BEHAVIOURAL AND NEURAL CORRELATES OF AUDITORY ATTENTION

Katherine Leonie Roberts, BA

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

September 2005

reflect the relative importance of spatial information in the two modalities, and differences in the neural coding of auditory and visual spatial information.

Acknowledgements

I am greatly indebted to my supervisors, Debbie Hall and Quentin Summerfield, who have been amazingly generous with their time and expertise, and have provided much-needed support and encouragement. I would also like to thank the computing and workshop staff at IHR for their impressive level of support. Many thanks to everyone else at IHR for making my time here so enjoyable, and especially to Paula for being an excellent office mate, both real and virtual.

Grateful thanks to Neil, who kept me sane and happy. For making endless cups of tea, patiently listening to my talks, and being great fun to live with.

Dedicated to my mum, who gave me her love of learning and enthusiasm for life, and whose strength and grace I can only aspire to. Who I miss every day of my life, and without whom the world is a poorer place.

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 analysed by the candidate, and all testing was conducted by the candidate, with the exception of Experiment 7 (Chapter 5), in which testing was conducted by a final-year project student.

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.

Roberts, K.L., Hall, D.A., & Summerfield, A.Q. (2005) Auditory spatial attention: Effect of different cueing strategies. Poster presented at ESCOP, Leiden, The Netherlands.

Roberts, K.L., Hall, D.A., Summerfield, A.Q. (2004) Conflict resolution in vision and audition: Behavioural and neuroimaging data. Poster presented at Human Brain Mapping, Budapest, Hungary.

Roberts, K.L., Hall, D.A., Summerfield, A.Q. (2004) Alerting, orienting, and executive control in vision and audition. Poster presented at Experimental Psychology Society meeting, London, UK.

Roberts, K.L., Hall, D.A., Summerfield, A.Q. (2004) Spatial orienting of auditory attention. Poster presented at British Society of Audiology meeting, London, UK.

Roberts, K.L., Hall, D.A., Summerfield, A.Q. (2003) Behavioural correlates of auditory attention. Poster presented at British Society of Audiology meeting, Nottingham, UK.

List of Abbreviations

ACC	Anterior cingulate cortex
ANOVA	Analysis of variance
ANT	Attention network test
APD	Auditory processing disorder
BA	Brodmann area
dB	Decibels
dB HL	Decibels (hearing level)
dB(A)	Decibels (A-weighted)
DLPFC	Dorsolateral prefrontal cortex
EEG	Electroencephalography
EPI	Echo-planar imaging
ERP	Event-related potential
f_0	Fundamental frequency
F1 (... FN)	1st formant (... nth formant)
FEF	Frontal eye field
fMRI	Functional magnetic resonance imaging
HRTF	Head-related transfer function
Hz	Hertz
ILD	Interaural level difference
IOR	Inhibition of return
IPL	Inferior parietal lobe
IPS	Intraparietal sulcus
ITD	Interaural time difference
MFG	Middle frontal gyrus
MMN	Mismatch negativity
MNI	Montreal Neurological Institute
PET	Positron emission tomography
rms	Root mean square
RT	Reaction time
rTMS	Repetitive transcranial magnetic stimulation
SEF	Supplementary eye field
SOA	Stimulus onset asynchrony
SPL	Superior parietal lobe
SSQ	The speech, spatial and qualities of hearing scale

TMS	Transcranial magnetic stimulation
TPJ	Temporo-parietal junction
VDU	Visual display unit

Table of Contents

Chapter 1: Introduction and Literature Review

Overview of Thesis

The research presented here investigates auditory attention. Behavioural and neuroimaging techniques are used to investigate the attentional skills of alerting, orienting, and attentional control (Posner & Petersen, 1990). Attention has been a primary topic for cognitive psychology research for a number of decades. However, the majority of this research has been conducted in the visual modality, and there has been considerably less interest in auditory attention. There are both practical and theoretical reasons for attempting to redress this balance. Firstly, there is emerging evidence of auditory attention difficulties in certain clinical groups. A sample of elderly hearing-impaired adults investigated by Gatehouse and Noble (2004) reported difficulties with attentionally-demanding listening tasks, such as following one person speaking and using the telephone at the same time, and following multi-talker conversations without missing the start of each new talker. Difficulties with these situations correlated significantly with self-reported handicap, even after accounting for the effects of hearing loss. A second applied problem is that of distinguishing auditory attention problems from other auditory disorders. For example, recent work investigating the nature of auditory processing disorder (Jerger & Musiek, 2000) has highlighted the importance of differentiating auditory attention difficulties from auditory processing difficulties. In both of these cases, a better understanding of auditory attention would be beneficial, as would a short, convenient test of auditory attention for use with clinical groups. On a more theoretical note, the preponderance of visual attention research leaves open the questions of whether attention is a supramodal facility, and if so, whether interactions with sensory modalities result in different behavioural effects. Whilst it is possible that visual attention research is readily applicable to auditory attention problems, and that tests of visual attention are appropriate for evaluation of auditory attention, a formal test of these possibilities seems timely.

This chapter gives a general introduction to attention research, and the pure and applied reasons for investigating auditory attention. The aim is to provide a brief overview of the background literature: more detailed information is presented within the relevant experimental chapters. To begin, an overview of the research conducted for this thesis is presented. There is then a short history of attention research, a discussion of sub-types of attention and their neural correlates, and a section on the relationship between attention and perceptual modalities. The introduction ends by touching on some of the practical applications of auditory attention research.

The first experimental chapter (Chapter 2) reports an initial experiment which compared performance on a test of visual attention, the Attention Network Test (ANT) (Fan, McCandliss, Sommer, Raz, & Posner, 2002), with performance on an auditory analogue by the same group of subjects. The test investigated three types of attention (alerting, orienting, and executive control), following Posner & Peterson's (1990) classification. Results revealed highly similar reaction time effects on the executive control measure, similar alerting effects, but substantially different spatial orienting effects across the auditory and visual tasks.

Chapter three reports research conducted to investigate whether the similar behavioural results on the executive control measure were the result of common cortical mechanisms. Conflict-resolution tasks are commonly used to investigate executive control. In these tasks, subjects must respond based on a relevant aspect of the target stimulus, while overcoming conflict generated by a competing irrelevant aspect of the same stimulus. A meta-analysis identified cortical areas reliably activated by conflict-resolution tasks. An fMRI study was then conducted in which the same group of subjects performed both auditory and visual conflict tasks. The results are consistent with a supramodal anterior network for conflict monitoring and resolution.

Research reported in Chapter four investigated alerting and orienting further using a vowel-identification task. Subjects were cued to the onset, location, pitch, or both location and pitch of a target vowel, and were asked to identify the target vowel whilst ignoring a concurrently presented distractor vowel. The results show reaction-time benefits from knowing when, where, and at what pitch to attend. However, there was no additive benefit to having both location and pitch information together, suggesting that subjects were orienting to an auditory object comprising both stimulus features. Within this experimental task orienting benefits were relatively slow to build, with the greatest effects occurring with 1050 ms between cue and target onsets.

Chapter five presents a series of seven experiments designed to investigate auditory spatial attention. The first six experiments test the spatial relevance hypothesis (McDonald & Ward, 1999). McDonald and Ward reasoned that auditory spatial orienting could only occur when a spatial representation was available to orient attention towards. They further hypothesised that a spatial representation would only be generated if space were relevant to the task in some way. The spatial relevance hypothesis therefore states that space must be task-relevant in order for auditory spatial orienting to occur. While the hypothesis accounts for a great deal of the variability in the literature, it does not account for the failure to find auditory spatial orienting effects in the auditory ANT, nor for the high inter-subject variability found in some studies. The results of the experiments reported in Chapter 5 suggest that exogenous (automatic) attention effects are robust, but that endogenous (voluntary) attention effects are weak and highly variable across subjects. It is proposed that these findings reflect the way in which auditory spatial information is represented in the midbrain and cortex. The final experiment investigates the influence of the way in which information is presented during auditory cueing tasks. Cues to target location could aid performance as a result of two different mechanisms. Attention could be oriented to a spatial location, in a manner comparable to covert orienting of visual attention. Alternatively, an 'ear selection' strategy could be used, in which attention is directed to an ear, rather than to a genuine spatial location. The results from the final experiment showed spatial-cue benefits only when the presentation method favoured an ear-selection strategy.

The thesis concludes with a summary of the main findings, a discussion of these findings in light of the research aims, and some suggestions for further research.

Section 1: Attention

A Short History of Attention Research

Typically, the following quote from William James is used to begin any review of attention research (e.g. Coull, 1998; Scholl, 2001).

"Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others."

William James (1890), pp. 403-404

This quote is ubiquitous because it appeals to a common-sense understanding of what attention is and how it works. And because a review of the literature rapidly reveals the inaccuracy of the statement that 'everyone knows what attention is'.

Attention became a popular research topic in cognitive psychology in the 1950s. The dichotic listening technique proved a valuable tool for investigating the ability to attend to one stream of information while ignoring another. In a typical experiment of this type (e.g. Cherry, 1953) subjects are presented with one stream of information to each ear, via headphones. Attention is directed to one of the streams by asking the subject to shadow (repeat aloud) the information heard at one ear. These early experiments demonstrated that very little is known about the information presented to the unattended ear, and gave rise to filter theories of attention. The first of these (Broadbent, 1958) proposed that unattended information was filtered out early in processing. According to this early-selection model, only attended information was fully processed. Unattended stimuli were processed for simple, physical properties, but not for semantic content. This highly influential theory contained specific predictions which sparked enthusiasm for further research. Evidence quickly accumulated which demonstrated that some unattended information is processed more comprehensively than the early-selection theory would suggest, and Broadbent's filter theory of attention gave way to a model in which unattended information was attenuated, rather than ignored (Treisman, 1960, , 1969). Further filter theories made small changes to this basic model (e.g. Deutsch & Deutsch, 1963), but essentially all models proposed a selective attention filter, in which unattended information was, at some point, and to some degree, processed to a lesser extent than attended information.

Filter models of attention have been highly influential in how attention has been conceptualised. Nilli Lavie recently continued and expanded on this with the perceptual load hypothesis (Lavie, 1995), according to which processing demands of the attended item will dictate how much processing unattended items receive. This theory successfully reconciled apparently contradictory studies which provided evidence in support of early- or late-selection theories (Driver, 2001). However, there have also been attempts to move away from filter-theory thinking. One alternative is the capacity model of attention (Kahneman, 1973), in which attention is seen as a limited resource. According to this view, the amount of attention available for processing is a function of overall arousal and task demands. If task demands exceed the available processing resources, then decrements in performance will arise. Therefore if two tasks are competing for resources, there will be decrements in performance on one or both tasks if task demands exceed capacity. This theory was rapidly developed to take account of the fact that some tasks can be carried out in parallel without decrements in performance, while others cannot. The new conceptualisation suggested multiple resource capacities, each specific to a perceptual modality or to a type of information processing (Navon & Gopher, 1979; Wickens, 1980).

An alternative concept of attention is that it serves as a link between perception and action. Action-selection views of attention (Allport, 1987; Neumann, 1987) propose that the limits of attention are not due to processing limitations, but result from the need to make appropriate behavioural responses. Once a response has been selected, other responses are necessarily inhibited. In support of this, Tipper, Lortie, and Baylis (1992) presented subjects with an array of buttons, each accompanied by red and yellow lights: red lights indicated that the button was a target; yellow that it was a distractor. The task was to press the target button, and ignore the distractors. When distractors were presented between the hand's starting position and the target, they interfered with task performance significantly more than if they were presented beyond the target, supporting the view that the motor demands of the task influence attentional processes.

While early research with the dichotic listening task was conducted in the auditory modality, the focus soon moved into the visual modality, where stimuli were easier to create and control. Visual search experiments became popular, and enabled researchers to investigate two key questions. Firstly, what are the units of visual information upon which attention operates? And secondly, which stimulus features can attention be directed towards? Addressing the first question, Anne Treisman proposed her feature integration theory (Treisman & Gelade, 1980). Visual search studies typically require subjects to search for a target item in an array of similar items. When the target item is distinguished from non-target items on just one feature (e.g. colour), search time does not increase with increasing numbers of non-targets (a 'pop-out' effect). However, when the target item is defined by more than one feature (e.g. colour and shape), and non-targets can match the target on one of those features, search time increases with increasing numbers of non-targets. Feature integration theory accounts for these findings by suggesting that certain stimulus features (such as colour) are processed preattentively, in parallel. However, features cannot be combined without attention, which must be applied to each item in a serial manner, accounting for the linear increase in search time with increased numbers of targets.

The second question addresses which stimulus features attention is oriented towards. Here there are three (non-exclusive) possibilities: that attention is space-based, feature-based, or object-based. Space-based attention is described using the spotlight metaphor (Eriksen & Eriksen, 1974; LaBerge, 1983). This suggests that attention acts as a spotlight which can be directed to areas of space. This is intuitively appealing given its similarity to the way in which the eyes are moved to foveate areas of interest. Feature-based attention is more general, in that attention can be directed to any stimulus feature, such as its colour or motion. Under this conceptualisation, space is sometimes seen as a special case (Tsal & Lavie, 1993). Support for object-based attention arose out of research investigating the spotlight metaphor, which revealed variations in the size and shape of the spotlight based on the item being attended. Further evidence has demonstrated benefits for processing two within-object features, compared with when these same features are presented between-objects (Duncan, 1984; Egly, Driver, & Rafal, 1994). Recent research, reviewed by Scholl (2001) suggests that task demands may dictate whether attention is directed to a location, a feature, or an object.

Types of Attention

The term 'attention' can be considered an umbrella term incorporating a number of different sub-processes. In general, attention enables the "appropriate allocation of processing resources to relevant stimuli" (Coull, 1998). However, while William James thought that 'everyone knows what attention is', both theories and taxonomies of attention vary between research groups. Spearman (1937) commented on the lack of consensus:

> "But [attention's] towering growth would appear to have been achieved at the price of calling down upon its builders the curse of Babel, 'to confound their language that they may not understand one another's speech'. For the word 'attention' quickly came to be associated ... with a diversity of meanings that have the appearance of being more chaotic even than those of the term 'intelligence'."
>
> Spearman, 1937, p.133, quoted in Scholl (2001)

In general, there are three key types of attention which appear in some form in most taxonomies: sustained attention, selective attention, and attentional control.

Alertness and Sustained Attention

Arousal levels vary over time. These variations can occur over a very short time scale - for example, heightened arousal following a warning cue, or over a much longer time scale, such as circadian variations. There are four key types of attentional control of arousal: intrinsic (tonic) alertness, phasic alertness, sustained attention, and vigilance (Sturm & Willmes, 2001). Intrinsic alertness refers to the cognitive control of arousal, and is assessed using simple reaction time tasks. Phasic alertness is the ability to increase arousal in response to an external stimulus. This is typically assessed in reaction time tasks in which the target stimulus is preceded by a warning cue (Posner, 1978). Sustained attention and vigilance are typically investigated by tasks which require subjects to detect rare targets over a period of some minutes or hours. These tasks involve prolonged attentive processing, but do not include attention-grabbing stimuli. Models of performance on these tasks suggest that they either reveal variations in arousal levels over time, that they reveal changes in signal detection parameters over time, or that they show a change from controlled behaviour to automatic behaviour (Robertson & Manly, 1999). According to the last of these theories, the term ‘sustained attention’ is often wrongly applied to situations which require sustained task performance, but low levels of attention. While sustained attention and vigilance are terms which are often used interchangeably, the difference is thought to lie in the frequency of target stimuli, with vigilance tasks presenting low frequency targets, requiring high levels of endogenous (voluntary) attention, while sustained attention tasks present targets more frequently (Sturm & Willmes, 2001).

Selective Attention and Orienting

Selective attention (also referred to as focused attention and targeted attention) refers to the ability to preferentially process a chosen aspect of the world. Orienting of attention is a term typically used to describe the act of directing attention to a particular spatial location, but studies have also investigated orienting attention towards other features and processes, such as objects, instants in time, and motor responses (Nobre, 2001). Early research using the dichotic listening paradigm with a shadowing task (e.g. Broadbent, 1958) revealed that it is possible to selectively attend to one message while being relatively unaware of another. Research has also demonstrated advantages of attending to stimulus features such as location (Tsal & Lavie, 1993) and colour (Laarni, 1999), and to specific objects (see Scholl (2001) for a review). Tsal and Lavie (1993) presented evidence that orienting to location is a special type of selective attention. In a series of experiments, subjects were presented with a cue followed by an array of letters. The cue indicated which letters in the array should be responded to. While the cue varied in both colour and location, only one of these attributes was relevant in any one condition. The results showed that subjects always attended to the cue location, even when colour was the relevant attribute and location was irrelevant. Tsal and Lavie concluded that attending to the location of a stimulus was a mandatory process which would occur irrespective of the feature subjects were trying to attend.

Studies of spatial orienting of attention are typically conducted using cueing tasks, in which the target stimulus is preceded by a cue which either occurs at a target location, or directs attention towards a target location (Posner, 1978). Selective attention to the cued location can be directed either endogenously (voluntarily, top-down), according to task demands, or exogenously (automatically, bottom-up), in response to an external stimulus. When the cue is predictive of target location, it draws endogenous attention, and is found to speed processing and improve accuracy for targets presented at the cued location (e.g. Fan et al., 2002; Posner, 1978). When the cue appears at a possible target location, but is not predictive of target location (i.e. when the target is equally likely to appear at the cued location as at the uncued location), exogenous attention is drawn to the cue location. With exogenous attention, there is a characteristic timecourse of attention. When the time between cue onset and target onset (stimulus onset asynchrony: SOA) is very short (around 100 ms), subjects are quicker to respond to targets presented at the cued location than at the uncued location. With increases in the SOA, this reaction-time benefit decreases, and is replaced by a reaction-time cost when the SOA is longer than around 300 ms. This cost at longer SOAs is known as 'inhibition of return' (see Klein (2000) for a review), and may act to facilitate visual search behaviour.

Attentional Control

Also referred to as executive control, attentional control refers to situations in which attention is used to process difficult or conflicting information, or to inhibit processing or responses. Conflict resolution and task switching are two tasks commonly used to investigate attentional control. Conflict resolution tasks (see MacLeod and MacDonald (2000) for a review) require subjects to respond according to one stimulus dimension, whilst ignoring an irrelevant stimulus dimension. The classic example of a conflict task is the Stroop task (Stroop, 1935), in which subjects are presented with colour-words written in coloured ink, and asked to name the colour of the ink, while ignoring the word meaning. When the word meaning and the ink colour match (e.g. "RED" written in red ink), or the word meaning is neutral (e.g. "LOT" written in red ink), the subject is faster to respond than when the word meaning and the ink colour are incongruent (e.g. "RED" written in blue ink). The additional time taken to respond is thought to reflect the time needed to inhibit the incorrect response, which is generated due to the rapid and automatic processing of written information. Task switching studies (see Monsell (2003) for a review) require subjects to switch frequently between tasks. Immediately following a task switch, subjects' responses are slower and less accurate, reflecting a 'switch-cost' comprising a carry-over of task set (and the need to inhibit the now inappropriate response), and the need to change task set.

Neural Correlates of Attention

Neuropsychology and neuroscience have proved highly useful to the study of attention in two ways. Firstly, they provide a new means for testing behavioural theories of attention. For example, in assessing the early- versus late-selection debate, ERPs have been used to identify the earliest point at which differences are found in the neural activity arising from attended and unattended items. The results suggest that attentional modulation can occur within 60ms of stimulus onset, indicating that attention can influence sensory coding (see Luck, Woodman and Vogel (2000) for a review of ERP studies of attention). Rees, Russell, Frith and Driver (1999) used fMRI to investigate the degree to which unattended words were processed. Behavioural studies can address this question only through indirect means such as surprise memory tests, which do not distinguish between information which was never processed, and information which was processed but not remembered. The fMRI study revealed no differentiation between processing of unattended words, and processing of random letter strings, providing evidence that the unattended words were not semantically processed.

The second use for neuropsychological and neuroscientific techniques is in identifying the neural mechanisms by which attention operates. There are two conceptualisations of the way in which attentional modulation is achieved. The more prevalent view is that of sources and sites of attention (Posner & Fan, in press): that there are control regions of the brain causing attentional modulations (sources), and other regions which are affected by this modulation (sites). In contrast, Duncan (e.g. Duncan, Humphreys, & Ward, 1997) has argued that there is no reason to assume such a modular system: that attentional modulation can be a product of integrated competition across populations of neurons. Here, I review evidence for the involvement of a number of cortical regions in attention-demanding tasks. It is difficult to differentiate between sites and sources of attention, particularly in light of feedforward and feedback interactions between cortical regions (Nobre, 2001). I have therefore drawn a simple distinction between sensory cortices, which are likely to be influenced by attentional modulation (sites), and other cortical regions, which are more likely to be sources of attentional modulation.

Sites of attention

Attention has been found to modulate activity in primary and secondary sensory cortices. In an early PET study (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990), subjects viewed the same stimuli, but were asked to respond based on only one stimulus characteristic (shape, colour, or velocity). Differential activity was found in extrastriate cortex based on the characteristic being attended. Similarly, Büchel et al. (1998) found that attention to motion led to enhanced activation in extrastriate areas V3a and V5: areas specialised for visual motion processing. Woodruff et al. (1996) presented subjects with simultaneous auditory and visual stimuli, and found that selectively attending to either stimulus enhanced activity in the respective sensory cortex. Attention to auditory stimuli has also been shown to result in greater activity in primary and secondary auditory cortices than passive listening (Jäncke, Mirzazade, & Shah, 1999), and attending to a signal presented to the left or right ear is associated with enhanced activation in the sensory cortex contralateral to the attended side (Alho et al., 1999).

