biology, 4*(7), 213-226.

Phillips, D. P., Semple, M. N., Calford, M. B., & Kitzes, L. M. (1994). Level dependent representation of stimulus frequency in the cat's primary auditory cortex. *Experimental Brain Research, 102*, 210-226.

Pickles, J. O. (2008). *An introduction to the physiology of hearing* (3rd ed.). Bingley: Emerald Group Publishing Limited.

Plack, C. J. (2005). *The sense of hearing*: Lawrence Erlbaum Associates.

Poirier, C., Collignon, O., DeVolder, A. G., Renier, L., Vanlierde, A., Tranduy, D., et al. (2005). Specific activation of the V5 brain area by auditory motion processing: An fMRI study. *Cognitive Brain Research, 25*(3), 650-658.

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., et al. (2006). Auditory motion perception activates visual motion areas in early blind subjects. *NeuroImage, 31*(1), 279-285.

Poirier, P., Jiang, H., Lepore, F., & Guillemot, J.-P. (1997). Positional, directional and speed selectivities in the primary auditory cortex of the cat. *Hearing Research, 113*(1-2), 1-13.

Posner, M. I. (1978). *Chronometric explanations of mind*. Hillsdale: NJ: Lawrence Erlbaum Associates Inc. .

Posner, M. I., & Fan, J. (2004). *Attention as an Organ System*. Cambridge UK: Cambridge University Press.

Posner, M. I., & Gilbert, C. D. (1999). Attention and primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 96*(6), 2585-2587.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*(2), 160-174.

Price, C. J., & Friston, K. J. (1997). Cognitive conjunction: a new approach to brain activation experiments. *NeuroImage, 5*, 261–270.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science, 268*(5207), 111-114.

Recanzone, G. H., Guard, D. C., Phan, M. L., & Su, T.-I. K. (2000). Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *Journal of Neurophysiology, 83*(5), 2723-2739.

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science, 278*(5343), 1616-1619.

Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience, 3*, 940-945.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience, 19*(5), 1736-1753.

Rhode, W. S., & Greenberg, S. (1994). Encoding of amplitude modulation in the cochlear nucleus of the cat. *Journal of Neurophysiology, 71*(5), 1797-1825.

Rivier, F., & Clarke, S. (1997). Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage, 6*, 288-304.

Roberts, K. L., Summerfield, A. Q., & Hall, D. A. (2009). Covert auditory spatial orienting: An evaluation of the spatial relevance hypothesis. *Journal of Experimental Psychology-Human Perception and Performance*.

Roder, B., Teder-Salejarvi, W., Sterr, A., Rosler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature, 400*(6740), 162-166.

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature, 395*(6700), 376-381.

Saalmann, Y. B., Pigarev, I. N., & Vidyasagar, T. R. (2007). Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science, 316*(5831), 1612-1615.

Sach, A. J., & Bailey, P. J. (2004). Some characteristics of auditory spatial attention revealed using rhythmic masking release. *Perception & Psychophysics, 66*(8), 1379-1387.

Sàenz, M., Buracas, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43*(6), 629-637.

Samuelsson, H., Jensen, C., Ekholm, S., Naver, H., & Blomstrand, C. (1997). Anatomical and neurological correlates of acute and chronic visuospatial neglect following right hemisphere stroke. *Cortex 33*, 271-285.

Schlauch, R. S., & Hafter, E. R. (1991). Listening bandwidths and frequency uncertainty in pure-tone signal detection. *The Journal of the Acoustical Society of America, 90*(3), 1332-1339.

Schoenfeld, M. A., Hopf, J. M., Martinez, A., Mai, H. M., Sattler, C., Gasde, A., et al. (2007). Spatio-temporal analysis of feature-based attention. *Cerebral Cortex, 17*(10), 2468-2477.

Schönwiesner, M., Von Cramon, D. Y., & Rubsamen, R. (2002). Is it tonotopy after all? *Neuroimage, 17*, 1144-1161.

Scott, B. H., Malone, B. J., & Semple, M. N. (2007). Effect of behavioral context on representation of a spatial cue in core auditory cortex of awake macaques. *Journal of Neuroscience, 27*(24), 6489-6499.

Sek, A., & Moore, B. C. J. (1995). Frequency discrimination as a function of frequency, measured in several ways. *The Journal of the Acoustical Society of America, 97*(4), 2479-2486.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science, 16*(2), 114-122.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences, 12*, 182-186.