However, some studies have failed to find attentional modulation of activity in primary sensory cortices. Frith and Friston (1996), using PET, found that attention to tones modulated activity in the right midthalamus, but not in auditory cortex. Since ERP studies demonstrate that the N100 arises in auditory cortex and is modulated by attention (Woldorff et al., 1993), Frith and Friston suggest that attention may lead to more synchrony between sources in auditory cortex, rather than more activity, and that the thalamus may be responsible for this synchrony. A review of the visual literature (Kanwisher & Wojciulik, 2000), reports a different pattern of results. Attentional modulation has been found to influence ERPs arising from extrastriate cortex, but not primary visual cortex, while fMRI evidence has successfully identified attentional modulation of activity in primary visual cortex. Martínez et al. (1999) hypothesised that this pattern might be due to the timecourse of attentional modulation: that primary visual cortex activity is modulated by attention, but via top-down influences which do not influence activity until after the sensory ERP components have occurred. An alternative explanation is that attention leads to changes in baseline activation, which are detectable using fMRI but not EEG. In support of this, increased activity in visual cortex has been demonstrated following the instruction to attend to a particular location, but in the absence of visual stimuli at that location (Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999).

Sources of attention

Alertness and Sustained Attention

Sustained attention tasks are consistently associated with activation in a right fronto-parietal network (Coull, Frith, Frackowiak, & Grasby, 1996; Pardo, Fox, & Raichle, 1991; Paus et al., 1997). The involvement of the right hemisphere in sustained attention is further demonstrated by patients with right-hemisphere lesions who experience difficulty with sustaining attention and using warning signals (see Posner and Petersen (1990) for a review). In addition to cortical activation, sustained attention tasks also influence activity in subcortical areas, including the reticular nucleus of the thalamus (Kinomura, Larsson, Gulyas, & Roland, 1996). Paus et al. (1997) used PET and EEG to investigate brain activity during an auditory vigilance task. Accuracy did not vary with time-on-task, but reaction times and EEG activity in the theta range (4 to 7 Hz) increased. Increased time on task also led to decreased activation in both right-hemisphere cortical areas (ventrolateral, dorsolateral, and orbital frontal cortex, parietal cortex and temporal cortex) and subcortical areas (thalamus, substantia innominata, and putamen). Paus and colleagues suggested that the decreased subcortical activity reflected changes in arousal over time, while the decreased cortical activity reflected a shift from controlled to automatic attentional processing.

Sturm and Willmes (2001) compared activation during tests of intrinsic and phasic alertness. The same simple reaction time task was used in both conditions, but in the phasic alertness condition targets were preceded by a warning cue. Intrinsic alerting was associated with right-hemisphere activation in the anterior cingulate cortex, dorsolateral prefrontal cortex, inferior parietal lobe, middle and superior temporal gyrus, right thalamus and dorsal pontomesencephalic tegmentum. When a warning cue was presented prior to target onset, additional activation was found in the thalamus, and left-hemisphere superior and ventrolateral frontal gyrus. Since left frontal activation is associated with selective attention, Sturm and Willmes suggest that it reflects inhibition of responses to the warning stimulus.

Similar results were found by Fan, McCandliss, Fossella, Flombaum, and Posner (2005), who investigated cortical activation following an alerting cue. They found the classic fronto-parietal activation pattern, along with activation of the thalamus. However, their results showed left-hemisphere dominance in both the parietal lobe and frontal areas. Fan and colleagues suggested that this change in the typical pattern may reflect the difference between sustaining attention from one trial to the next (which would show the typical right-hemisphere dominance) and specific activation in response to an alerting cue, which has been associated with left-hemisphere activation (Coull, Nobre, & Frith, 2001).

Selective Attention and Orienting

Selectively attending to a spatial location is associated with a fronto-parietal network of activation (Kanwisher & Wojciulik, 2000). Kastner et al. (1999) used fMRI to reveal specific regions involved in orienting visual attention to a cued location, even in the absence of target stimuli at that location. The study revealed activation in the frontal eye fields (FEF), supplementary eye fields (SEF), superior parietal lobe (SPL), and around the intraparietal sulcus (IPS). There was also less reliable activation in the inferior parietal lobule (IPL) and middle frontal gyrus (MFG). While this study found activation bilaterally, other studies have found enhanced activation in right-hemisphere parietal cortex (e.g. Gitelman et al., 1999).

Spatial neglect is an attentional disorder which can arise following unilateral brain injury (such as that resulting from a stroke). Patients with neglect fail to attend to the contralesional side of space. Much research has been conducted to identify the common locus of lesions which result in deficits in spatial attention. The results show clear hemispheric asymmetry, with neglect more likely to occur following lesions to the right hemisphere (Vallar, 1998). However, within the right hemisphere more than one site has been suggested as the locus of the attentional deficit. The area most frequently associated with neglect is the supramarginal gyrus in the inferior parietal lobule at the temporo-parietal junction (Halligan, Fink, Marshall, & Vallar, 2003). Other potential regions are the lateral premotor cortex in the frontal lobe, and subcortical structures such as the thalamus and basal ganglia (Halligan et al., 2003). The superior temporal gyrus has also been proposed as the crucial lesion site (Karnath, Ferber, & Himmelbach, 2001). Since lesions are generally large, and the extent of functional damage can extend beyond the lesion site (Hillis et al., 2005), conclusions regarding the functional role of lesioned sites are complicated. Evidence from neglect patients is broadly consistent with evidence from neuroimaging studies in indicating a neural network for spatial attention which includes frontal premotor and posterior parietal regions (Halligan et al., 2003). However, the parietal region implicated by neglect studies appears inferior to that implicated by imaging studies (Nobre, 2001). This may reflect the difficulty involved in isolating regions critical to neglect, or may indicate that neglect arises from disruption to specific selective-attention processes. Corbetta, Kincade, Ollinger, McAvoy, and Shulman (2000) reported that when subjects attended to a cued target location, activation was found around the IPS, even in the absence of target stimuli. However, when subjects were cued to the wrong location, activation was found in the right temporo-parietal junction (TPJ). This suggests different functional roles for these two regions, with IPS responsible for orienting attention, and TPJ involved in disengaging attention.

Orienting attention endogenously (voluntarily) and exogenously (automatically) give rise to similar patterns of activation. Rosen et al. (1999) compared activation when subjects were cued to target location by an informative central arrow (endogenous) and by a non-informative peripheral cue (exogenous). Overlapping activation was found bilaterally in the posterior parietal cortex and frontal eye fields. Activation was enhanced in the endogenous condition, reflecting more effortful orienting than in the exogenous condition (Rosen et al., 1999). Right dorsolateral prefrontal cortex (DLPFC) was additionally activated in the endogenous condition, possibly suggesting the involvement of working memory during the centrally-cued trials. Similarly, Peelen, Heslenfeld, and Theeuwes (2004) found no difference in cortical activity when subjects performed exogenous and endogenous orienting tasks. The superior colliculus is a small region located in the midbrain that is responsible for reflexive head and eye movements (Sparks, 1999), and has been shown to be involved in exogenous orienting in macaque monkeys (Robinson & Kertzman, 1995). Peelen et al. (2004) hypothesised that the superior colliculus may have been active during their exogenous orienting task, but that the small size of the region may have resulted in changes in activation being too small to detect using fMRI.

Attention can be oriented to a target location either overtly (with a corresponding physical movement, such as an eye or head movement) or covertly (without a corresponding physical movement). Corbetta et al. (1998) directly compared overt and covert visual attention, and found that covert shifts of visual attention and saccadic eye movements are associated with overlapping areas of parietal, frontal, and temporal lobes, implying that a similar mechanism is used to orient visual attention in both cases. Nobre, Gitelman, Dias, and Mesulam (2000) also compared activation during a covert orienting task and a task in which subjects made saccadic eye movements. Consistent with Corbetta et al.'s study, they found extensive overlap of activation in the two tasks, but with enhanced activation bilaterally in the FEFs and around the IPS in the covert orienting task. The covert orienting task additionally activated right DLPFC, which may reflect the involvement of working memory in this task, but not in the saccadic eye-movement task.

Attentional Control

It is consistently observed that the anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (DLPFC) are involved in attentional control (Smith & Jonides, 1999). The precise role of these areas remains unclear, but there is evidence to suggest that the ACC performs a monitoring role, while the DLPFC influences perceptual or response processes.

ACC is the region most reliably active during attentional control tasks. Examples include response conflict arising from Stroop (MacLeod & MacDonald, 2000) and flanker tasks (van Veen, Cohen, Botvinick, Stenger, & Carter, 2001); task set-switching and response suppression (Swainson et al., 2003); divided attention (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991); dual task performance (D'Esposito et al., 1995), monitoring for errors

(Carter et al., 1998); and tasks which require a response which is not fully constrained, such as verb generation (Barch, Braver, Sabb, & Noll, 2000). In verb generation tasks subjects are presented with a noun and asked to generate a verb associated with that noun. Nouns associated with a small number of verbs (such as 'bell') are highly constrained, while nouns associated with a large number of verbs (such as 'ball') are weakly constrained. Barch et al. found enhanced activation in ACC when subjects were presented with weakly-constrained nouns, compared with highly-constrained nouns. DLPFC is frequently, although not always, active during attentional control tasks such as these. It has been suggested that its role is to bias processing in favour of task-relevant responses (Badre & Wagner, 2004).

One influential theory of ACC function is that it performs a conflict-monitoring role, recruiting other brain regions (such as DLPFC) to then resolve this conflict (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999). This theory is based on the finding that ACC is active during three different types of situation in which conflict occurs (Botvinick, Cohen, & Carter, 2004). First, situations in which a prepotent response must be overridden, (e.g. Stroop tasks). Second, situations in which the response is underdetermined, such that several possible responses present themselves simultaneously and one must be selected (e.g. verb generation). Third, situations in which errors are likely. In this situation, Botvinick and colleagues suggest, even when an incorrect response has been selected, ongoing processing of the stimulus may lead to a correct response also being generated which will then conflict with the initial response.

Kerns et al. (2004) investigated the role of ACC in resolving Stroop conflict by analysing how the interaction between current and previous trial types affected behavioural responses and cortical activity. They hypothesised that responding to an incongruent trial would engage enhanced cognitive control, and that this would be reflected in performance on the subsequent trial. The results support this hypothesis. Reaction time costs associated with incongruent stimuli were smaller following an incongruent trial than following a congruent trial. In addition, ACC activity was reduced on incongruent trials which followed incongruent trials (iI) relative to those which followed congruent trials (cI). They additionally found that greater ACC activation resulted in a greater reduction in the reaction-time cost on the subsequent trial. This study also provided suggestive evidence that ACC recruits DLPFC to resolve the conflict. ACC activity was correlated with DLPFC activity on the subsequent trial, and trials on which the behavioural effects of conflict were most reduced were associated with increased activity in DLPFC.

Section 2: Attention & Perceptual Modalities

Supramodal vs. Intramodal debate

Opinion varies over whether attention is a supramodal facility or whether there are modality-specific resources. At one extreme, separate attentional facilities are suggested for each perceptual modality (Wickens, 1980). At the other extreme, one attentional resource is thought to operate regardless of perceptual modality (Farah, Wong, Monheit, & Morrow, 1989). Less extreme hypotheses have also been suggested. There may be separate attentional systems with strong crossmodal links (Spence & Driver, 1996). Alternatively, there may be separate attentional systems which are subservient to a higher-level supramodal system (Posner, 1990). Since spatial information is coded differently across the different senses, spatial orienting of attention is the prime candidate for modality-specific effects. However, some research has also investigated modality-specific attentional effects in sustained attention and alerting studies, and there is tentative evidence from attentional control studies.

Selective Attention and Orienting

Researchers have hypothesised that if there is one supramodal system controlling shifts of spatial attention, then it would be impossible to direct attention to one location in one modality, and to another location in a different modality. To test this hypothesis, Spence and Driver (1996) devised a series of spatial orienting experiments, in which subjects responded to auditory and visual targets in an orthogonal cueing paradigm. Subjects were cued to either their left or right, but made a response based on whether a target was presented from above or below head-height. In Experiment 2, subjects were presented with cues to target location. On approximately three quarters of trials, an auditory target was presented, with visual targets presented on the remaining quarter. The cue accurately predicted the location of seven out of eight auditory targets, but only one out of three visual targets. Subjects would therefore benefit from directing their auditory attention in the cued direction, and their visual attention in the uncued direction. In fact, subjects were faster to respond to both auditory and visual targets at the cued, rather than uncued, location. However, they received less benefit from cues in the secondary modality (i.e. vision). Experiment 3 reversed the modalities and found similar effects. In a further demonstration, Experiment 6 provided cues which were 80% correct for one modality, and 80% incorrect for the other, encouraging participants to direct their attention in different directions for different modalities (but this time with no bias towards one modality). There were no differences in reaction times to cued and uncued targets, suggesting that subjects were unable to achieve this. However, a small change in task design led to a different pattern of results. Experiment 7 used a blocked design in which 80% of auditory targets were presented to one side, and 80% of visual targets were presented to either the same side, or the opposite side. This allowed subjects to sustain their attention at the likely location(s). Using this design, they found faster reaction times on the more likely side for both auditory and visual targets. When the likely side for auditory targets was opposite to that for visual targets, effects were smaller than when the likely side was common to both, but still significant. Similar findings were found in a series of comparable studies investigating vision and touch (Spence, Pavani, & Driver, 2000). The finding that attention can be directed to different locations in different modalities argues against the supramodal hypothesis, while the weaker effects, and the tendency to direct attention to one location for both modalities argues against entirely separate attentional systems. Spence and Driver argue that their results are consistent with separate-but-linked attentional systems.

Eimer (1999) recorded ERPs while subjects detected auditory and visual targets. Subjects were asked to respond to auditory targets presented to one side, and visual targets presented to either the same side or the opposite side. Attentional modulation of sensory-specific ERP components occurred only when subjects attended to the same side in both modalities. This result favours the theory that there is a supramodal attentional system for spatial orienting. The study also found attentional modulation of a late ERP component (> 200ms post-stimulus) both when subjects attended to the same side in both modalities, and when they attended to different sides (although the effect was smaller in the latter condition). Eimer suggested that this late component may reflect post-perceptual processing, and may explain the results found by Spence and Driver.

A second approach which has been used to investigate whether attention is a supramodal facility is to present cues in one modality, and targets in a different modality. Research following this line of questioning has shown a mixed pattern of results. Spence and Driver (1997) found that an uninformative visual, auditory, or somatosensory cue could draw attention to a target in any one of those modalities, with the exception that a visual cue did not draw attention to an auditory target, at least when eye movements were prevented. In contrast, Ward (1994) found that while visual cues could draw attention to auditory targets, auditory cues could not draw attention to visual targets. These studies used different methodologies, which may account for the differences in the pattern of results. Broadly, the asymmetries found in both studies may reflect differences in sensory processing of spatial information across modalities. While visual and somatosensory information is processed spatiotopically, the location of auditory sound sources must be computed from non-spatial properties of the stimulus, including interaural time and level differences, and spectral cues introduced by the head and external ears. Auditory information is initially coded tonotopically, and while the superior colliculus contains spatiotopic maps of auditory space, there is no evidence for similar maps in the cortex (Middlebrooks, 2000).

Finally, ERP and fMRI studies have been conducted to investigate whether common cortical areas are involved in spatial orienting within different modalities. Macaluso, Eimer, Frith, and Driver (2003) used fMRI to investigate attention-related modulation of processing of visual and tactile targets, but also activation associated with the preparatory interval between cue and target. In response to a symbolic auditory cue directing attention to the left or right, activation was enhanced in contralateral occipital areas when a visual target was anticipated, and in contralateral somatosensory cortex when a tactile target was anticipated, irrespective of whether a target was actually presented. In addition, activation was found in contralateral intraparietal sulcus when targets were anticipated in either modality. Further activation was found in bilateral premotor cortex, left inferior parietal lobule, superior frontal gyrus and the precuneus irrespective of the attended side and modality. These results demonstrate both supramodal and modality-specific aspects of the attention system.

Using a similar approach, Eimer and van Velzen (2002) measured ERP components when subjects were cued to both target side and target modality (vision or touch) by a symbolic auditory cue. During the interval between cue and target, anterior negative components and occipital positive components were found contralateral to the cued side, irrespective of target modality. In addition, attention modulated early (sensory) ERP components to a similar extent for both the relevant and irrelevant modalities. Macaluso et al. (2003), and Eimer and van Velzen (2002) both concluded that their results demonstrate supramodal control of spatial attention processes.

Transcranial magnetic stimulation (TMS) provides a way of investigating areas which are necessary for spatial orienting, rather than simply involved. Chambers, Stokes, and Mattingley (2004) applied repetitive TMS (rTMS) while subjects oriented their attention to either visual or somatosensory stimuli. They found that rTMS applied to the inferior parietal cortex interfered with spatial orienting to visual, but not tactile, stimuli. This result suggests that modality-specific attentional processing occurs in the parietal cortex, and that supramodal activation found in fMRI and ERP studies in fact reflects synchronisation between shifts of attention to targets in different modalities, rather than necessary recruitment of those regions. This finding supports Spence and Driver's 'separate-but-linked' theory for how attention may operate across different perceptual modalities. However, Chambers et al. acknowledge that their results do not exclude the possibility of supramodal areas located elsewhere in the cortex.

Overall, it appears that there are both supra- and intra-modal aspects of spatial orienting of attention. Different methodologies produce different results, and while it appears unlikely that attention operates in an entirely supramodal or modality-dependent manner, there is not yet sufficient evidence to draw any firm conclusions about the exact mechanism by which attention interacts with sensory modalities.

Alertness and Sustained Attention

Neuroimaging studies provide evidence that sustained attention may be a supramodal facility. Pardo et al. (1991) investigated cortical activity during visual and somatosensory tasks of sustained attention. They found increased activation in prefrontal and superior parietal cortex, primarily in the right hemisphere, regardless of the modality of the sensory input. Similarly, Kinomura et al. (1996) found similar activation associated with intrinsic alertness during visual and somatosensory tasks. Similar (but not identical, cf. Sturm et al., 2004) activation patterns were also found from intrinsic alertness studies which presented the same tasks, but with visual (Sturm et al., 1999) and auditory (Weis et al., 2000) stimuli. Sturm and Willmes (2001) suggest that this pattern reveals a "supramodal right-hemisphere network for the control of intrinsic alerting". Additional experiments using auditory and visual warning cues with auditory targets revealed similar activation patterns for auditory and visual phasic alertness (Sturm & Willmes, 2001).

Attentional Control

Overwhelmingly, attentional control studies are conducted using visual stimuli (MacLeod, 1991; Monsell, 2003). There are however a few examples of studies which present stimuli in different modalities. For example, Green and Barber (1983) created an auditory conflict task in which subjects heard the words 'man', 'girl', 'mill' and 'game' spoken by male and female voices. Subjects responded by saying 'man' when the voice was male, and 'girl' when the voice was female. Their results show typical effects found in visual conflict tasks (i.e. longer reaction times and less accuracy with incongruent stimuli, compared with neutral stimuli). Similarly, McClain (1983) found typical interference effects when subjects heard the words 'high' and 'low' spoken in a high- or low-pitched voice, and were asked to respond to the pitch of the voice with a verbal or button-press response.

Section 3: Practical Applications of Auditory Attention Research

Hearing impairment, disability, and handicap

Recent research has shown that some elderly, hearing-impaired adults may experience auditory attention difficulties, and that these difficulties have a negative impact on their self-reported handicap. The Speech, Spatial and Qualities of Hearing Scale (SSQ) (Gatehouse & Noble, 2004) is a self-report questionnaire addressing a number of challenges for hearing in everyday life.

In developing the questionnaire, Gatehouse and Noble (2004) hoped to address two key questions: What is disabling about hearing impairment, and how do these disabilities determine the experience of handicap? Their results suggest that difficulty with listening situations involving divided or rapidly switching attention may enhance feelings of handicap.

The SSQ comprises three sections. Each section contains a number of questions to which subjects respond on a visual-analogue scale, ranging from no disability to a great deal of disability. The *speech hearing* section contains 14 questions about the ability to hear speech in a number of different contexts and environments. For example, "You are in a group of about five people, sitting round a table. It is an otherwise quiet place. You can see everyone else in the group. Can you follow the conversation?" The *spatial hearing* section asks 17 questions about the ability to hear where the source of a sound is located, how far away it is, and in which direction it is moving. *Qualities of hearing* is a heterogeneous category which explores the personal experience of sound, for example, the clarity and naturalness of voices and sounds, and the effort involved in listening to speech. There are 19 items in this section.

To assess which aspects of hearing difficulty have the strongest impact on quality of life, scores on the SSQ were correlated with scores on a hearing-handicap questionnaire. The hearing-handicap questionnaire addressed emotional distress and discomfort, social withdrawal, and general restriction on participation. In the *speech hearing* section, the items which correlated most highly with handicap are those which describe attentionally-demanding situations. These items include following one person speaking and using the telephone at the same time, following multi-talker conversations without missing the start of each new talker, and talking with one person whilst the television is on. This relationship remains even when hearing impairment, i.e. hearing thresholds, is partialled out.

As attention is implicated in the self-perceived handicap of adults with hearing impairment, there is a need for an objective measure of auditory attentional skills which would enable attention to be assessed as part of a standard audiological assessment. At present, typical tests of hearing impairment assess the ability to listen to a voice or sound in either a quiet room or in noise. However, both the voice or sound, and the background noise, are presented at a predictable place and time, meaning that there is no comprehensive assessment of attentional skills.

The exact link between attentional (or other cognitive) deficits, age, and hearing impairment is yet to be determined. The average age of respondents in Gatehouse and Noble's study was 71, and each had presented with a hearing impairment. It is likely that a decline in attentional abilities is associated with ageing, and that it is particularly noticeable in adults with hearing impairments because of an interaction between attention and hearing impairment. An objective test of auditory attentional skills would enable progress to be made in understanding this relationship.

Auditory processing disorder

A similar situation arises with a different type of auditory problem. People with auditory processing disorder (APD) experience difficulty processing auditory information. This disorder is newly-identified, and currently very little is known about it. During this early period of trying to identify diagnostic tests and criteria (e.g. Jerger & Musiek, 2000; Keith, 2000), it is particularly important to reliably exclude other potential problems. Auditory attention is one such possible confound (Jerger & Musiek, 2000). Current tests of attentional skills used with clinical populations are either entirely visual (Fan et al., 2002) or contain both visual and auditory sub-tests (Manly et al., 2001; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996), with no clear differentiation between the sub-types of attention being measured and the modality of the test stimuli. The link between auditory and visual attentional skills is not yet sufficiently understood to make these tests an ideal choice for differentiating between auditory processing difficulties and auditory attention difficulties, particularly when an interaction between the two may exist.

Chapter 2: The Attention Network Test

Introduction

The Attention Network Test (ANT) (Fan et al., 2002) is a test of visual attention designed to evaluate the attentional skills of alerting, spatial orienting, and executive control. Separate measures of each skill are derived by comparing performance across different trial types. A cueing task (Posner, 1978) is used to assess alerting and spatial orienting, while a flanker task (Eriksen & Eriksen, 1974) is used to assess executive control. In a cueing task subjects are presented with cues prior to target onset. The cues provide information about the target, such as when or where it will occur. In a flanker task subjects are required to respond to a central target while ignoring distracting information presented on either side of the target. In the original behavioural study (Fan et al., 2002) 40 subjects were tested using the ANT. These subjects were significantly faster to respond to the target following a warning cue (alerting), and gained an additional benefit from a warning cue which also cued target location (orienting). Target stimuli were horizontal arrays of five arrows, in which the central target arrow was flanked by two distractor arrows on each side (see Figure 2.1c, page 45, for example stimuli). The task was to press a button to indicate whether the central arrow pointed to the left or to the right. Flanking arrows were either congruent (pointed the same way as the central arrow), neutral (straight lines), or incongruent (pointed the opposite way to the central arrow). Subjects were significantly slower to respond to incongruent stimuli compared with congruent stimuli (executive control). The response cost reflects the need to inhibit the incorrect response elicited by the distractor items. Importantly, Fan et al. (2002) reported no significant correlations between the three measures of attention, indicating that these types of attention may be independent of each other.

The studies conducted for this chapter were a re-implementation of the visual ANT, an auditory analogue of the ANT, and another visual conflict task: the colour-word Stroop task. These tests were used to investigate similarities and differences between auditory and visual attention. For example, is the benefit obtained from being cued to target location equivalent across modalities? The ANT was selected for two main reasons. First, it purports to test three fundamental and independent types of attention, and clearly defines the skills being measured. Second, it tests all three types of attention within a short, simple test (around 30 minutes), and is suitable for use with children (Mezzacappa, 2004; Rueda et al., 2004) and with clinical groups (Posner et al., 2002; Wang et al., 2005).

Independence of the Networks?

While Fan et al. (2002) found evidence for functional independence between the attentional networks, there is also evidence to suggest that at the very least there is some cross-talk between them. Evidence which supports the independence of the attentional networks comes from neuroimaging and neurochemical studies, which suggest that each type of attention is associated with specific cortical regions and neurotransmitters. Studies of sustained attention (increased arousal over a long time period) have identified a right fronto-parietal network (Coull et al., 1996), and a role for the thalamus (Kinomura et al., 1996). However, studies which presented warning cues of the type used in the ANT have revealed left-hemisphere frontal activation (Coull et al., 2001; Sturm & Willmes, 2001). This may reflect inhibition of responses to the warning stimulus (Sturm & Willmes, 2001) or specific activation in response to the alerting cue (Fan et al., 2005). Neurochemical studies have shown that sustained attention and increased arousal following warning cues are influenced by changes to levels of norepinephrine (Marrocco & Davidson, 1998). Orienting of attention to a spatial location has been associated with a right fronto-parietal network of activation (Kastner et al., 1999), which includes activity around the intraparietal sulcus (IPS), temporo-parietal junction (TPJ) and frontal eye fields (FEF) (Corbetta et al., 1998; Corbetta et al., 2000). Neurochemical studies associate selective attention with the cholinergic system (Marrocco & Davidson, 1998). Executive control, as measured by conflict resolution tasks such as the flanker task used in the ANT, has been associated with activity in anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (DLPFC) (Smith & Jonides, 1999). There is some suggestion that dopamine may play a role in executive control (Posner & Fan, in press).

To investigate the independence of the cortical networks, Fan et al. (2005) used event-related fMRI to measure cortical activity while subjects performed the ANT. This allowed a direct comparison of cortical activity associated with each attentional network within the same set of subjects. The results for each attention network were consistent with those found when each type of attention is investigated in isolation (as described above). A conjunction analysis revealed no overlap between the orienting network and either the alerting or executive control networks, but there were two areas of overlap between the alerting and executive control networks: in the thalamus and left fusiform gyrus. This result suggests that the three components of performance in the ANT may not engage completely separable cortical networks.

In addition to this neuroimaging evidence, there is also behavioural evidence showing interdependence between the attentional networks. In the original ANT experiment Fan et al. (2002) found that their measure of executive control (slower responses to incongruent compared with congruent stimuli) was reduced on no cue and spatial cue trials relative to trials which contained warning cues but no spatial information. Callejas, Lupiáñez, and Tudela (2004) adapted the ANT for the specific purpose of investigating interactions between the networks. A warning tone was presented on half of trials. After a 400ms pause a spatial cue was presented at either the correct (valid) or incorrect (invalid) target location. The cue was not predictive of target location (i.e. the target was equally likely to occur at the cued and uncued locations). The interval between the spatial cue onset and the target onset (stimulus onset asynchrony: SOA) was 100 ms. Under these circumstances, Callejas et al. (2004) found interactions between measures of all three types of attention. A measure of alerting was calculated by subtracting reaction times on trials with a warning tone from those without. Similarly, orienting was assessed by subtracting reaction times on trials with a valid cue from those with an invalid cue, and executive control was evaluated by measuring the amount of conflict introduced by congruent stimuli relative to incongruent stimuli. Callejas et al. (2004) found that an alerting cue increased both the orienting and executive control effects, and that the conflict effect was reduced following a valid cue. They concluded that while the networks are independent, as demonstrated by the absence of significant correlations between them in the original ANT study, and the evidence for a largely independent functional neuroanatomy, they are however influenced by each other in order to produce 'efficient and adaptive behaviour' (p. 227).

Comparison of auditory and visual attentional networks

There is very little research comparing auditory and visual alerting and executive control. What there is, though, does not suggest substantial differences between modalities (Green & Barber, 1983; Kinomura et al., 1996; Pardo et al., 1991). In contrast, spatial orienting of attention has been the subject of a great deal of intra- and cross-modal research, which has shown differences between modalities. This might be expected based on differences in the way in which perceptual information is coded. The visual and somatosensory systems code information spatiotopically, while the auditory system predominantly codes information tonotopically. The location of the source of auditory stimuli needs to be calculated using non-spatial information such as differences in the timing and level of the signals arriving at each ear, and spectral differences introduced by the head and external ears. Research investigating purely auditory spatial attention (Bédard, El Massioui, Pillon, &

Nandrino, 1993; Buchtel & Butter, 1988; Mondor, Breau, & Milliken, 1998; Posner, 1978; Spence & Driver, 1994) has produced variable results: some studies show benefits from orienting auditory attention to a cued location, others show no benefits. Much of this variability is accounted for by McDonald and Ward's (1999) 'spatial relevance hypothesis', which proposes that spatial cues will only be beneficial if the task requires subjects to generate a spatial representation of the auditory stimuli. In other words, only when space is relevant to the task will subjects generate a spatial representation of the task, and only when subjects have generated a spatial representation will they be able to benefit from spatial orienting of auditory attention. The hypothesis suggests that space can be relevant to a task through either of two mechanisms. First, if the task requires subjects to make a spatial discrimination, such as deciding whether a stimulus was played from a top or bottom speaker. Second, if subjects are cued to a spatial location using informative cues, thereby inviting subjects to orient attention voluntarily to a spatial location. The ANT presents informative spatial cues, and should therefore be suitable for eliciting auditory spatial attention benefits.

This study compares performance by the same group of subjects on the visual ANT and on an auditory analogue of the ANT which I have developed. Posner's (1978) cueing paradigm is used to generate a measure of alerting (improved performance following an alerting cue) and orienting (an additional improvement following a cue which also cues target location). In the visual task targets were presented above and below a central fixation point, and cues were provided in the form of asterisks. In the auditory analogue targets were presented monaurally to the left and right ear over headphones, and cues were short noise bursts. A flanker task (Eriksen & Eriksen, 1974) was used to evaluate executive control in the visual modality. The flanker task requires subjects to respond to a central arrow while ignoring flanking arrows. The auditory analogue used a pitch-word Stroop task, in which subjects responded to the pitch of a voice while ignoring the word being spoken. The words were 'high', low', and 'day', and were spoken on a high or low pitch. Both the flanker task and the Stroop task are well-established tests of executive control. In both tasks the congruent condition is that in which the to-be-ignored information leads to the same response as the relevant information. The incongruent condition requires subjects to inhibit the incorrect response elicited by irrelevant information. The neutral condition provides irrelevant information which does not elicit a possible response, and therefore should not affect task performance. Two measures of executive functioning can be calculated from these trial types: interference (the cost associated with ignoring incongruent information), and facilitation (the benefit derived from concurrently presented congruent information).

Interference and facilitation are asymmetric effects, in that the costs incurred by incongruent stimuli are much larger than the benefits obtained from congruent stimuli. MacLeod and MacDonald (2000) report that facilitation effects are typically around 20 ms or less, while interference effects are more likely to be 100 ms or more. Since facilitation effects are typically small, and the exact mechanism by which facilitation arises is the subject of some debate

(MacLeod & MacDonald, 2000; Wright & Wanley, 2003), I will use interference as the primary measure of executive control.

While the flanker task and Stroop task are both commonly used tests of executive control, they may not be directly comparable. Fan, Flombaum, McCandliss, Thomas, and Posner (2003) compared cortical activation while subjects performed colour-word Stroop, flanker, and spatial conflict tasks. While all three tasks were successful in generating behavioural correlates of response conflict, these response-conflict measures did not correlate across tasks. Moreover, while some cortical areas were commonly activated by the tasks, there were also a number of areas which showed task-specific activation. In addition, Hazeltine, Poldrack, and Gabrieli (2000) suggest that the flanker task can be performed by selectively attending to the relevant item in the display, introducing a visuospatial element which is not present in Stroop tasks. To address these differences, subjects will also be tested on a visual colour-word Stroop task. This will permit a comparison of measures of executive control elicited by visual tests (colour-word Stroop and flanker tasks) and measures elicited by Stroop tasks (colour-word and pitch-word Stroop tasks).

The measures of alerting, orienting, and executive control will be compared across modalities to determine whether these attentional skills are supramodal or relate specifically to the perceptual modality in which the stimuli are presented. A supramodal mechanism would be implied if auditory and visual measures correlate significantly with each other, even if the absolute magnitude of those measures varies. To further address the issue of independence between the networks, correlations between the networks within each modality will be investigated, as will interactions between the cueing and conflict conditions.

Method

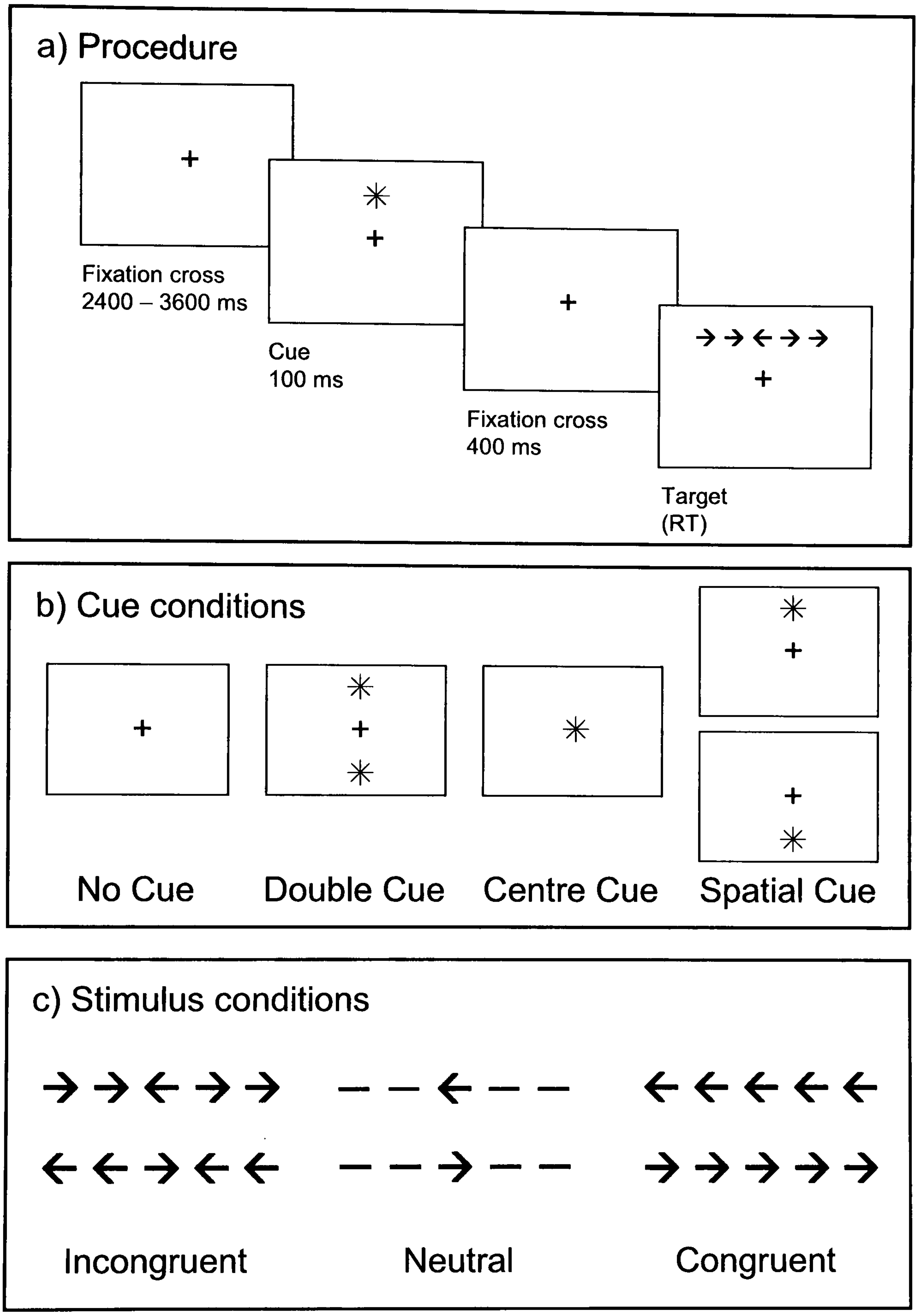
Visual Attention Network Test

The visual ANT was presented in the form described by Fan et al. (2002). Subjects observed stimuli on a VDU. Each trial began with a fixation cross at the centre of the display screen for a short, variable period of time (between 2400 and 3600 ms). A cue was then presented in the form of a briefly presented (100 ms) asterisk, followed by a 400-ms pause during which the fixation cross was again visible. Then the stimulus appeared, either above or below the fixation cross. The subject's task was to indicate with a button press whether a target arrow was pointing to the right or to the left. The time course of these events is shown in Figure 2.1a.

Performance with different cue types provides information about the ability of the subject to alert and to orient. There are four cue types (Figure 2.1b): *no* cue; a single *central* cue; a *double* cue (an asterisk at both possible stimulus locations: above and below the fixation cross); and a *spatial* cue (presented at one of the possible stimulus locations). The alerting effect is calculated by subtracting the subject's median reaction time on the double-cue trials from their median reaction time on the no-cue trials. Fan et al. (2002) selected the double cue, rather than the single cue, because it keeps visual attention relatively diffuse across the two possible stimulus locations, as it would be in the no-cue condition. The orienting effect is calculated by subtracting the subject's median reaction time on the spatial-cue trials from their median reaction time on the single-cue trials. The single cue is selected because it focuses attention on one area, as does the spatial cue.

There are six possible stimuli (Figure 2.1c). Each stimulus consists of five items which are presented against a uniformly grey background. The central item of each stimulus is the target arrow, which points to the right or to the left. The flanking items are either two straight lines (without arrow heads) on either side of the target (neutral condition), arrows pointing in the same direction as the target arrow (congruent condition) or arrows pointing in the opposite direction from the target arrow (incongruent condition). The subjects' task is to indicate with a button press whether the target arrow points to the right or to the left. When viewed from a distance of 65cm, a single arrow subtends 0.55° of visual angle, the spaces between the items subtend 0.06° of visual angle and the entire stimulus (target arrow plus four flankers) subtends a total of 3.08° of visual angle. Each stimulus appears 1.06° above or below the fixation cross.

Figure 2.1: The visual ANT procedure (a), cue conditions (b), and stimulus conditions (c).

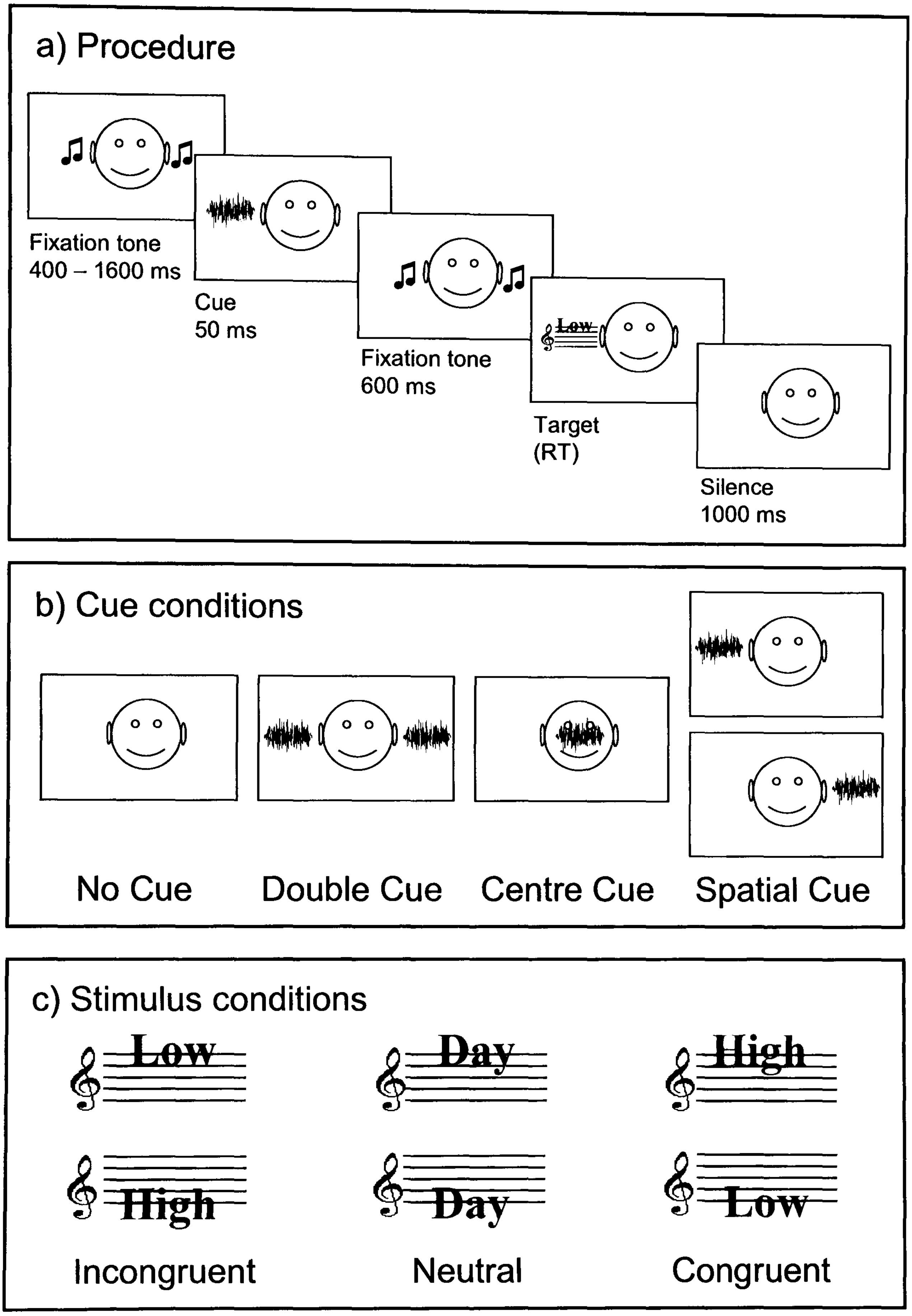


The incongruent, neutral, and congruent stimuli provide a measure of interference (incongruent trials - neutral trials) and facilitation (neutral trials - congruent trials). The interference effect gives a measure of executive control, in that the more the subject's reaction time is slowed down by the incongruent stimulus, the more difficulty they have responding to the target signal in the presence of conflicting information. Fan et al. (2002) obtained their measure of executive control by subtracting the subject's average reaction time in the congruent condition from their average reaction time in the incongruent condition. However, this method confounds the effects of interference and facilitation, so the analyses outlined above will be conducted.

Auditory Attention Network Test

The auditory ANT follows a similar protocol to that of the visual ANT (Figure 2.2a). Each trial began with a 500-Hz fixation tone, presented diotically for a variable period of time (between 400 and 1600 ms). The fixation tone was identical at the two ears, and would therefore be perceived as having a compact source located at the centre of the head. A 50-ms auditory cue was then presented, after which the fixation tone was presented for a further 600 ms (giving an SOA of 650 ms). The target word was then presented to the left or right ear, and the subjects' task was to indicate with a button press whether the target word was spoken on a high or low pitch. Each trial ended with a one second period of silence prior to the onset of the next trial.

Figure 2.2: The auditory ANT procedure (a), cue conditions (b), and stimulus conditions (c).