Shmuel, A., Augath, M., Oeltermann, A., & Logothetis, N. K. (2006). Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature Neuroscience, 9*(4), 569-577.

Shulman, G. L., Astafiev, S. V., & Corbetta, M. (2004). Two cortical systems for the selection of visual stimuli. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 114-126). New York: Guilford Press.

Shulman, G. L., Astafiev, S. V., McAvoy, M. P., d'Avossa, G., & Corbetta, M. (2007). Right TPJ deactivation during visual search: functional significance and support for a filter hypothesis. *Cerebral Cortex, 17*(11), 2625-2633.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks .2. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience, 9*(5), 648-663.

Shulman, G. L., McAvoy, M. P., Cowan, M. C., Astafiev, S. V., Tansy, A. P., d'Avossa, G., et al. (2003). Quantitative analysis of attention and detection signals during visual search. *Journal of Neurophysiology, 90*(5), 3384-3397.

Shulman, G. L., Ollinger, J. M., Akbudak, E., T.E., C., Snyder, A. Z., Petersen, S. E., et al. (1999). Areas involved in encoding and applying directional expectations to moving objects. *Journal of Neuroscience, 19*, 9480–9496.

Simon, S. R., Meunier, M., Piettre, L., Berardi, A. M., Segebarth, C. M., & Boussaoud, D. (2002). Spatial attention and memory versus motor preparation: Premotor cortex involvement as revealed by fMRI. *Journal of Neurophysiology, 88*(4), 2047-2057.

Slotnick, S. D., Schwarzbach, J., & Yantis, S. (2003). Attentional inhibition of visual processing in human striate and extrastriate cortex. *NeuroImage, 19*(4), 1602-1611.

Smith, A. T., Cotillon-Williams, N. M., & Williams, A. L. (2006). Attentional modulation in the human visual cortex: the timecourse of the BOLD response and its implications. *NeuroImage 29*(1), 328-334.

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. H. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 96*(4), 1663-1668.

Soroker, N., Calamaro, N., Glicksohn, J., & Myslobodsky, M. S. (1997). Auditory inattention in right-hemisphere-damaged patients with and without visual neglect. *Neuropsychologia, 35*(3), 249-256.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science, 240*(4850), 338-340.

Stecker, G. C., Harrington, I. A., Macpherson, E. A., & Middlebrooks, J. C. (2005). Spatial sensitivity in the dorsal zone (area DZ) of cat auditory cortex. *Journal of Neurophysiology, 94,* 1267-1280.

Stecker, G. C., & Middlebrooks, J. C. (2003). Distributed coding of sound locations in the auditory cortex. *Biological Cybernetics, 89,* 341-349.

Studholme, C., Hill, D., & Hawkes, D. (1998). *A normalized entropy measure of 3-D medical image alighment.* Paper presented at the Proceedings of Medical Imaging, San Diego, CA.

Stufflebeam, S. M., Poeppel, D., Rowley, H. A., & Roberts, T. P. L. (1998). Peri-threshold encoding of stimulus frequency and intensity in the M100 latency. *NeuroReport, 9*, 91-94.

Sussman, E. S., Bregman, A. S., Wang, W. J., & Khan, F. J. (2005). Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. *Cognitive, Affective, & Behavioral Neuroscience, 5*(1), 93-110.

Sylvester, C. M., Jack, A. I., Corbetta, M., & Shulman, G. L. (2008). Anticipatory suppression of nonattended locations in visual cortex marks target location and predicts perception. *Journal of Neuroscience, 28*(26), 6549-6556.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.

Talavage, T. M., Ledden, P. J., Benson, R. R., Rosen, B. R., & Melcher, J. R. (2000). Frequency-dependent responses exhibited by multiple regions in human auditory cortex. *Hear. Res., 150*, 225-244.

Talavage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neurophysiology, 91*, 1282-1296.

Teder-Salejarvi, W. A., & Hillyard, S. A. (1998). The gradient of spatial auditory attention in free-field. An event-related potential (ERP) study. *Perception and Psychophysics, 60*(7), 1228-1242.

Teder-Salejarvi, W. A., Hillyard, S. A., Roder, B., & Neville, H. J. (1999). Spatial attention to central and peripheral auditory stimuli as indexed by event-related potentials. *Cognitive Brain Research*, *8*, 213-227.