The pattern of cue conditions (Figure 2.2b) follows that of the visual ANT, but with the spatial cues being presented to the left or right ear of the subject, rather than above or below the fixation cross. Auditory cues were 50-ms bursts of speech-shaped noise, gated with a 10-ms cosine window at the start and end, and otherwise of constant amplitude. The central cue was created by presenting a noise burst diotically (identically to each ear, heard at the centre of the head). The double cue was created by presenting two statistically independent noise bursts, one to each ear. Listeners report hearing uncorrelated noise as either two separate sounds at the two ears, or a spatially diffuse sound (Blauert & Lindeman, 1986). In both cases, attention would be kept relatively diffuse across the possible target locations, as with the visual double-cue condition. The spatial cues were created by presenting a single monaural noise burst to either the left or the right ear. A monaural sound is clearly lateralised to one side or the other.

Ideally, for tight experimental control, auditory conflict would have been generated through an auditory equivalent of the flanker task. However, since concurrent acoustic signals tend to fuse and be perceived as a complex sound originating from a single source, it was not possible to produce a conflict effect using a target sound with concurrent flanking sounds. It was also not possible to separate the target and flanking sounds in time, since the initial sound would act as an alerting and orienting cue. For these reasons, verbal stimuli were selected, in which semantic information could conflict with the target pitch information. The auditory targets (Figure 2.2c) consisted of the words 'high', 'low' and 'day', spoken on either a high or low pitch by a female talker, i.e. with either a high or low fundamental frequency. The subject's task was to indicate with a button press whether the word was spoken on a high or low pitch. The incongruent condition consisted of the word 'high' spoken on a low pitch, and the word 'low' spoken on a high pitch. McClain (1983), using this design, found a significant interference effect of 105 ms with button press responses. Stimuli with matching word and pitch information formed the congruent condition, and the word 'day' spoken on either pitch formed the neutral condition. Stimuli were created by recording the spoken words, then digitising them at a sampling rate of 44,100 16-bit samples per second. Three examples of each stimulus were included, with all stimuli being presented in random order. The six groups of three examples of each word (on each pitch) had been chosen from a larger corpus to have approximately equal duration (average 457 ms; range: 403 ms - 507 ms) and intensity (average rms: -20.26 dB re full scale; range: 18.03 - 22.39). The high-pitched words had an average fundamental frequency of 290 Hz and the low-pitched words had an average fundamental frequency of 178 Hz, giving an average difference in fundamental frequency of 112 Hz. Fundamental frequency was measured as the frequency of the lowest harmonic calculated when a Fast Fourier Transform was applied to a 1024-sample Hanning windowed segment, selected from the central section of each word.

Colour-word Stroop Task

A classic colour-word Stroop task was included to provide a visual analogue for the auditory pitch-word Stroop task. The four colour words RED, BLUE, GREEN, and YELLOW were presented in either the colour they describe (congruent condition) or one of the other three colours (incongruent). The words LOT, SHIP, KNIFE, and FLOWER in any colour formed the neutral condition (Fan et al., 2003). A single letter subtended 0.58° in height and 0.49° in width. The gap between the letters subtended 0.08°. Stimuli were presented in a pseudo-random counterbalanced order, so that every condition followed every other condition equally often. Each trial began with a stimulus presented on screen, which remained there until the subject made a response. There was then a two-second pause, during which the fixation cross was visible, before the next stimulus appeared.

Subjects

Forty volunteers (19 male, mean age 23.7 years, range 16 - 42) participated in the study. Participants gave informed consent prior to the study, and were paid £25 for participation in a battery of tests, including the visual and auditory ANTs and colour-word Stroop task. Subjects all spoke English as their native language, and had normal or corrected-to-normal vision, as tested using a Snellen chart. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed 34 subjects with normal hearing (thresholds below 20 dB HL) and six subjects with thresholds no greater than 25 dB HL, which was considered acceptable for this study. Three additional participants were rejected: two for unacceptable audiograms, and one who consistently scored more than three standard deviations from the mean.

Apparatus and Stimuli

Testing was conducted in a sound-attenuating chamber. Stimuli were presented under the control of Visual Basic programmes implemented at the MRC Institute of Hearing Research on an IBM-compatible personal computer running Windows 98. Visual stimuli were presented on a 15-inch flatscreen monitor; auditory stimuli were presented via Sennheisser HD 480II headphones. All auditory stimuli were presented above threshold, in the range 70-80 dB(A). Participants viewed the screen from a distance of 65 cm. When auditory stimuli were presented the screen was uniformly grey. Subjects responded by pressing buttons on a response box situated on the desk in front of them. The buttons were arranged in a left to right horizontal array during the visual ANT and colour-word Stroop task, but turned sideways to provide a front to back vertical array for the auditory ANT. This arrangement aligned the physical location of the buttons with the intuitive locations of left/right and high/low responses.

Procedure

Subjects participated in the visual and auditory ANTs and the colour-word Stroop task, plus an additional study, in a single session. Subjects participated in two blocks of the visual ANT, two blocks of the auditory ANT, and one block of the Stroop task. These were presented using an ABBA counterbalance, with the Stroop task presented half way through the session. Each block of the ANT contained 144 trials presented in a random order. Prior to each auditory and visual block, subjects were provided with instructions and given a 24-trial practice session with feedback on whether they responded correctly. The Stroop task contained 288 trials presented in a counterbalanced order. Prior to the experiment, subjects were given instructions and a 32-trial practice session with feedback. During the experimental blocks, no feedback was provided. Subjects were instructed to respond as quickly and as accurately as possible. Each practice block lasted approximately two minutes, and each experimental block approximately eight minutes.

Results

Within-subjects t-tests were used to compare reaction times (RTs) and error rates with no cue versus a double cue (alerting effect) and a centre cue versus a spatial cue (orienting effect). One-way within-subjects ANOVAs were used to compare RTs and error rates in the incongruent, neutral, and congruent conditions (interference and facilitation effects). Only RTs from correct trials were analysed. RTs were trimmed to exclude responses quicker than 100 ms and slower than 2000 ms. This resulted in the removal of 1.1% of responses from the ANTs and 0.9% of responses on the Stroop task. Analyses were conducted on the median of the remaining RTs. The median was selected because RTs are not normally distributed. Planned comparisons were conducted using t-tests with a Bonferroni correction (dividing the critical p value by the number of comparisons being made). Group means and standard deviations are presented in Table 2.1 (visual ANT) and Table 2.2 (auditory

ANT). Alerting, orienting, and executive control RT and accuracy effects can be seen in Figures 2.3 and 2.4, respectively.

Table 2.1: Reaction time (RT) and accuracy data from the visual ANT.

	Warning Type			
Congruency	*None*	*Double*	*Centre*	*Spatial*
(a) Mean RTs (ms) and standard deviations:				
Incongruent	662 (101)	655 (98)	640 (94)	585 (95)
Neutral	546 (69)	521 (84)	522 (79)	486 (80)
Congruent	571 (86)	513 (86)	518 (75)	474 (65)
(b) Error rate (%) and standard deviations:				
Incongruent	5.10 (6.70)	7.81 (9.25)	5.31 (5.58)	3.96 (5.58)
Neutral	0.94 (1.76)	1.15 (2.67)	1.15 (2.11)	0.15 (2.67)
Congruent	0.31 (1.46)	0.63 (1.78)	0.73 (1.60)	0.52 (1.40)

Table 2.2: Reaction time (RT) and accuracy data from the auditory ANT.

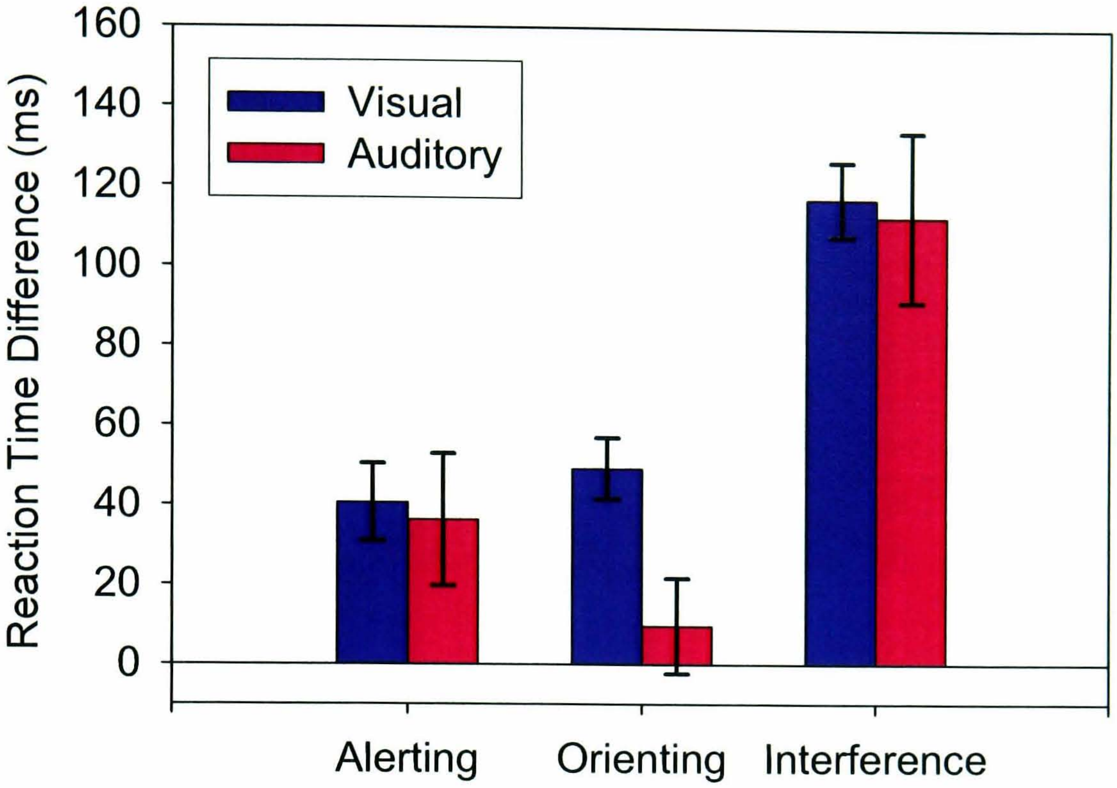
	Warning Type			
Congruency	*None*	*Double*	*Centre*	*Spatial*
(a) Mean RTs (ms) and standard deviations:				
Incongruent	780 (175)	747 (181)	756 (196)	751 (185)
Neutral	671 (159)	641 (164)	639 (161)	617 (140)
Congruent	641 (129)	603 (134)	602 (141)	603 (131)
(b) Error rate (%) and standard deviations:				
Incongruent	10.10 (9.04)	10.42 (8.39)	10.10 (10.71)	11.35 (10.29)
Neutral	3.96 (5.08)	2.50 (4.41)	2.60 (4.59)	2.92 (4.54)
Congruent	1.56 (2.78)	0.83 (1.69)	1.15 (2.67)	0.63 (1.51)

Alerting, Orienting, and Interference

In the visual condition, subjects were significantly faster when there was an alerting cue (mean difference between no cue and double cue conditions = 41 ms; $t_{39}=8.4$; $p<0.001$). This was also true in the auditory condition (36 ms; $t_{39}=4.4$; $p<0.001$). The overall percentage of errors in each test was very low (auditory: 4.8%; visual: 2.4%). In the visual condition, subjects made a significantly higher proportion of errors with a double cue than with no cue (mean difference = 1%; $t_{39}=-2.597$; $p<0.05$), indicating a speed-accuracy trade-off, while in the auditory condition there was no significant difference in error rates (mean difference = -1%, $t_{39}=1.148$; $p=0.258$).

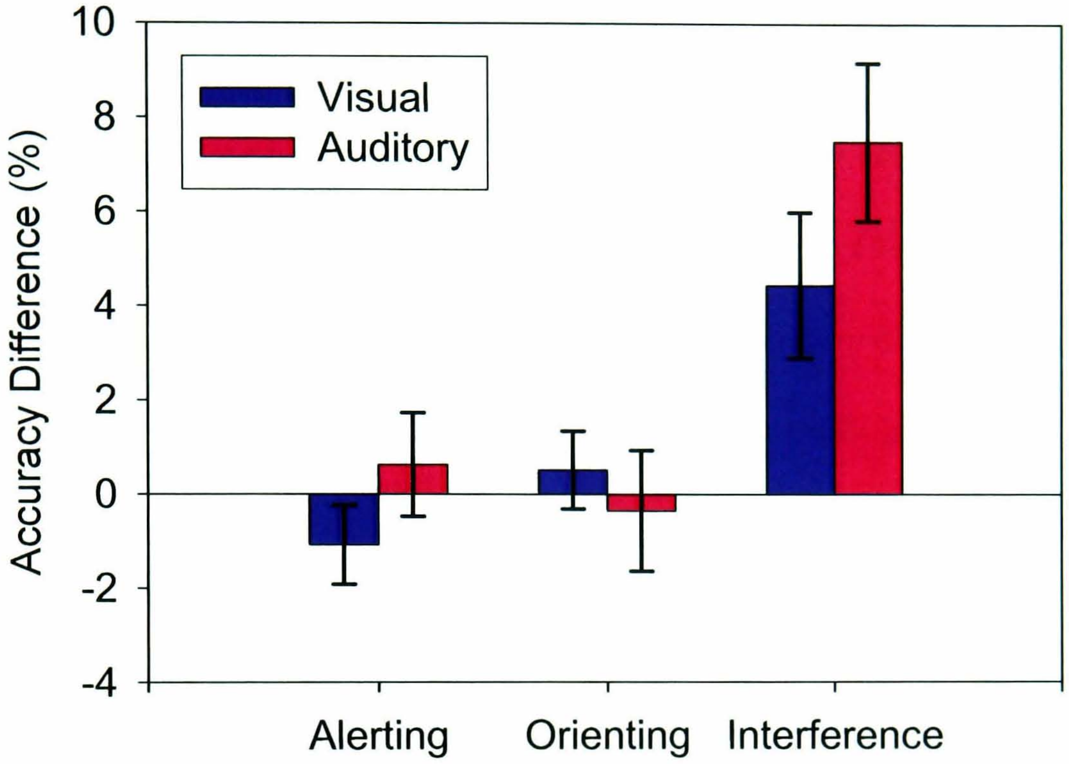
An orienting effect was found in the response times from the visual test (mean difference between single cue and spatial cue conditions = 49 ms; $t_{39}=12.8$; $p<0.001$). In the auditory condition, subjects were again faster, but the effect did not reach significance (mean difference = 10 ms; $t_{39}=1.6$; $p=0.110$). There were no significant differences in accuracy.

Figure 2.3: Reaction time effects for alerting (no cue - double cue), orienting (central cue - spatial cue) and interference (incongruent - neutral). Error bars show 95% confidence intervals.



One-way within subjects ANOVAs showed significant differences in RTs between stimulus conditions (incongruent, neutral, congruent) in the visual ($F_{1.3,52.1}$=450.2; $p<0.001$) and auditory ($F_{1.4,56.4}$=95.6; $p<0.001$) modalities. Planned comparisons ($p_{critical}$=0.05/2 = 0.025) on the RTs from the visual task showed significantly slower responses in the incongruent condition compared to the neutral condition (t_{39}=25.096, $p<0.001$), but there was no significant difference between the congruent and neutral conditions (t_{39}=0.450, p=0.655). In the auditory test responses were significantly slower in the incongruent condition compared with the neutral condition (t_{39}=10.713, $p<0.001$), and in the neutral condition compared to the congruent condition (t_{39}=4.094, $p<0.001$). The interference effects were of a similar size in both tasks (visual = 117 ms; auditory = 113 ms). The facilitation effect was small in the auditory task (32 ms), and not present in the visual task (1 ms).

Figure 2.4: Accuracy effects for alerting (no cue - double cue), orienting (central cue - spatial cue) and interference (incongruent - neutral). Error bars show 95% confidence intervals.



One-way within subjects ANOVAs conducted on the accuracy data also showed significant differences between stimulus conditions for the visual ($F_{1.1,43.5}=33.4$; $p<0.001$) and auditory ($F_{1.2,46.5}=66.8$; $p<0.001$) tasks. Planned comparisons on the error rates (with $p_{critical} = 0.05/2 = 0.025$) from the visual task showed that subjects made more errors in the incongruent condition than the neutral condition ($t_{39}=5.805$, $p<0.001$), but that there was no significant difference between the neutral and congruent conditions ($t_{39}=2.319$, $p=0.026$). The same tests on data from the auditory task showed that subjects made significantly more errors in the incongruent condition compared with the neutral condition ($t_{39}=9.018$, $p<0.001$), and more errors in the neutral condition than the congruent condition ($t_{39}=3.956$, $p<0.001$).

Correlations and Test Reliability

Having established effect sizes for each of the attention types in the visual and auditory tasks, analyses were then conducted to investigate correlations between RT measures of each type of attention, both within and between modalities (Table 2.3). There were no significant correlations between the three measures of attention, suggesting that they are independent of each other. Alerting and orienting scores did not correlate significantly across tasks, but auditory and visual interference costs were significantly correlated.

Table 2.3: Correlations between reaction time measures of alerting, orienting, and executive control (Exec.). *p<0.05, **p<0.01

	Auditory				*Visual*		
	Alert.	Orient.	Exec.	Mean	Alert.	Orient.	Exec.
Auditory							
Orient.	-.15						
Exec.	-.08	.11					
Mean	-.08	.29	.20				
Visual							
Alert.	.09	.02	.24	-.27			
Orient.	.03	.05	-.25	.04	-.06		
Exec.	-.19	.30	.33*	.35*	.16	-.14	
Mean	-.28	.26	.25	.76**	-.18	.23	.38*

The relative variability of the auditory and visual measures of each type of attention was also assessed. This showed significantly more variability across subjects for the auditory alerting ($t_{38}=3.39$, $p<0.05$), orienting ($t_{38}=2.76$, $p<0.05$), and interference ($t_{38}=5.90$, $p<0.05$) effects, relative to the visual effects.

A simple measure of test reliability was calculated by comparing performance on the first and second testing blocks. Alerting did not correlate significantly between blocks for the auditory ($r=0.12$, $p=0.449$) or visual ($r=0.17$, $p=0.301$) tasks. The orienting effect did not correlate significantly between blocks for the auditory task ($r=-0.11$, $p=0.518$), but there was a trend for the visual task ($r=0.29$, $p=0.073$). Interference correlated significantly across blocks for both the auditory ($r=0.34$, $p<0.05$) and visual ($r=0.44$, $p<0.01$) tasks. Overall reaction times were significantly correlated across blocks for both tasks (auditory: $r=0.87$, $p<0.001$; visual: $r=0.68$, $p<0.001$).

Interactions between the Attention Networks

To investigate any interaction between alerting, orienting, and interference, 4 (cue condition: no cue, centre cue, double cue, spatial cue) x 3 (stimulus condition: incongruent, neutral, congruent) ANOVAs were conducted on the RT and accuracy data from both the auditory and visual tests. Where Mauchley's test of sphericity indicated that sphericity could not be assumed, a Greenhouse-Geisser correction was applied. This is evident from non-integer degrees of freedom.

The RT data from the visual ANT showed a significant main effect of cue type ($F_{2.2,86.2}$=133.458, $p<0.001$), a significant main effect of stimulus type ($F_{1.3,49.8}$=441.478, $p<0.001$), and a significant interaction ($F_{4.9,190.2}$=10.614, $p<0.001$). Three one-way ANOVAs were then conducted to examine the influence of stimulus type (incongruent, neutral, congruent) on the size of the alerting and orienting measures, and the influence of cue type (no, double, centre, spatial) on the size of the interference effect. A Bonferroni correction was applied, giving a corrected p value of 0.05 / 3 = 0.0167.

Alerting varied significantly as a function of stimulus type ($F_{2,78}$=23.326, $p<0.001$). Paired t-tests, with a Bonferroni correction ($p_{critical}$=0.05/3 = 0.0167) showed that the alerting effect was significantly larger with congruent stimuli than with neutral stimuli (t_{39}=5.192, $p<0.001$) or with incongruent stimuli (t_{39}=5.946, $p<0.001$). Orienting did not vary significantly as a function of stimulus type ($F_{2,78}$=3.562, p=0.033). Interference varied significantly as a function of cue type ($F_{3,117}$=6.400, $p<0.001$). Paired t-tests, with a Bonferroni correction ($p_{critical}$=0.05/6 = 0.00833) showed that the interference effect was significantly greater with a double cue than with a spatial cue (t_{39}=4.270, $p<0.001$).

The ANOVA conducted on the accuracy data from the visual ANT showed the same pattern of results: a significant main effect of cue type ($F_{2.5,98.7}$=3.947, $p<0.05$), a significant main effect of stimulus type ($F_{1.1,43.5}$=33.370, $p<0.001$), and a significant interaction ($F_{3.4,130.9}$=3.574, $p<0.05$). Planned comparisons were conducted as before. Neither alerting ($F_{1.2,46.6}$=3.614, p=0.056) nor orienting ($F_{1.3,51.5}$=1.552, p=0.218) varied as a function of stimulus type. However, interference did vary as a function of cue type ($F_{3,117}$=3.651, p<0.0167). As with the RT data, paired t tests (with $p_{critical}$=0.05/6 = 0.00833) showed a significantly greater interference effect with a double cue than with a spatial cue (t_{39}=3.124, p<0.00833).

Reaction time data from the auditory ANT showed significant main effects of cue type ($F_{2.5,96.0}$=15.251, p<0.001) and stimulus type ($F_{1.4,\ 53.2}$=112.634, p<0.001), but the interaction was not significant ($F_{4.6,179.6}$=0.958, p=0.440). The accuracy data showed no main effect of cue type ($F_{3,117}$=0.547, p=0.651), a significant main effect of stimulus type ($F_{1.2,46.5}$=66.768, p<0.001), but again the interaction was not significant ($F_{3.3,130.5}$=0.813, p=0.500).

Colour-word Stroop Results

A one-way ANOVA showed a significant difference in RTs between conditions (incongruent, neutral, congruent) in the Stroop task ($F_{1.7,66.2}$=95.724, p<0.001). Planned comparisons ($p_{critical}$=0.05/2 = 0.025) showed significantly slower responses in the incongruent condition compared with the neutral condition (t_{39}=8.640, p<0.001) and significantly faster responses in the congruent condition compared to the neutral condition (t_{39}=6.466, p<0.001). There were no significant differences in accuracy across conditions ($F_{1.7,67.2}$=2.854, p=0.072). The size of the RT interference effect was smaller in the Stroop task (66 ms) than the visual (117 ms) or auditory (113 ms) ANT.

However, there was a trend towards correlations between interference on the Stroop task and interference on the auditory ANT (r=0.280, p=0.081) and visual ANT (r=0.298, p=0.062).

Discussion

The re-implementation of the visual ANT produced similar results to those reported by Fan et al. (2002). There are significant reaction time effects of visual alerting, orienting, and interference, of a similar magnitude to those found by Fan et al. (2002) (alerting: 41 ms in this study vs. 47 ms previously; orienting: 49 ms vs. 51 ms; interference: 117 ms vs. 84 ms, although note that Fan et al. compared incongruent and congruent conditions, while here the comparison is between incongruent and neutral conditions). However, while the auditory ANT revealed significant alerting (36 ms) and interference (113 ms) effects, there was no significant benefit from an orienting cue (10 ms). Moreover, measures of all three types of attention were more variable across subjects in the auditory task than in the visual task. Interference was significantly correlated between the auditory and visual tasks, but the alerting and orienting measures were not correlated between tasks.

Consistent with the results found in the Fan et al. study, there were no significant correlations between the different measures of attention within each test. However, only the interference measure was significantly correlated between the two blocks of testing, suggesting that the alerting and orienting measures may not be highly robust in terms of test-retest reliability. Test-retest reliability was also measured in the original ANT study, in which all three measures correlated significantly across two sessions. Each session in the original study was longer than the two test blocks on which the current analysis is based, and therefore may be a more powerful measure of test reliability.

As in the original study, there was a significant interaction between cue type and stimulus type in the visual task (although this was not present in the auditory task). The interaction arises from an increased alerting effect with congruent stimuli compared with neutral or incongruent stimuli, and increased interference with a double cue relative to a spatial cue. The original study found that in the no cue and spatial cue conditions the interference effect was reduced, which is broadly in line with the results found here.

Alerting

Neuroimaging research has provided evidence that sustained attention may be a supramodal facility (Kinomura et al., 1996; Pardo et al., 1991). In addition, Sturm and Willmes (2001) showed similar patterns of activation when auditory targets were cued using auditory and visual warning cues. Consistent with this evidence, we found significant benefits from an alerting cue in both the auditory and visual tasks, of a similar magnitude in both (41 ms in the visual task and 36 ms in the auditory task). However, the effect was significantly more variable across subjects in the auditory task (standard deviation: 31 ms in the visual task; 52 ms in the auditory task). In addition, the size of the alerting effect did not correlate significantly between the auditory and visual tasks.

There was also no correlation between the size of the alerting effect in the two testing bocks, for either the visual or auditory task, suggesting that this measure may not be highly robust.

Orienting

Significant benefits from an orienting cue were found in the visual ANT (49 ms) but not the auditory ANT (10 ms). Again, there was much more variability across subjects in the auditory test than in the visual test (standard deviation: 24 ms in the visual task; 37 ms in the auditory task). This result is contrary to the spatial relevance hypothesis (McDonald & Ward, 1999), which predicted that auditory spatial orienting effects would be found with a non-spatial task in which cues were informative about target location. However, perhaps this effect is not entirely unexpected given that auditory information is not coded spatiotopically in the same way as visual or somatosensory information (Middlebrooks, 2000). While spatial information may be critical for visual information processing (Tsal & Lamy, 2000), in the auditory system other features may be more critical, such as the temporal and spectral characteristics of the sound, and this difference may be reflected in the usefulness of spatial cues in the two modalities. As with the alerting measure, there was no correlation between the size of the orienting effect on the two testing blocks.

Interference

Significant interference effects were found in both the visual (117 ms) and auditory (113 ms) ANTs. These effects correlated significantly (r=0.33) across tasks, although the auditory effect was more variable across subjects than the visual measure (standard deviation: 30 ms in the visual test; 67 ms in the auditory test). There is very little previous research on which to base predictions about how executive control might operate across sensory modalities. The interference effect measured by the auditory and visual ANTs correlated significantly between the studies, and also between testing sessions, indicating that it is a reliable measure of this attentional skill. It is interesting to note that interference is the measure which correlated most highly across test sessions in the original ANT, and that it is the skill which was found to be influenced by genetic variation (Fossella et al., 2002). It may be that executive control is a highly robust attentional skill relative to alerting and orienting.

The colour-word Stroop task produced a smaller interference effect (66 ms) than the conflict tasks in the visual and auditory ANTs. However, there was a trend towards significant correlations between the measure of colour-word Stroop interference and the measures of interference from the ANT. The cueing conditions in the ANTs may have introduced greater variability into the ANT interference measures, which were not present in the colour-word Stroop task, and this may account for the non-significance in the correlations between these tasks.

Future Studies

The design and procedure of the auditory ANT were intended to be as similar as possible to those of the visual ANT. However, the failure of the auditory ANT to elicit an orienting effect, and the increased variability in the auditory measures compared with the visual measures, suggests that this may not be the ideal strategy. There are four key changes which would be interesting to implement. First, in an attempt to faithfully adapt the visual ANT, a fixation tone was used in place of a fixation cross. However, the fixation tone may have acted to distract subjects from the task, or to cue them to an incorrect location by acting as a spatial cue. Removing the fixation tone may give more accurate measures of subjects' ability to alert and to orient. Second, the time course of alerting and orienting may not be comparable across modalities. It is possible that auditory alerting and orienting effects may be more substantial over a shorter or longer time course. This question could be addressed by comparing performance over different cue-target intervals. Third, it was necessary to use lexical stimuli in the auditory ANT in order to elicit an interference effect. However, subjects may have been able to alert and to orient to the initial consonant sound of the stimulus, before sufficient information was available with which to make a response. This may have led to an underestimation of subjects' abilities to alert and orient in the auditory domain. Using steady-state target stimuli might elicit more reliable measures of alerting and orienting since all information on which to base a response would be available from target onset. Finally, in line with the visual ANT, the auditory ANT used spatial orienting cues. Given that auditory information is initially coded tonotopically, it is possible that cueing target frequency may be more efficient than cueing target location. For example, Scharf, Quigley, Aoki, Peachey and Reeves (1987) reported that sounds presented at an expected frequency were detected more often than sounds presented at unexpected frequencies, and Mondor and Bregman (1994) found that valid cues to the frequency of a target facilitated judgments of target duration.

Interference generated by the auditory and visual tasks was of a similar magnitude in both tasks, and correlated significantly between them. It is now interesting to ask whether this similar behavioural effect is the result of common cortical mechanisms. This can be addressed using fMRI to investigate whether the same cortical areas are active when subjects perform auditory and visual conflict-resolution tasks.

Conclusions

A comparison of performance on auditory and visual ANTs revealed greater variability in measures of auditory alerting, orienting, and executive control, compared with visual measures of the same skills. Spatial orienting effects were substantially different in the two tasks, although the nature of the difference is not clear from this study. Executive control appears to be unaffected by stimulus modality, which may reflect supramodal components of conflict resolution. Additional behavioural studies would be beneficial in furthering understanding of auditory orienting of attention, while fMRI studies may be able to identify common cortical mechanisms involved in conflict monitoring and resolution.

Chapter 3: Neural Correlates of Auditory and Visual Conflict Resolution

Introduction

A comparison of attentional skills measured by the auditory and visual attention network tests (ANTs) revealed similarities between measures of auditory and visual executive control (Chapter 2). Reaction time (RT) interference effects elicited by incongruent compared with neutral stimuli in the auditory and visual tasks were of a similar magnitude and significantly correlated with each other. To investigate whether this similar behavioural effect is the result of common cortical mechanisms, the study reported in this chapter used fMRI to investigate cortical activity during performance on a visual colour-word Stroop task and an auditory pitch-word Stroop task. The results reported in Chapter 2 show that interference elicited by the pitch-word Stroop task was more highly correlated with interference from the flanker task than the colour-word Stroop task. However, all three tasks elicited reliable interference effects, and any differences between them may have simply reflected additional variability introduced by the cueing conditions in the flanker and pitch-word Stroop tasks. The colour-word and pitch-word Stroop tasks are theoretically well matched in that they both require subjects to respond to a relevant non-linguistic component (colour / pitch identification) while ignoring an irrelevant linguistic component (word meaning). The colour-word Stroop task is therefore a closer analogue of the pitch-word Stroop task than the flanker task, and comparisons between the two can be used to identify modality-independent and modality-specific activation.

Conflict arises when a stimulus gives rise to more than one possible response, requiring inhibition of the incorrect response(s) in order to respond correctly. The colour-word Stroop task is the classic paradigm used to elicit conflict. In this task, subjects typically view colour words printed in colours which are congruent (e.g. 'RED' in red) or incongruent (e.g. 'RED' in blue). A neutral condition is provided by presenting non-colour words, or, less commonly, non-words or rows of symbols. The subject's task is to report the colour the word is printed in, while ignoring the competing semantic information. Compared with neutral stimuli, subjects are slower to respond to incongruent stimuli, and faster to respond to congruent stimuli. The two effects are asymmetric, with the interference created by incongruent stimuli being larger than the facilitation created by congruent stimuli (see MacLeod (1991) for a review of Stroop studies). Other conflict tasks include the flanker task (Eriksen & Eriksen, 1974), where flanking information interferes with processing of central, target information, and spatial conflict tasks where, for example, right hand responses are made to left-hand stimuli, otherwise known as the 'Simon' task (e.g. Simon & Berbaum, 1990). Stroop, flanker, and Simon tasks belong to a family of tasks which require some form of conflict resolution. In a review of these studies, Nee, Jonides, and Wager (2004) identified four types of interference, and classified tasks based on which of these types were present. The four types were: 1) irrelevant stimulus conflict, 2) response selection conflict, 3) response execution conflict, and 4) response mapping conflict. Stroop, flanker, and

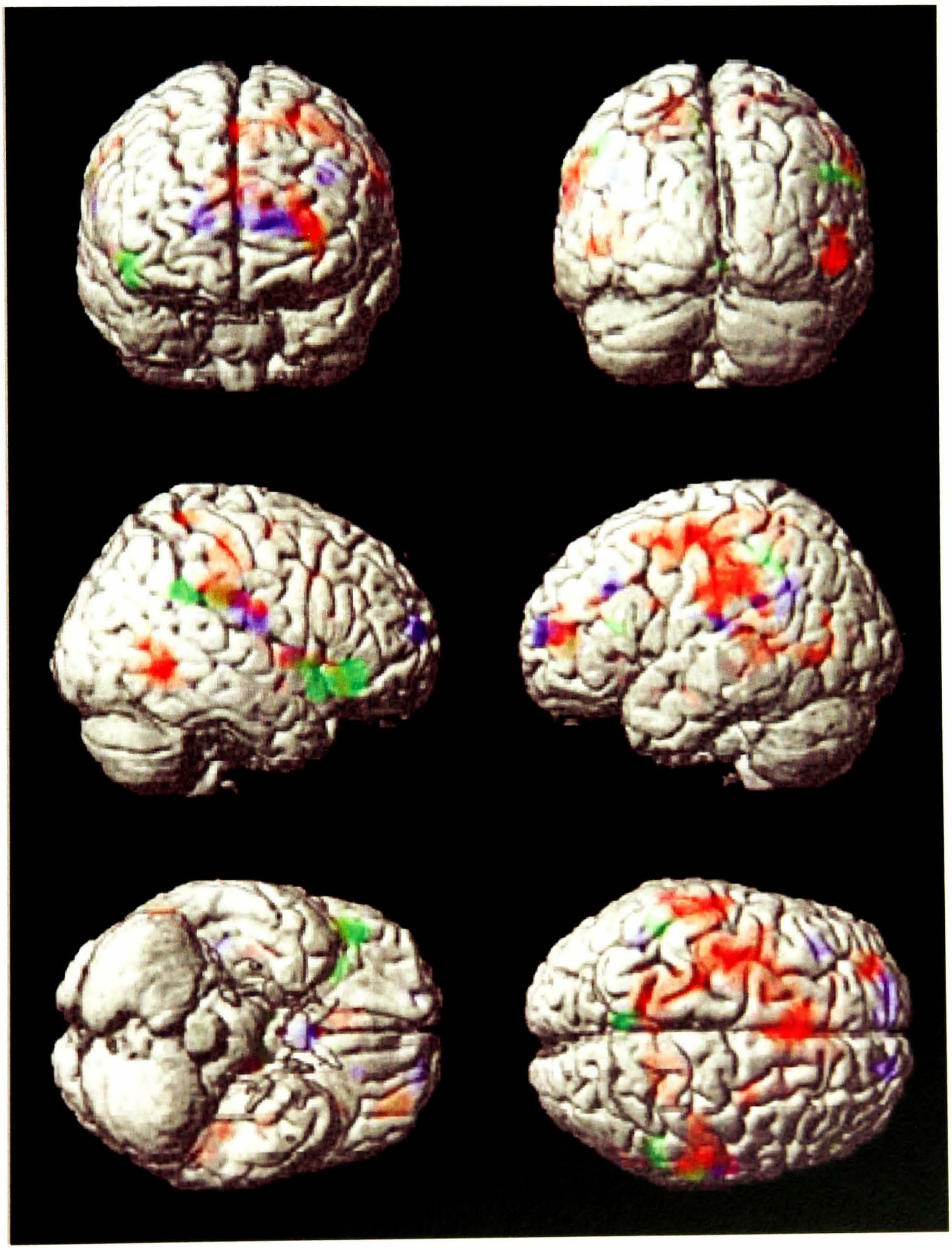
Simon tasks all include irrelevant stimulus conflict and response selection conflict. Alternative types of conflict task show a different pattern of interference. For example, a 'go/no-go' task, in which subjects must inhibit responding to certain stimuli, involves response selection conflict and response execution conflict.

Imaging studies investigating cortical activation associated with conflict resolution have identified an anterior network which incorporates the anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (DLPFC) (Smith & Jonides, 1999). However, a range of additional areas have also been found to be involved, and these vary from study to study. Some consistently reported areas are the frontal polar cortex (Bench et al., 1993), inferior temporal gyrus (Bush et al., 1998; Carter, Mintun, & Cohen, 1995), superior parietal lobe (Bush et al., 1998) and inferior parietal lobe (Carter et al., 1995; George et al., 1994).

While ACC and DLPFC are frequently identified as being involved in conflict resolution their exact role is unclear. Several theories about the functional interpretation of ACC activation can be found in the literature. These identify two components of the conflict-resolution process: monitoring and active resolution. Two monitoring roles have been suggested for ACC: monitoring for conflict (Barch et al., 2000; Carter et al., 2000), and monitoring for errors (Gehring, Goss, Coles, Meyer, & Donchin, 1993). In contrast to these monitoring theories, ACC has also been hypothesised to have an active role in conflict resolution (Peterson et al., 1999; Posner & DiGirolamo, 1998). An intermediate theory is that ACC monitors for conflict and uses that information to activate other areas which actively resolve the conflict (Pardo, Pardo, Janer, & Raichle, 1990). The DLPFC is the main candidate for this active conflict resolution (Casey et al., 2000). The role of other areas which are active during conflict resolution is also unclear, although there is some evidence that superior parietal lobe activation may reflect selective attention processes (Casey et al., 2000).

In an attempt to identify a common network activated by conflict tasks, Fan et al. (2003) compared cortical activation associated with conflict arising from three different visual conflict tasks which were performed by a single group of subjects. The tasks were a colour-word Stroop task, a flanker task, and a spatial conflict task. The colour-word Stroop task and flanker task were as previously described. The spatial conflict task presented response labels to the bottom left and right of the screen prior to each trial. The target then appeared at the top left, middle, or right, and subjects were required to press the right or left button, whichever corresponded to the correct response label. In the congruent condition the target and response label were on the same side of the screen. In the neutral condition the target was presented centrally, and in the incongruent condition the target was presented on the opposite side to the response label. Fan et al. (2003) contrasted cortical activation associated with incongruent and congruent trials and found that all three tasks activated areas of ACC (BA 32) and prefrontal cortex (BA 10), as well as a number of areas specific to each task (Figure 3.1). Since each task activates a large number of cortical areas, and areas of overlap are relatively small, it is difficult to see the regions of common activation in Figure 3.1. All three tasks were successful in generating behavioural correlates of response conflict. However, when comparing reaction times effects both from this study and a larger study with 40 participants, Fan et al. failed to find significant correlations between measures of interference.

Figure 3.1: Surface maps showing activations from the flanker task (red), colour-word Stroop task (green) and spatial conflict task (blue). Taken from Fan et al. (2003).



Since the three tasks engage common cortical regions, it would be reasonable to expect that these regions might perform the same function across tasks. However, the failure to find correlations between the behavioural measures suggests otherwise. Fan et al. propose that this disjunction reflects the multiple operations that must be performed in response to conflict. They hypothesise that the areas common to all three conflict tasks are involved in monitoring conflict, while areas of different activation are responsible for actively resolving conflict, which may be a task-specific process. An alternative explanation is that some of the activation (either common or different) does not reflect interference, but instead represents the neural correlate of facilitation. This interpretation cannot be ruled out by the present data since the contrast is a subtraction of incongruent and congruent trials.

Meta Analysis

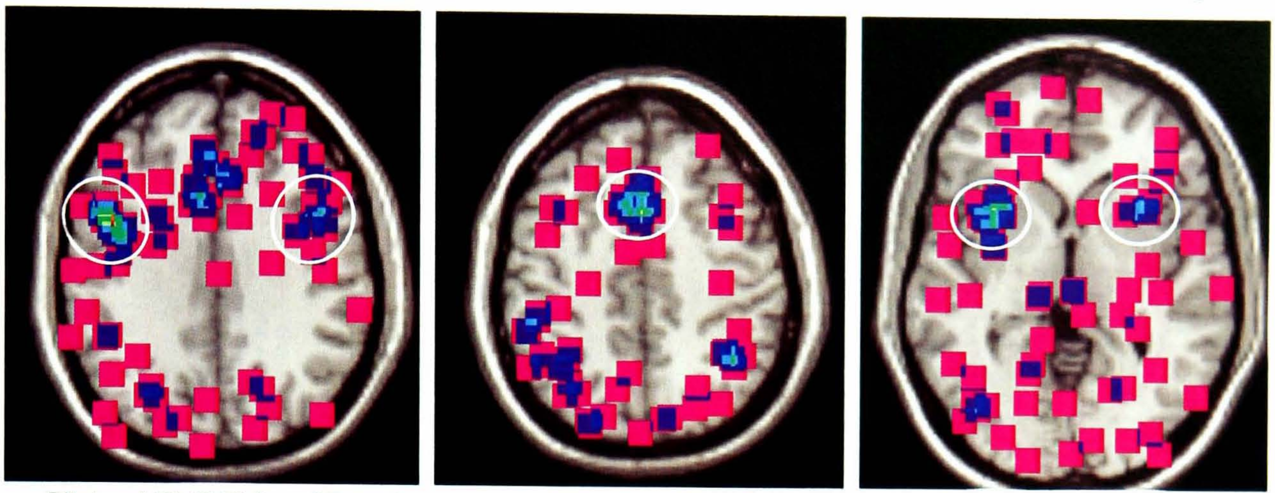
Given the wide range of brain areas activated by conflict tasks, a meta-analysis was conducted to identify regions reliably associated with conflict monitoring and resolution. Papers were identified through searches of ScienceDirect, Ingenta, PsycINFO, and the Proceedings of the National Academy of Sciences in the United States of America (PNAS) on August 5th 2004. The words ['Stroop' OR 'Simon' OR 'flanker'] AND ['fMRI' OR 'functional MRI' OR 'functional magnetic resonance imaging' OR 'PET' OR 'positron emission tomography'] were entered. ScienceDirect and Ingenta searched for these words in the article title, abstract and keywords, and PsycINFO and PNAS searched the article title and abstract. The searches identified 75 English-language journal articles available in the UK. This list of articles was then cross-checked against reviews by Nee et al. (2004), Cabeza and Nyberg (2000), Duncan & Owen (2000), Barch et al. (2001), and Bush et al. (1998). This revealed a further five articles which reported PET or fMRI measures of the Stroop, flanker or Simon tasks. The final 80 articles were then reviewed, and were included if they 1) reported a conflict generating task, 2) analysed their data with a simple incongruent - neutral or incongruent - congruent comparison, 3) reported results for normal adult subjects, including studies which reported results from a normal adult control group independently from those from a patient group, 4) reported a whole-brain analysis, and 5) published peak coordinates for activated regions. Thirty-four papers, reporting a total of forty experiments, met these criteria (Table 3.1). Bench et al. (1993) report data from two experiments. For the first of these, the incongruent condition was compared against two neutral baselines: neutral words and neutral crosses. To avoid using the same incongruent condition data twice, only the neutral word contrast has been included. Similarly, Carter et al. (1995), Mead et al. (2002), and Zysset, Müller, Lohrmann, and von Cramon (2001) report contrasts against both neutral and congruent baselines. Only the contrasts against the neutral baselines are included here. For their first experiment, Taylor, Kornblum, Lauber, Minoshima, and Koeppe (1997) report contrasts against neutral-word and false-font baselines, of which only the neutral-word contrast is included.