Teder, W., & Naatanen, R. (1994). Event-related potentials demonstrate a narrow focus of auditory spatial attention. *NeuroReport, 5,* 709-711.

Thivard, L., Belin, P., Zilbovicius, M., Poline, J. B., & Samson, Y. (2000). A cortical region sensitive to auditory spectral motion. *Neuroreport, 11*(13), 2969-2972.

Tootell, R. B. H., & Hadjikhani, N. (2000). Attention -brains at work! *Nature Neuroscience, 3*(3), 206-208.

Tootell, R. B. H., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron, 21*(6), 1409-1422.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12,* 242-248.

Treue, S., & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature, 382*(6591), 539-541.

Treue, S., & Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*(6736), 575-579.

Tsuchitani, C. (1977). Functional organization of lateral cell groups of cat superior olivary complex. *J Neurophysiol, 40*(2), 296-318.

Turgeon, M., & Bregman, A. S. (1997). Rhythmic masking release: A paradigm to investigate auditory grouping resulting from the integration of time-varying intensity levels across frequency and across ears. *Journal of the Acoustical Society of America, 102(Suppl. 1),* 3160.

Turner, R., Jezzard, P., Wen, H., Kwong, K. K., Le Bihan, D., Zeffiro, T., et al. (1993). Functional mapping of the human visual cortex at 4 and 1.5 tesla using deoxygenation contrast EPI. *Magnetic Resonance in Medicine, 29*(2), 277-279.

Vallar, G., & Perani, D. (1986). The anatomy of unilateral neglect after right-hemisphere stroke lesions. A clinical/CT-scan correlation study in man. *Neuropsychologia, 24*(5), 609-622.

Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.-F., & Pernier, J. (1995). Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis. *Electroencephalography and Clincal Neurophysiology, 96,* 143-156.

Voisin, J., Bidet-Caulet, A., Bertrand, O., & Fonlupt, P. (2006). Listening in silence activates auditory areas: A functional magnetic resonance imaging study. *Journal of Neuroscience, 26*(1), 273-278.

von Békésy, G. (1960). *Experiments in hearing*. New York: McGraw-Hill.

Vuilleumier, P., & Driver, J. (2007). Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society, 362*, 837-855.

Wallace, M. N., Johnston, P. W., & Palmer, A. R. (2002). Histochemical identification of cortical areas in the auditory region of the human brain *Experimental Brain Research, 143*, 499-508.

Wang, X., Lu, T., Snider, R. K., & Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. *Nature, 435*(7040), 341-346.

Wang, X., & Sachs, M. B. (1995). Transformation of temporal discharge patterns in a VCN stellate cell model: Implications for physiological mechanisms. *Journal of Neurophysiology, 73*, 1600-1616.

Wardak, C., Ibos, G., Duhamel, J.-R., & Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *Journal of Neuroscience, 26*(16), 4228-4235.

Warnking, J., Dojat, M., Guérin-Dugué, A., Delon-Martin, C., Olympieff, S., Richard, N., et al. (2002). fMRI Retinotopic Mapping--Step by Step. *NeuroImage, 17*(4), 1665-1683.

Warren, J. D., Zielinski, B. A., Green, G. G. R., Rauschecker, J. P., & Griffiths, T. D. (2002). Perception of sound-source motion by the human brain. *Neuron, 34*(1), 139-148.

Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *The Journal of the Acoustical Society of America, 61*(1), 178-184.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America, 90*(18), 8722-8726.

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature Neuroscience, 9*(9), 1156-1160.

Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *Journal of Neuroscience, 28*(36), 8934-8944.

Woodruff, W., Benson, R. R., Bandettini, P. A., Kwong, K. K., Howard, R. J., Talavage, T., et al. (1996). Modulation of auditory and visual cortex by selective attention is modality-dependent. *Neuroreport, 7,* 1909-1913.

Woods, T. M., Lopez, S. E., Long, J. H., Rahman, J. E., & Recanzone, G. H. (2006). Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *Journal of Neurophysiology, 96*, 3323-3337.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research, 39*(2), 293-306.

Yin, T. C. T., & Chan, J. K. (1990). Interaural time sensitivity in medial superior olive of cat. *Journal of Neurophysiology, 64*, 465-488.

Zatorre, R. J., Mondor, T. A., & Evans, A. C. (1999). Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage, 10*(5), 544-554.