Peak coordinates were accepted as reported, irrespective of the statistical threshold or number of peaks reported per region. Between 1 and 40 peaks were reported for each experiment. In total, 406 peaks of activation are included. Coordinates reported in Talairach space (Talairach & Tournoux, 1988) were converted into MNI space (Evans et al., 1993) using Matthew Brett's Matlab routine (http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml). To account for the spread of activation around a peak, and spatial variations in normalisation across subjects and studies introduced by image transformations into standard brain space, each peak was extended for 5mm in each direction. This resulted in 11mm^3 cubes around each reported peak of activation. For each experiment, an 'activation' map was created in which each 1 x 1 x 1 mm voxel within these cubes was assigned a value of one, and all other voxels were assigned a value of zero. The 'activation' maps were then summed together to produce probability maps (Figure 3.2). One advantage of this technique is that each voxel can only be counted once from each study. This partly alleviates the problem of some studies reporting multiple peaks in the same region, while others report only one.

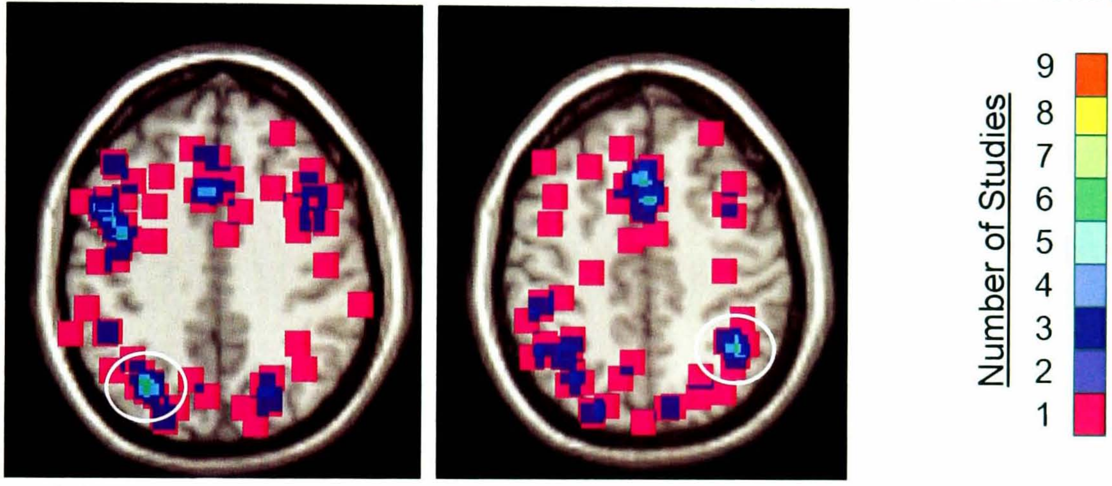
Table 3.1: Studies included in the meta-analysis, including task description, number of subjects (n), and number of reported peaks of activation (peaks).

	Study	Task	n	Peaks
1	Adleman et al. (2002)	Colour-word Stroop	11	3
2	Banich et al. (2000)	Colour-word Stroop	10	4
		Colour-object Stroop	10	3
3	Bench et al. (1993)	Colour-word Stroop	6	2
		Colour-word Stroop	6	5
4	Brass et al. (2001)	Inhibition of imitative resps.	10	5
5	Brown et al. (1999)	Colour-word Stroop	7	4
6	Bunge et al. (2002)	Flanker inc. go/no-go cond.	16	11
7	Bush et al. (1998)	Counting Stroop	9	7
8	Bush et al. (1999)	Counting Stroop	8	7
9	Carter et al. (1995)	Colour-word Stroop	15	6
10	Compton et al. (2003)	Colour-word Stroop	12	5
11	de Zubicaray et al. (2001)	Picture-word Conflict	8	9
12	Fan et al. (2003)	Flanker	12	14
		Colour-word Stroop	12	13
		Spatial Conflict	12	11
13	George et al. (1994)	Colour-word Stroop	21	10
14	George et al. (1997)	Colour-word Stroop	11	3
15	Hazeltine et al. (2000)	Flanker	8	4
16	Liu et al. (2004)	Simon	11	34
		Spatial Stroop	11	15
17	Maclin et al. (2001)	Spatial Conflict	8	5
18	Matthews et al. (2004)	Counting Stroop	18	5
19	Mead et al. (2002)	Colour-word Stroop	18	1
20	Milham et al. (2001)	Colour-word Stroop	16	7
21	Milham et al. (2002)	Colour-word Stroop	12	14
22	Norris et al. (2002)	Adapted colour-word Stroop	7	10
23	Pardo et al. (1990)	Colour-word Stroop	8	13
24	Peterson et al. (1999)	Colour-word Stroop	34	40
25	Potenza et al. (2003)	Colour-word Stroop	11	10
26	Ruff et al. (2001)	Colour-word Stroop	12	10
27	Steel et al. (2001)	Colour-word Stroop	7	26
28	Tamm et al. (2002)	Counting Stroop	14	3
29	Taylor et al. (1994)	S-R Compatibility	8	3
30	Taylor et al. (1997)	Colour-word Stroop	12	10
		Colour-word Stroop	6	10
31	Ullsperger et al. (2001)	Flanker	9	34
32	van Veen et al. (2001)	Flanker	12	8
33	Videbech et al. (2004)	Colour-word Stroop	46	13
34	Zysset et al. (2001)	Adapted colour-word Stroop	9	9

Figure 3.2: Probability maps showing areas most likely to be reported as peaks of activation in the meta analysis. Reported peaks were extended for 5mm in all directions, hence the squares. Areas in which more than 5 peaks overlap are circled in white. Coordinates are in MNI space.



Bilateral DLPFC (z = 35 mm) ACC (z = 47 mm) Bilateral Insula (z = 3 mm)



Left Parietal Lobe (z = 40 mm) Right Parietal Lobe (z = 46 mm)

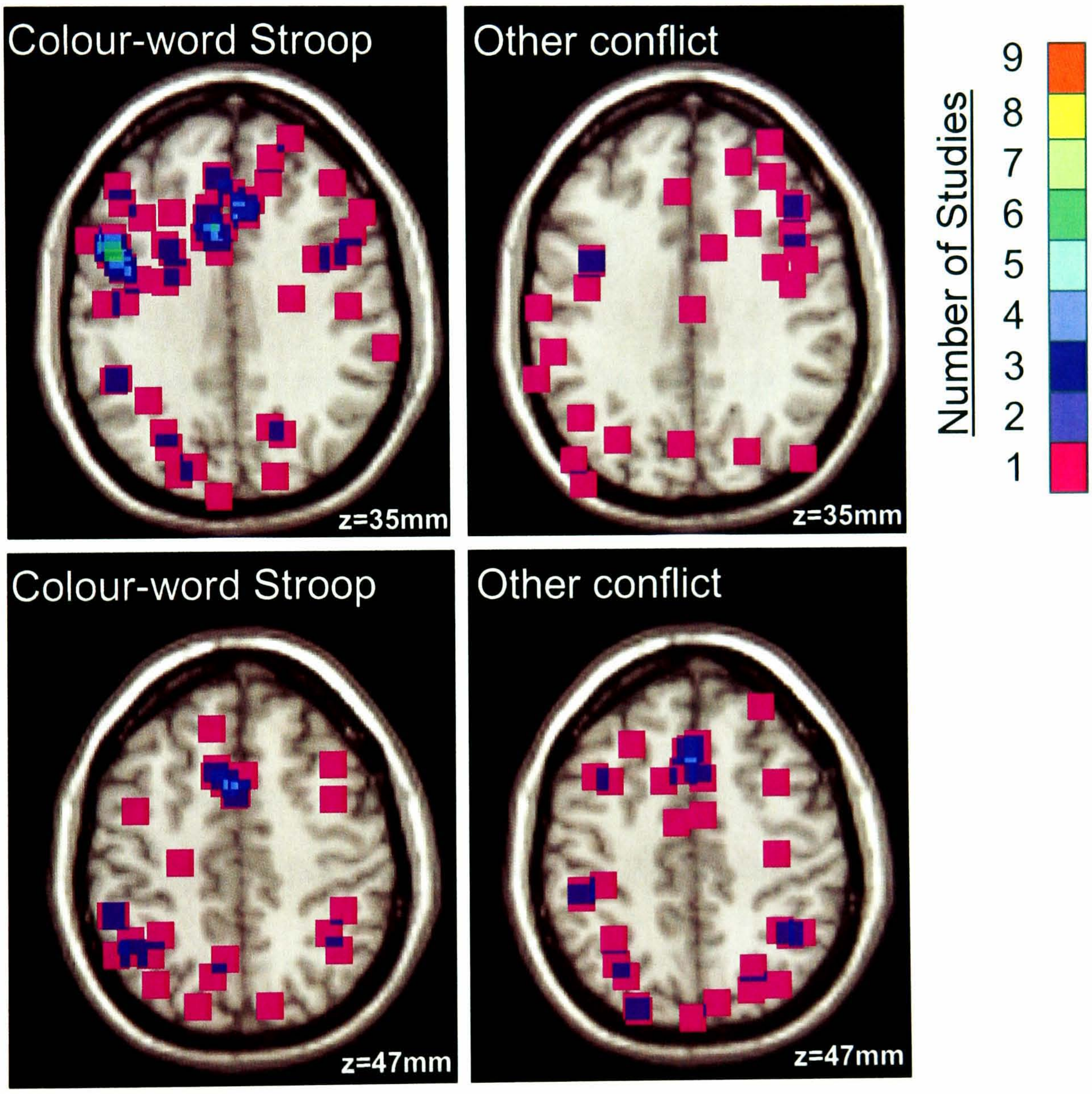
The probability maps show that activation is most reliably found in left dorsolateral prefrontal cortex (DLPFC). Around the centre of this cluster, nine studies report peaks of activation. Other clusters of activation resulting from the overlap of five or more peaks are right DLPFC, bilateral inferior frontal gyrus (insula), anterior cingulate cortex (ACC), and left and right parietal lobes. The approximate coordinates of the centre and extent of these clusters are reported in Table 3.2. While not all studies contribute to these clusters, there is still remarkable consistency considering differences in task demands, subjects, data processing techniques, normalisation templates and coordinate frames, and statistical thresholding.

Table 3.2: Meta analysis results: Extent of overlapping activation from five or more studies, and approximate peaks of overlap, in MNI coordinates.

Area	Max	Peak	Extent (x)			Extent (y)			Extent (z)		
L DLPFC	9	-43, 5, 35	-51	to	-37	-3	to	14	28	to	41
ACC	7	1, 12, 47	-8	to	5	11	to	27	35	to	52
L Insula	7	-31, 14, 1	-36	to	-26	6	to	16	0	to	5
L Parietal	6	-28, -67, 40	-31	to	-25	-71	to	-65	39	to	45
R Parietal	5	41, -51, 46	41	to	42	-53	to	-50	46	to	47
R DLPFC	5	48, 10, 31	46	to	51	8	to	12	27	to	35
R Insula	5	31, 13, 5	31	to	31	10	to	16	0	to	5

To investigate whether regions commonly activated by conflict tasks were task-dependent, the tasks were divided into those which used a colour-word Stroop task (n=21) and those which used a different type of conflict task (n=19). The probability maps for each of group separately are shown in Figure 3.3. Of the nine studies which contributed to the left DLPFC activation, seven were colour-word Stroop tasks. The remaining two were adapted colour-word Stroop tasks (Norris, Zysset, Mildner, & Wiggins, 2002; Zysset et al., 2001) in which subjects were shown two rows of letters and decided whether the colour of the top row of letters corresponded to the colour name on the bottom row. The colour name on the bottom row was always written in black ink. The word on the top row was either 'XXXX' (neutral condition), a colour name in a congruent colour ink, or a colour name in an incongruent colour ink. Although the task differs from a traditional colour-word Stroop task, it still requires the subject to inhibit a prepotent response generated by fast and automatic word reading.

Figure 3.3: Probability maps for colour-word Stroop tasks (n = 21) and other conflict tasks (n = 19) show differential activation in DLPFC (top row) but similar activation in ACC (bottom row).



One possible interpretation of the task-specific activation in left DLPFC is that there is greater consistency between colour-word Stroop tasks than between 'other conflict' tasks, which by definition involve varying stimuli and task demands. However, ACC activation is fairly equally contributed to by both groups of studies. Of the seven studies which contribute to the peak, four were from colour-word Stroop tasks and three from other conflict tasks. This provides tentative evidence that the ACC is involved in a general process required during conflict tasks, such as monitoring for conflict, while the DLPFC is involved in a task-specific process, such as conflict resolution relating specifically to inhibiting a prepotent response arising from automatic word reading.

An fMRI study of conflict resolution

The study reported here investigates the commonalities of the conflict resolution network by comparing behavioural and cortical correlates of conflict resolution across tasks. This study is novel in two ways. First, conflict resolution is measured in visual and auditory tasks which are otherwise well matched. Second, incongruent, neutral, and congruent trials are presented so that interference can be separated from facilitation. The behavioural study reported in Chapter 2 demonstrated significant correlations between auditory and visual measures of conflict resolution, suggesting that conflict resolution may be a supramodal facility. However, data from Fan et al. (2003), and from the meta analysis reported here, suggest task-specificity in some cortical areas active during conflict resolution. It is hypothesised that the close similarity between the pitch-word and colour-word Stroop tasks will result in a clearer assessment of the commonalities of the conflict-resolution network. Both tasks contain a relevant non-linguistic component (colour / pitch identification) and an irrelevant linguistic component (word meaning). Presenting tests in different modalities should aid identification of task-specific activation as it will relate to modality differences as well as task differences. In addition, comparing cortical correlates of auditory and visual conflict resolution should allow supramodal and intramodal areas of the conflict resolution network to be identified. A subset of 16 of the subjects that took part in the ANT experiment (Chapter 2) returned to take part in the fMRI study. Data are reported from the original ANT study, and from behavioural and neuroimaging measures taken during the fMRI study.

Method

Subjects

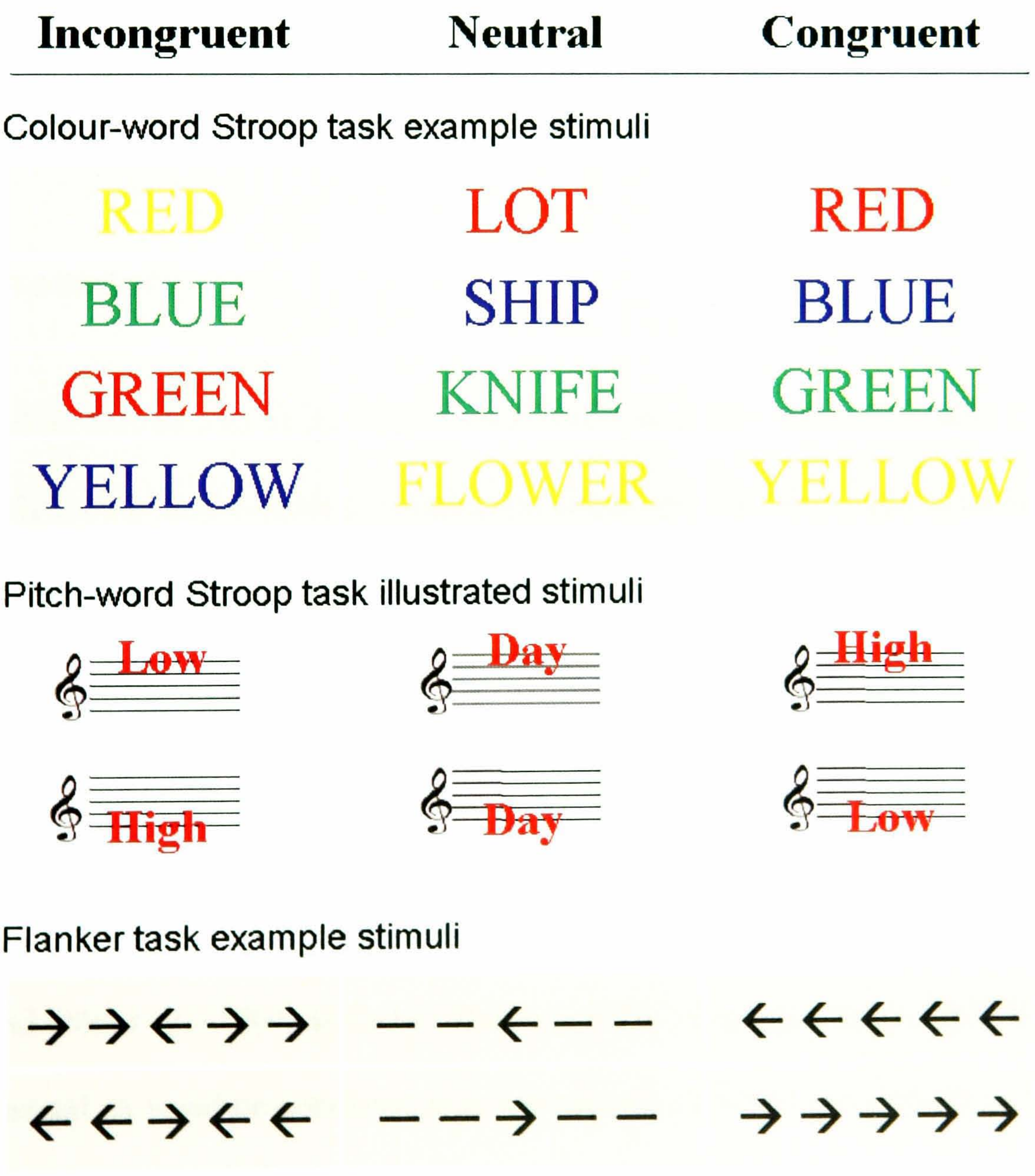
Behavioural data (Chapter 2) were collected from 40 healthy volunteers (19 male, mean age 23.7 years, range 16-42). Sixteen of these subjects (9 male, mean age 24.3 years, range 16-42) then returned to take part in the fMRI study. Subjects were right-handed native English speakers, with normal or corrected-to-normal vision. Pure tone air-conduction audiometry conducted at frequencies between 250 Hz and 8000 Hz, inclusive, confirmed that all subjects had thresholds lower than 25 dB HL. Participants gave informed consent prior to the study and were paid for their time. Six additional subjects took part in the fMRI study but were rejected for moving a distance that exceeded two voxels during a single task.

Apparatus and Stimuli

The behavioural study was conducted in a sound-attenuating chamber. Visual stimuli were presented on a flat-screen monitor, while auditory stimuli were presented via Sennheisser HD 480II headphones. During the imaging study subjects wore prism goggles to enable them to see a projector screen positioned at the end of the scanner bed. Images were back-projected onto the screen. Subjects wore ear plugs and ear defenders. Specially modified electrostatic transducers were fitted into the ear defenders to present the sound stimuli and to enable the experimenter to communicate with the subject (Palmer, Bullock, & Chambers, 1998). During both studies subjects responded by pressing one of four buttons on a response box. In-house software was used to present the stimuli and to record response times and accuracy.

Visual stimuli were created following Fan et al. (2003; 2002) and presented on a light grey background (Figure 3.4). Flanker stimuli comprised a central arrow pointing to the left or the right, with two flanking elements on either side. These elements could be arrows or straight lines. A single arrow subtended 0.55° of visual angle, and an entire stimulus (target arrow plus four flankers) subtended 3.08° of visual angle. Colour-word Stroop stimuli comprised a word ('RED', 'BLUE', 'GREEN', 'YELLOW', 'LOT', 'SHIP', 'KNIFE', or 'FLOWER') presented in one of four colours (red, blue, green, or yellow). A single letter of the colour-word Stroop stimuli subtended 0.58° of visual angle in height and 0.49° in width. The gap between letters subtended 0.08° of visual angle. Auditory stimuli were created by recording a female talker saying the words 'HIGH', 'LOW', and 'DAY' on a high and low pitch. These were then digitised at a sampling rate of 44,100 16-bit samples per second. Three examples of each stimulus were included. The six groups of three examples of each word (on each pitch) had been chosen from a larger corpus to have approximately equal duration (average 457 ms; range: 403-507 ms) and intensity (average rms: -20.26 dB re full scale; range 18.03-22.39). The high-pitched words had an average fundamental frequency of 290 Hz and the low-pitched words had an average fundamental frequency of 112 Hz.

Figure 3.4: Example stimuli from the colour-word Stroop, pitch-word Stroop, and flanker tasks.



Imaging Protocol

MR imaging was performed on a dedicated echo-planar imaging (EPI) 3-Tesla scanner with purpose-built head gradient coils (Bowtell & Peters, 1999) and a TEM head coil (Nova Medical Inc.). Functional images were collected using a T2*-weighted sequence (flip angle 90°; in-plane resolution 3 x 3 mm; 128 x 64 matrix; TR=2.992s). Twenty-two contiguous coronal slices, 8 mm thick, were acquired. For the majority of subjects this volume covered the whole head, but for some the occipital pole fell outside the field of view. One full head 64 slice T2*-weighted image at thinner slice thickness (4 mm; in-plane resolution: 3 x 3 mm) was collected to facilitate normalisation.

Procedure

Subjects took part in the behavioural study, and then returned to take part in the fMRI study at a later date. During the behavioural study, subjects completed the colour-word Stroop task, and cued versions of the flanker task and pitch-word Stroop task. The colour-word Stroop task stimuli were presented in pseudo-random order, so that every condition followed every other condition equally often. The flanker and pitch-word Stroop task stimuli were presented in random order. For the fMRI study, subjects took part in two tasks (pitch-word and colour-word Stroop tasks) within a single imaging session. Both tasks were uncued. A baseline condition was introduced, in which subjects either saw a briefly presented cross, or heard a 100-ms 1000-Hz tone. To improve sensitivity, the fMRI study was blocked, with each 24-second block comprising 14 stimuli. Forty-four blocks were presented (11 each of incongruent, neutral, congruent and baseline) in pseudo-random order, counterbalanced so that every condition followed every other condition equally often. The order of the two tasks was counterbalanced across subjects.

Behavioural Data Analysis

One-way within-subjects ANOVAs were used to analyse the median reaction times (RTs) and accuracy in the incongruent, neutral, and congruent conditions. Only RTs from correct trials were analysed. RTs from the behavioural study were trimmed to exclude responses quicker than 100 ms and slower than 2000 ms. This resulted in removal of 1.1% of responses from the ANT study with all 40 subjects and 0.6% with the subset of 16 subjects. Figures for the behavioural Stroop task are 0.90% and 0.73% respectively. During the fMRI study 14 stimuli were presented evenly over a 24-second block. If subjects did not respond within the allotted time for a stimulus (1.7 seconds) their response was not accepted. Analyses were conducted on the median of the remaining RTs. Where the assumption of sphericity was violated, a Greenhouse-Geisser correction was applied. This is evident from non-integer degrees of freedom. Planned comparisons were conducted using t-tests with a Bonferroni correction (dividing the critical p value by the number of comparisons being made).

fMRI Data Analysis

Image analysis was performed on a Sun Ultra 2 computer (Sun Microsystems) using SPM99 software (Friston, Holmes et al., 1995: http://www.fil.ion.ucl.ac.uk/spm) running in MATLAB v6.5 (Mathworks Inc., Natick, MA, USA). The first two scans were acquired before the onset of the first stimulus epoch and discarded to allow for steady-state saturation. For each subject, the remaining 352 scans were realigned to the last scan of the first task, and the first scan of the second task to correct for 3-dimensional movement of the head between image acquisitions. Realignment involved determining the values for a 6-parameter, rigid-body, affine transformation that minimised the sum of squared differences between the reference scan and each of the scans in the experimental sequence (Friston, Ashburner et al., 1995). Low-frequency artefacts, corresponding to aliased respiratory and cardiac effects and other cyclical variations in signal intensity, were removed by high-pass filtering the time series using cosine basis functions up to a maximum frequency of half a cycle per minute.

Realigned images were normalised by spatially transforming the realigned images into a standard brain space and re-sampling voxels to a size of 3 x 3 x 4 mm. This procedure involved a nine-parameter affine transformation, followed by non-linear deformations using discrete cosine transform basis functions that matched the 64-slice scan to an EPI template using a least-squares algorithm (Friston, Ashburner et al., 1995). These parameters were then applied to each of the coregistered scans in the functional data. The brain template was defined in a space specified by the Montreal Neurological Institute (Evans et al., 1993). The template was modified to match the signal loss in the inferior temporal cortex found during image acquisition at 3-Tesla. The normalised scans were spatially smoothed using a Gaussian kernel of 10 mm full-width-at-half-maximum.

For each subject, the three conditions were modelled using three box-car regressors where the box car was convolved with the haemodynamic response function (with a delay of six seconds). Six additional regressors were introduced to model movement-correlated changes in the signal, as variables of no interest. As the interscan interval was shorter than the haemodynamic response we used an AR(1) to account for temporal autocorrelation in the data. Data from individual subjects were then entered into higher-level random effects analyses with an activation threshold of $p<0.001$ uncorrected and an extent threshold of 14 voxels (0.5 cm^3). Conjunction analysis to identify areas of common activation was conducted using SPM2 software (http://www.fil.ion.ucl.ac.uk/spm/spm2.html) using the conjunction null hypothesis (Nichols, Brett, Andersson, Wager, & Poline, 2005). This tests the null hypothesis of no significant activation in contrast A OR contrast B, and therefore identifies areas which are significantly active in both contrast A AND contrast B.

Results

Behavioural Results

Median reaction times (RTs) on correct trials and proportion of correct responses (Table 3.3) were analysed using one-way within subjects ANOVAs and planned contrasts with a Bonferroni correction for multiple comparisons (Table 3.4). The planned contrasts investigate two effects: i) interference, where responses are expected to be slower and less accurate to incongruent targets than to neutral targets, and ii) facilitation, where responses are expected to be faster and more accurate to congruent targets than to neutral targets. Three datasets are analysed: from the original behavioural study (reported in Chapter 2), data are analysed for the full 40-subject sample, and for the subset of 16 subjects who went on to take part in the fMRI study. In addition, behavioural data were collected during the fMRI study. RT interference (incongruent - neutral) effects for all three groups are shown in Figure 3.5.

Behavioural study (n=40 - as reported in Chapter 2)

There were significant differences in RTs across conditions for the flanker task, pitch-word Stroop task, and colour-word Stroop task (Table 3.4). Planned contrasts (with $p_{critical}=0.05/2 = 0.025$) showed that there was significant interference from incongruent trials compared with neutral trials for all three tasks, but a facilitation effect (neutral versus congruent) was only found in the pitch-word and colour-word Stroop tasks.

Accuracy also differed significantly across conditions for the flanker task and pitch-word Stroop task, but not the colour-word Stroop task. Relative to the neutral condition, planned contrasts showed decreased accuracy for the incongruent condition for the flanker task and the pitch-word Stroop task, and increased accuracy for the congruent condition for just the pitch-word Stroop task.

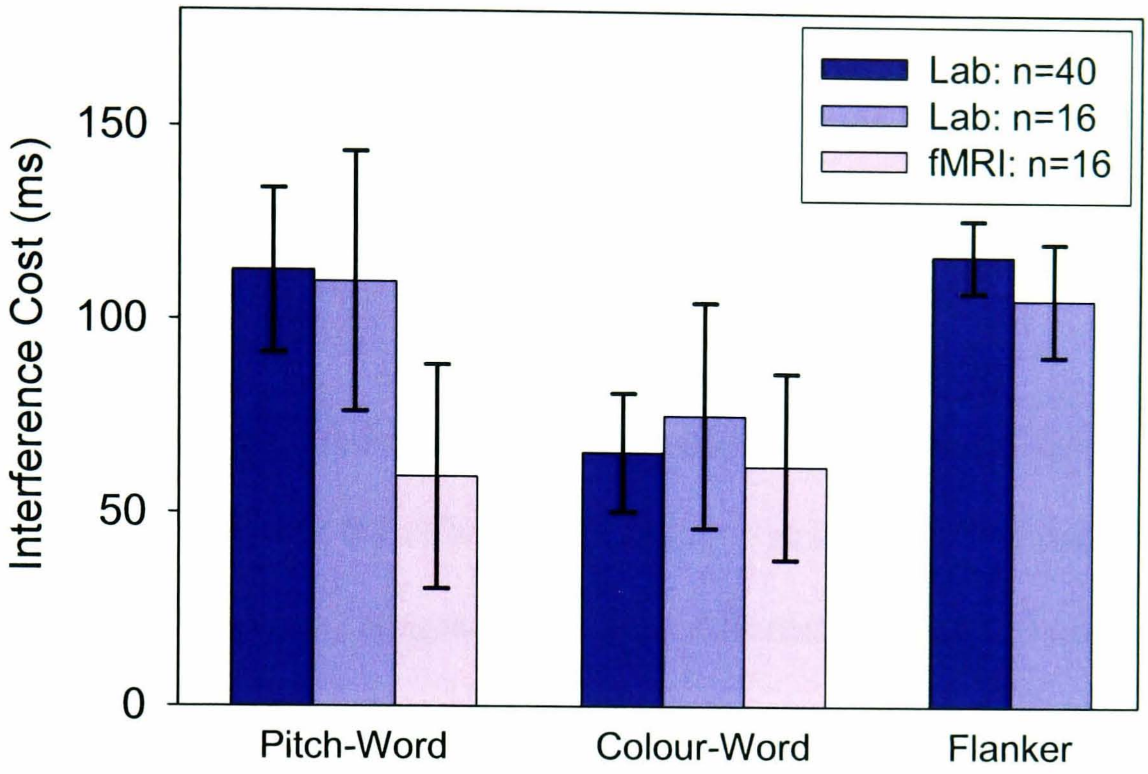
Table 3.3: Mean RTs (ms) and accuracy (% errors) to incongruent, neutral and congruent stimuli from the three tasks. Figures relate to the behavioural study where n=40 (and n=16), and the fMRI study.

	Flanker Task	Colour-word Stroop	Pitch-word Stroop
(a) Mean RTs (ms). Behavioural study n=40 (n=16), fMRI study.			
Incongruent	635 (594)	743 (737), 723	756 (729), 578
Neutral	518 (488)	677 (662), 661	643 (619), 518
Congruent	517 (490)	636 (633), 615	611 (590), 495
(b) Mean accuracy (% errors). Behavioural study n=40 (n=16), fMRI study.			
Incongruent	5.6 (5.8)	6.3 (7.2), 7.0	10.5 (8.9), 6.1
Neutral	1.1 (1.2)	7.3 (8.9), 5.0	3.0 (2.8), 3.2
Congruent	0.6 (0.6)	6.0 (6.8), 3.9	1.0 (1.0), 2.4

Table 3.4: ANOVA results: Main effects and planned contrasts (with $p_{critical}=0.05/2 = 0.025$) investigating interference (incongruent - neutral) and facilitation (neutral - congruent). $^{*}p<0.025$, $^{**}p<0.01$, $^{***}p<0.001$.

	Flanker Task	Pitch-word Stroop	Colour-word Stroop
Behavioural Study (n=40): RTs			
Main effects	$F_{1.3,52.1}=450.2^{***}$	$F_{1.4,56.4}=95.6^{***}$	$F_{1.7,66.2}=95.7^{***}$
Interference	$t_{39}=25.1^{***}$	$t_{39}=10.7^{***}$	$t_{39}=8.6^{***}$
Facilitation	n.s.	$t_{39}=4.1^{***}$	$t_{39}=6.5^{***}$
Behavioural Study (n=40): Accuracy			
Main effects	$F_{1.1,43.5}=33.4^{***}$	$F_{1.2,46.5}=66.8^{***}$	n.s.
Interference	$t_{39}=5.8^{***}$	$t_{39}=9.0^{***}$	---
Facilitation	n.s.	$t_{39}=4.0^{***}$	---
Behavioural Study (n=16): RTs			
Main effects	$F_{1.3,19.5}=188.4^{***}$	$F_{1.3,20.2}=41.9^{***}$	$F_{1.2,18.5}=35.4^{***}$
Interference	$t_{15}=15.4^{***}$	$t_{15}=7.0^{***}$	$t_{15}=5.5^{***}$
Facilitation	n.s.	$t_{15}=2.8^{*}$	$t_{15}=4.5^{***}$
Behavioural Study (n=16): Accuracy			
Main effects	$F_{2,30}=9.5^{***}$	$F_{2,30}=10.3^{***}$	n.s.
Interference	$t_{15}=2.7^{*}$	$t_{15}=3.0^{*}$	---
Facilitation	n.s.	n.s.	---
fMRI Study (n=16): RTs			
Main effects		$F_{1.2,18.4}=23.2^{***}$	$F_{2,30}=68.1^{***}$
Interference		$t_{15}=4.4^{**}$	$t_{15}=5.5^{***}$
Facilitation		$t_{15}=3.8^{**}$	$t_{15}=6.1^{***}$
fMRI Study (n=16): Accuracy			
Main effects		$F_{2,30}=10.3^{***}$	$F_{2,30}=9.5^{**}$
Interference		$t_{15}=3.0^{*}$	$t_{15}=2.7^{*}$
Facilitation		n.s.	n.s.

Figure 3.5: RT interference effects from the three tasks: data from the behavioural study for the initial group of 40 subjects and the subset of 16 subjects, and data collected during scanning. Error bars show 95% confidence intervals.



Behavioural study (n=16)

Data from the behavioural study for the subset of 16 subjects who returned to take part in the fMRI study show the same pattern of results (Table 3.4), with the exception that with only 16 subjects there was no longer a significant facilitation effect in the accuracy data from the pitch-word Stroop task.

fMRI study

Reaction times acquired during the fMRI study also revealed significant condition-specific differences (Table 3.4). Planned contrasts show significant interference and facilitation effects for both tasks. As with the behavioural study, interference costs (pitch-word: 60 ms; colour-word: 62 ms) were greater than facilitation benefits (23 ms and 46 ms respectively). There were also significant differences in accuracy across conditions. Planned contrasts show these to be due to interference costs only in both tasks.

Correlations

Correlations were conducted both within and between studies. Table 3.5 shows correlations between RT effects (interference = incongruent - neutral; facilitation = neutral - congruent) from the different tasks from the behavioural study (n=40).

Table 3.5: Correlations between reaction-time effects from the three tasks in the behavioural study (n=40).

	Colour-word Stroop	Pitch-word Stroop
Interference		
Colour-word Stroop	---	0.280 (p=0.081)
Flanker Task	0.298 (p=0.062)	0.326 ($p<0.05$)
Facilitation		
Colour-word Stroop	---	0.043 (p=0.793)
Flanker Task	-0.163 (p=0.316)	0.206 (p=0.203)

The tasks performed during the fMRI study differed slightly from those performed during the behavioural study. In particular, while stimuli were presented in random or pseudo-random order during the behavioural study, a blocked design was used for the fMRI study. To test whether these differences, and the additional effects of lying in a scanner, affected performance, correlations were calculated between RT measures obtained from the two studies. RT measures of interference correlated significantly (or near-significantly) between the fMRI and behavioural studies for the sixteen subjects that took part in both (pitch-word Stroop task: $r=0.567$, $p<0.05$; colour-word Stroop task: $r=0.492$, $p=0.053$). RT measures of facilitation did not correlate significantly (pitch-word Stroop task: $r=0.282$, $p=0.290$; colour-word Stroop task: $r=-0.303$, $p=0.254$). There were no significant correlations between interference or facilitation measures for the colour-word and pitch-word tasks performed during the fMRI study. However, with only 16 subjects it is difficult to interpret non-significant correlations.

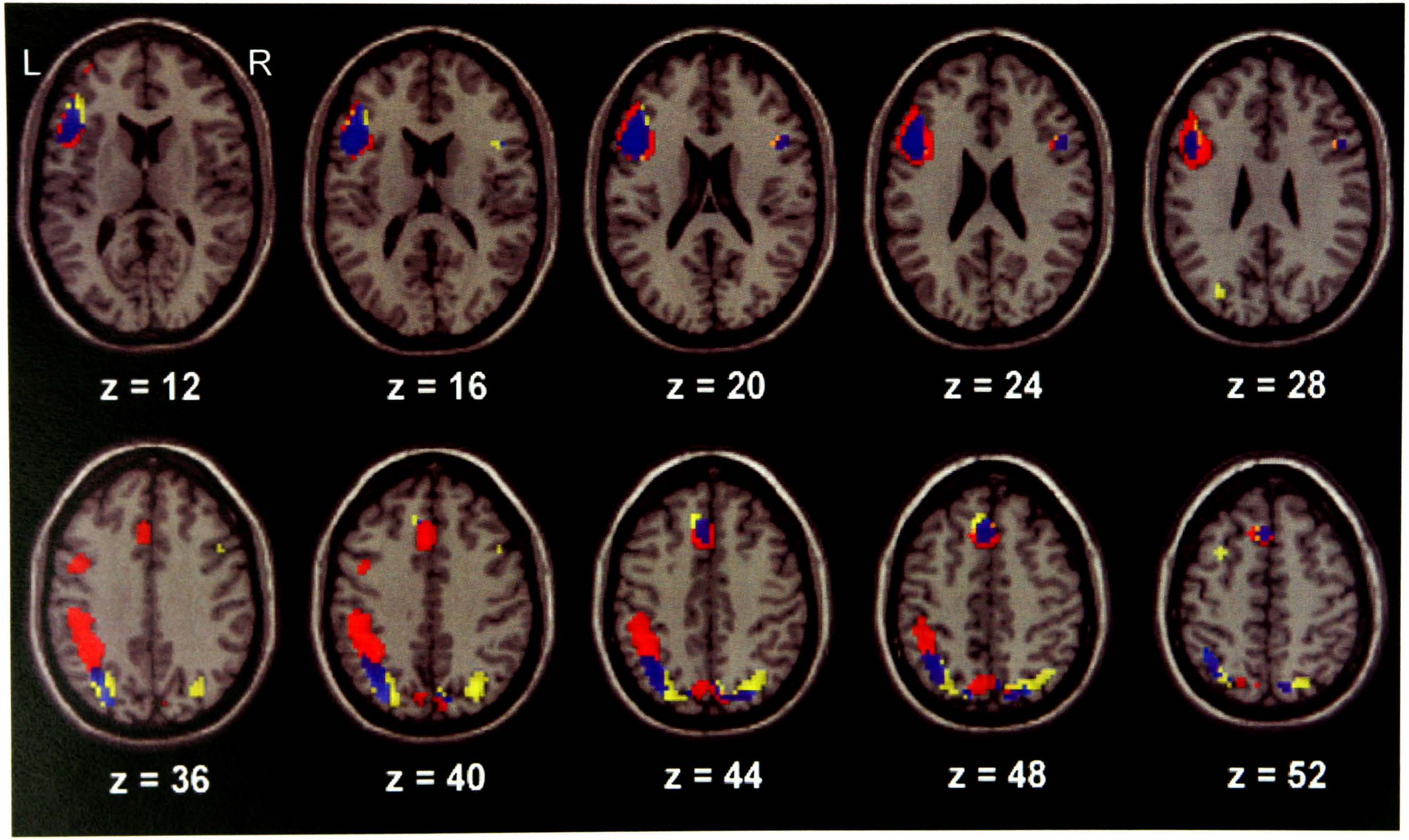
fMRI Results

Table 3.6 shows the peaks of regions significantly more active during the incongruent blocks than the neutral blocks for each of the two tasks, and from a conjunction analysis identifying areas of common activation. Most of the common activation was found in the anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (DLPFC) (Figure 3.6). There is some common activation in the parietal lobe centred on the intraparietal sulcus (IPS), but there are also large areas of modality-specific activation.

Table 3.6: MNI coordinates of peaks of activity associated with interference (incongruent - neutral) for each of the two tasks, and from a conjunction analysis identifying areas of common activation (p<0.001 uncorrected, extent threshold 14 voxels).

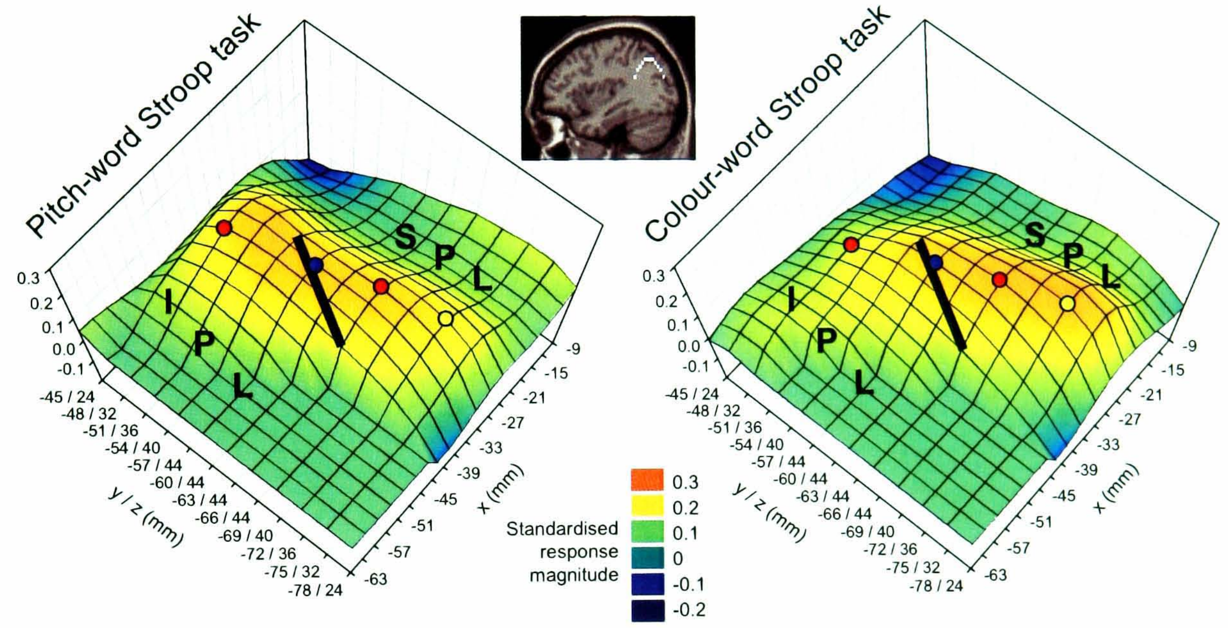
		Colour-word Stroop			Pitch-word Stroop			Conjunction		
Region	BA	x	y	z	x	y	z	x	y	z
Anterior cingulate	32	-6	30	48	-3	21	44	-3	24	48
L DLPFC	45	-45	27	20	-45	12	24	-42	42	8
		-48	18	20	-48	18	8	-51	18	20
					-54	21	20			
R DLPFC	44	48	15	24	48	15	24	48	15	24
R DLPFC	8	45	12	40						
L Premotor Cortex	6	-30	9	56						
		-39	0	56						
R Middle frontal gyrus	11	24	30	-16						
L Superior parietal lobe	7	-30	-72	36	-33	-66	44	-36	-60	44
		-12	-78	48				-27	-78	36
R Superior parietal lobe	7	30	-72	40	24	-78	44	24	-75	48
					15	-75	48			
L Inferior parietal lobe	40				-39	-51	36			
					-45	-30	36			
Precuneus	7	6	-72	44	-6	-72	44	9	-78	40
								12	-72	52

Figure 3.6: Activation associated with pitch-word Stroop conflict (red), colour-word Stroop conflict (yellow), and a conjunction of the two (blue). The top row shows activation in DLPFC; the bottom row shows activation in ACC and the parietal lobe.



The overlapping parietal lobe activation is located around the IPS. Activation extended downwards, anteriorly and laterally for the auditory task, and posteriorly and medially for the visual task. This differential bias has two alternative explanations. First, interference from both tasks results in activation in one area, but which decreases in different directions. Second, interference from the auditory and visual tasks results in activation in separate areas, which intersect due to the spatial smoothness of the data. The graphs in Figure 3.7 plot the activation data in an alternative format in order to try to address this question. Standardised response magnitudes based on beta weights were obtained for points along a curved section of left-hemisphere cortex passing through the peaks of activation in the parietal lobe. Beta weights were from the incongruent - neutral contrast for the auditory and visual tasks. Beta weights represent the condition-specific effects in the MR signal in each voxel, controlling for the contributions of other condition effects, and for non-task related factors such as head movement and cardiac variability. Beta weights from the incongruent - neutral contrast therefore reflect the strength of conflict-related activation. The curve intersects the first peaks reported in Table 3.6 for the left superior parietal and left inferior parietal lobe, and has the equation $z = -0.0793y^2 - 9.7747y -256.34$. Values for points along this curve were obtained for pitch-word and colour-word incongruent - neutral contrasts separately, for values of x between -63 mm and -9 mm (in 3 mm steps). These voxel beta values were then plotted in 3D mesh graphs (Figure 3.7). The graphs for both contrasts show a ridge of activation across the IPS (illustrated by the black line). However, the peak of the colour-word Stroop contrast is clearly shifted in position relative to those for the pitch-word Stroop. This shift across the axis of the IPS indicates functional segregation. The peak of the conjunction reflects the intersection of the spread of maximal activation located in two adjacent areas, rather than a common maximum that is present in the IPS in both auditory and visual contrasts.

Figure 3.7: Strength of conflict-related activation (incongruent - neutral) in the left parietal lobe resulting from the two tasks. The yellow dot indicates the location of the colour-word Stroop peak of activation, the red dots shows pitch-word Stroop peaks of activation, while the blue dot shows the peak of the conjunction. The black line shows the approximate location of the intraparietal sulcus. The example sagittal slice is at x = -36 mm; the white line shows the shape and location (in y and z) of the curve.



A paired t-test showed no voxels which were significantly more active during pitch-word Stroop conflict than during colour-word Stroop conflict. One area was significantly more active during colour-word Stroop conflict: an area of 14 voxels in the right middle frontal gyrus (BA 11), with a peak at 27, 30, -16 mm. This peak is 3 mm Euclidean distance from the peak for the colour-word Stroop task (24, 30, -16) indicating that this region responds differentially across the two conflict comparisons.

To investigate neural correlates of facilitation, two comparisons were made: neutral - congruent and congruent - neutral. Neither comparison revealed any significant differences for the pitch-word Stroop task, but there were areas which differed significantly between conditions for the colour-word Stroop task. These areas are described in Table 3.7.

Table 3.7: Peak MNI coordinates of areas significantly different in the congruent and neutral conditions in the colour-word Stroop task.

Comparison	Region	BA	x	y	z
Congruent - Neutral	L Middle occipital lobe	19	-45	-75	16
Neutral - Congruent	L Fusiform gyrus	37	-42	-51	-20
	R Fusiform gyrus	37	39	-66	-20
			39	-54	-20

Discussion

All tasks were successful in generating response conflict, in both the behavioural and fMRI studies. Reaction time measures of interference were positively correlated between all three tasks in the behavioural study, and between corresponding measures from the fMRI study and behavioural study. These correlations are themselves evidence of a common mechanism engaged during both auditory and visual conflict tasks. Facilitation effects were less consistent. In the behavioural study the flanker task failed to generate a facilitation effect. Measures of facilitation were not significantly correlated in the behavioural or fMRI studies, or between the behavioural and fMRI studies. While the effects of interference and facilitation are known to be asymmetric, with interference costs being larger than facilitation benefits (MacLeod &

MacDonald, 2000), the variability in the facilitation measure highlights the problems inherent in contrasting incongruent and congruent trials rather than obtaining measures of interference and facilitation separately, even when the facilitation effect is small or absent behaviourally.

Contrasting activation on incongruent and neutral trials revealed conflict-related activity in ACC and bilaterally in DLPFC. These areas comprise the anterior network for conflict resolution identified in a number of other studies (e.g. Casey et al., 2000; Fan et al., 2003). Peak ACC activation from the auditory and visual tasks fell inside, or adjacent to, the cluster of peaks revealed by the meta analysis. DLPFC activation was anterior and ventral to the region identified in the meta analysis. However, there was substantial overlap between regions active during auditory and visual conflict, with peaks of these regions located in close proximity to each other. Activation was also found bilaterally in the superior parietal lobe (SPL) and in the precuneus for both tasks, while auditory activation additionally extended from left SPL into left inferior parietal lobe (IPL). While the auditory and visual peaks of activity in SPL were also close together (no more than two voxels apart in any dimension), there was substantially less overlapping activation than in DLPFC. Activation in left SPL was adjacent to the region identified by the meta analysis, but right SPL activation was more medial than that found in the meta analysis, and left IPL and precuneus activation did not map onto any regions identified by the meta analysis. Two areas of the frontal lobe were additionally active during visual conflict: the left premotor cortex (BA 6) and right middle frontal gyrus (BA 11). Neither of these regions was identified by the meta analysis. While the meta analysis showed reliable conflict-related activation in bilateral insula, these regions were not more active during incongruent trials compared with neutral trials in either the auditory or visual task. While consistency with the results of the meta-analysis is encouraging where it occurs, differences in subject groups, normalisation templates, smoothing kernels, and statistical thresholding mean that inconsistencies are to be expected (Brett, Johnsrude, & Owen, 2002).

A stringent pairwise analysis revealed significant modality-specific effects of conflict only in right MFG, where activation was present only in the colour-word Stroop contrast. The main effects also revealed that the premotor cortex was engaged by conflict in the colour-word Stroop task but not the pitch-word Stroop task. Task-specific processing may account for the activation found in the premotor cortex for visual conflict alone. This area is associated with motor planning, which may have been more difficult in the colour-word Stroop task which had four response buttons and an arbitrary response-to-button mapping, unlike the pitch-word Stroop task which had two response buttons and a more intuitive response-to-button mapping (to respond ‘high’, subjects pressed the button furthest from them - the top button in the array - while to respond low, they pressed the button nearest to them - the bottom button). The task-specific processing view may also account for the activation in middle frontal gyrus associated with visual but not auditory conflict resolution. In a review of a number of neuroimaging studies of diverse cognitive tasks, Duncan and Owen (2000) found that peaks of activation were approximately grouped in three regions: dorsal ACC, mid-dorsolateral prefrontal cortex, and mid-ventrolateral prefrontal cortex. While the review did not identify specific roles for each region, it did suggest that functional specificity could account for the groupings. The area of MFG associated with visual conflict resolution in this study is located in the mid-ventrolateral region.

Another striking observation was the different spatial distribution of parietal lobe activation across auditory and visual tasks. These differences could be meaningfully related to one of two mechanisms: differential flow of information from sensory to higher-order areas, or selective attention processes.

In support of the first hypothesis, the pattern of activation is consistent with flow of information from sensory areas to the parietal lobe. Anatomical links between auditory and visual cortices and the parietal lobe are well documented (e.g. Mishkin, Ungerleider, & Macko, 1983; e.g. Romanski et al., 1999). Auditory conflict-related activity was located anterior and lateral to that found for the visual task. This bias may relate to the locations of auditory and visual sensory cortices (which are respectively anterior and lateral, and posterior and medial to IPS), and their projections to the parietal cortex. In addition, Bushara et al. (1999) found modality-specific activation in the parietal lobe during auditory and visual spatial localisation tasks, and Nishitani, Nagamine, and Shibasaki (1998) found modality-specific activation in the inferior parietal lobes during auditory and visual oddball tasks. While the pattern of activation in these studies does not precisely match that found in this study, they do provide further evidence of modality-specific processing in the parietal lobe.

The second hypothesis is that the differential activation relates to selective attention processes. This is particularly likely for the activation around the intraparietal sulcus, which has been shown to play a role in selective attention (Corbetta et al., 2000). The observed differences in activation in this area may reflect differences in strategies used by subjects to overcome conflict in the two tasks. For example, focusing on a small section of the word is an effective selective-attention strategy for overcoming colour-word Stroop conflict, which would not be beneficial during pitch-word Stroop conflict. Casey et al. (2000) varied the probability of an incongruent stimulus being presented in order to differentiate between selective attention and conflict resolution processes during a flanker task. An effective strategy for overcoming conflict from flankers is to narrowly attend to the central item in the array. Casey et al. hypothesised that if an incongruent stimulus was expected, then this attentional strategy would be engaged. However, if an incongruent stimulus was not expected, this would be a less efficient strategy and would therefore be less likely to be engaged. Their results show differential activity related to the different expectations, and are consistent with an anterior system involved in conflict resolution, incorporating ACC and DLPFC, and a posterior system involved in selective attention, including the superior parietal lobe. The results presented here are also consistent with these functional roles.

To identify areas associated with the facilitation effect, two comparisons were conducted. First, to identify areas more active during congruent trials than neutral trials. No areas were found for the auditory task, but for the visual task an area of left middle occipital lobe (BA 19) was revealed. In a review of neuroimaging studies of word reading, Fiez and Petersen (1998) reported that there is converging evidence from language-related studies that this area is involved in visual analysis specific to word-like stimuli. This functional role would explain the lack of activation in the auditory task, although it is less clear why this area should be more active during congruent trials than neutral trials. The second comparison identified areas more active during neutral trials than congruent trials. In other words, areas showing reduced activity during congruent trials. As with the previous comparison, there were no areas active for the auditory task, but the visual task showed bilateral activation in the fusiform gyrus with reduced activation during congruent trials. The fusiform gyrus is associated with visual processing in the ventral 'what' stream, and is specifically associated with face processing (Grill-Spector, Knouf, & Kanwisher, 2004) and, in the left hemisphere, visual word form processing (McCandliss, Cohen, & Dehaene, 2003). Modality-specificity has been demonstrated for visual word form processing, as this area does not respond to spoken words (Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002). However, the mechanism by which fusiform gyrus is bilaterally involved in facilitation on this task is not clear.

Conclusions

The results are consistent with a supramodal anterior network involved in conflict monitoring and resolution which incorporates the ACC and DLPFC. Common neural processing within this network probably contributes to the behavioural similarities in the RT interference effect. The auditory and visual Stroop tasks differentially activated the parietal lobe, premotor cortex, and right middle frontal gyrus. We identify two possible explanations for these dissociations: differential flow of information from sensory to higher order areas; and selective attention processes.

Chapter 4: Orienting to spatial and non-spatial stimulus features

Introduction

The auditory attention network test (ANT) described in Chapter 2 demonstrated that subjects gained a significant, but highly variable, benefit from an alerting cue, and a small, variable, and non-significant benefit from a cue to spatial location. These effects may have been influenced by the experimental design and so the study reported in this chapter (the Vowels study) uses a different design which addresses some of the limitations of the design of the auditory ANT. First, the linguistic stimuli used in the auditory ANT may have allowed subjects to alert and to orient to the stimulus using the sound of the initial consonant, before sufficient information was available on which to make the pitch judgment. In contrast, the Vowels study used brief steady-state synthesised-vowel stimuli. Since the spectro-temporal structure of the stimuli did not vary over time, all information on which to base a response was available from the onset of the stimulus. Second, it is possible that the SOA chosen for the auditory ANT (650 ms) was not optimal for revealing effects of alerting and orienting, as these effects may have been maximal at an earlier or later point in time. The Vowels study measures benefits of alerting and orienting at three different stimulus onset asynchronies (SOAs) to enable the time course of attentional effects to be investigated. Third, the auditory ANT investigated orienting to a spatial location. Since auditory information is initially processed tonotopically, a cue to target frequency or pitch may prove more beneficial than a cue to location. The Vowels study was designed to investigate not only alerting and orienting to a location, but also orienting to pitch, and orienting to a combination of location and pitch. Finally, benefits from attending to an auditory target may be more robust if there is a competing stimulus to be ignored, so that attention acts both to attenuate the unattended information and to enhance the attended information. This hypothesis was explored by presenting target and distractor vowels concurrently.

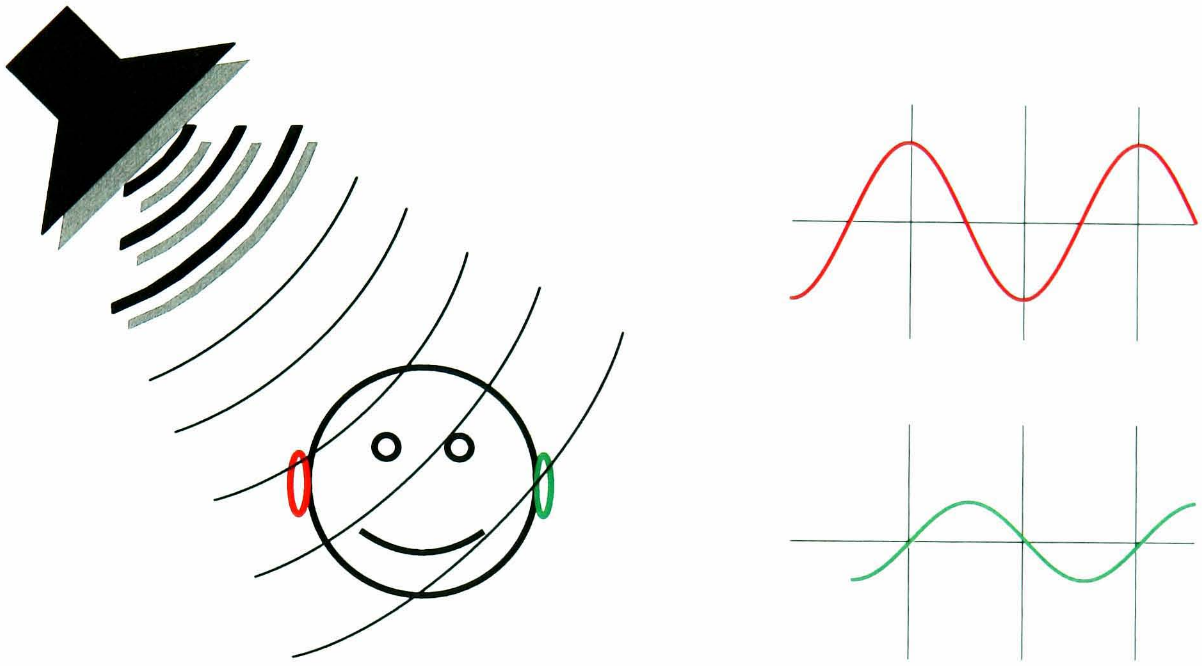
Effect of SOA

The period of time between cue onset and target onset influences the attentional effects being assessed. At short SOAs (around 100 ms), cues presented at a target location automatically capture attention (exogenous orienting). At longer SOAs, when subjects are presented with informative cues they are able to voluntarily orient their attention to the cued location (endogenous orienting). The auditory ANT presented stimuli with an SOA of 650 ms, which was not necessarily optimal for detecting endogenous orienting effects. Maximum benefit from an orienting cue may have occurred at an earlier or later point during the trial.

Spatiotopic vs. tonotopic organisation

In vision and touch the sensory epithelia (the retinae and skin) are organised spatiotopically. There are variations in acuity across the receptors so that visual or tactile information is coded with finer spatial precision if it falls onto an area of high acuity (the fovea or, for example, the lips or fingertips). Visual and somatosensory information is then represented spatiotopically throughout the neural network, where there are neurons narrowly tuned for spatial location, and organised in eye-centred, head-centred, body-centred or limb-centred coordinate frames (Cohen & Andersen, 2002). In contrast, the auditory sensory epithelium, the cochlea, is organised tonotopically, and neurons in auditory cortex are broadly organised according to their sensitivity to frequency. The spatial location of the source of auditory stimuli must be calculated from acoustic cues. The primary cues are differences in the time at which a sound reaches the two ears (interaural time difference: ITD), differences in the level of the sound at the two ears (interaural level difference: ILD), and variations in spectral characteristics introduced by the head and pinnae. Figure 4.1 illustrates the interaural difference cues. The signal arrives at the listener's right ear slightly earlier, and at a higher level, than at the left ear. Spatial location across the azimuthal plane is primarily coded by ITD and ILD, with ITD cues being prominent for low-frequency sounds (up to around 1500 Hz) and ILD cues being prominent for high-frequency sounds (higher than around 1500 Hz). Computation of the location of an auditory sound source takes place in the brainstem and midbrain. A spatiotopic map of auditory space is found in the superior colliculus (Cohen & Knudsen, 1999), a midbrain site involved in reflexive head and eye movements which also contains maps of visual and somatosensory space. However, as yet no spatiotopic maps of auditory space have been found in the cortex (Middlebrooks, 2000). Neurons broadly tuned for space are found in posterior auditory cortex (Tian, Reser, Durham, Kustov, & Rauschecker, 2001) and there is some evidence for a putative dorsal 'where' pathway analogous to the visual 'where' pathway (Arnott, Binns, Grady, & Alain, 2004). However, neurons tuned for auditory space do not show the narrow spatial tuning of visual neurons, and are not organised in a strictly spatiotopic fashion (Middlebrooks, 2000).

Figure 4.1: Diagram to illustrate interaural time and level cues. The sine-wave signal will arrive at the listener's right ear earlier and at a higher level (louder) than at the left ear.



Orienting to frequency

Since the auditory system is primarily organised according to frequency, it is possible that frequency cues would be more beneficial in orienting attention than location cues. Scharf et al. (1987) reported that sounds presented at an expected frequency were more likely to be detected than sounds presented at unexpected frequencies. Using a discrimination task, Mondor and Bregman (1994) found that valid cues to the frequency of a target facilitated judgments of target duration, for both reaction times and accuracy.

Psychoacoustical evidence demonstrates that listeners are unable to segregate concurrent sounds on the basis of differences in ITDs alone. If a sound is presented over headphones so that the signal to one ear arrives slightly earlier (typically less than 1 ms) than the signal to the other ear, the listener will perceive a single sound, which is lateralised to the side the signal arrived at first. Culling and Summerfield (1995) presented two synthesised vowels concurrently, but with ITDs which lateralised one vowel to the left and one vowel to the right. They found that listeners were unable to segregate the two vowels when ITD was the only cue available. Listeners were able to segregate the vowels when additional cues were available, such as a differences in the interaural intensity of the vowels (Culling & Summerfield, 1995) or differences in the fundamental frequency (pitch) of the vowels (Summerfield & Akeroyd, 1998). Hill and Darwin (1996) created complex tones comprising seven harmonics of a fundamental frequency of 100 Hz (200, 300, 400, ..., and 800 Hz) and presented them with an ITD of +1500 μs so that they were lateralised to the right. When the 500-Hz component was presented with a different onset time or at a mistuned frequency, it was perceived as being lateralised to the left ear, contralateral to the remaining harmonics. Since the 500-Hz component had the same ITD as the other harmonics, this result suggests that grouping according to common onset time or common harmonicity precedes computation of the location of the grouped components (Summerfield & Akeroyd, 1998). Hence when two synthesised vowels are presented with different ITDs and different fundamental frequencies (f_0s), segregation will initially be conducted based on the difference in f_0. Once this has been achieved, the locations of the two segregated auditory objects can be determined. This stepwise process might suggest that auditory attention could be most usefully directed to a fundamental frequency, rather than to a location.

Orienting to location

Despite the previous arguments, there are some reasons to hypothesise that attention might be more suited to orienting to a spatial location. Attention plays a critical role in the processing of visual and tactile sensory information, both of which are processed spatiotopically. Covert orienting to a spatial location (i.e. without a head or eye movement) has been shown to operate using the same cortical areas as overt orienting, in which the eye gaze is moved to the attended location (Corbetta et al., 1998; Nobre et al., 2000). Nobre and colleagues (Griffin & Nobre, 2003; Nobre et al., 2004) used a cueing task in which subjects were presented with cues either before or after a target stimulus appeared, so that they were cued to either a location in the external world, or to an internal representation. Behavioural data showed that both types of cue led to spatial orienting benefits. Overall, ERP (Griffin & Nobre, 2003) and neuroimaging (Nobre et al., 2004) data showed extensive similarities between the two types of orienting. Both types of orienting were associated with a lateralised early posterior ERP component, followed by a later lateralised frontal component. fMRI data showed overlapping activation in a number of regions in the parietal, frontal and visual cortices, including superior parietal lobe and around the intraparietal sulcus bilaterally, left-hemisphere inferior parietal lobe, bilateral dorsolateral prefrontal / premotor cortex, and frontal eye fields. In addition to this common activation, both ERP and fMRI data showed frontal components which were present when subjects oriented to internal, but not external, representations of space. The overlapping activation illustrates the attention system's use of the overt orienting system, even when an overt orientation is not possible. Since there is evidence that eye movements do play a role in spatial orienting to the location of auditory sound sources (Rorden & Driver, 1999), it is possible that attention is more suited to attending to the location of auditory stimuli than to their spectral characteristics.

Orienting to an auditory object

A further possibility is that attention can operate most effectively upon an auditory object comprising both location and frequency information. There is debate over whether auditory attention can be directed to frequency and location separately, or whether attention is (also) directed to an auditory object (such as a voice) which encompasses both features. ERP evidence (Woods & Alain, 1993; Woods, Alho, & Algazi, 1994) reveals that location and frequency are initially processed separately, and are conjoined around 110-120 ms after stimulus onset. Furthermore, Woldorff et al. (1993) used combined MEG and MRI techniques while subjects attended to either frequency or location. They report modulation of activity in the ranges 20-50 ms and 80-130 ms following stimulus onset, and localised the source of the modulated signal to auditory cortex. This suggests that attention can modulate sensory processing before location and frequency information are conjoined at around 110 ms. While these studies provide evidence that attention can operate on location and frequency separately, there is also evidence to suggest that attention can operate upon auditory objects. Zatorre, Mondor, and Evans (1999) used PET to measure cortical activity while subjects heard tones which varied in frequency and location, and made responses based on either the frequency or location of the tones. No difference was found in cortical activity based on whether subjects were attending to frequency or location. Zatorre et al. concluded that auditory attention 'operates at a level at which separate features have been integrated into a unitary representation' (p. 544). Similarly, Mondor, Zatorre, and Terrio (1998) found that listeners were unable to attend to the location of an auditory stimulus independent of its spectral characteristics, and vice versa. These studies suggest that attention can operate upon an auditory object, rather than separate stimulus features.

Darwin and Hukin (1999) presented listeners with two carrier sentences, in which they embedded two target words. Listeners were instructed to attend to one of the sentences, and report the target word from that sentence. The target words could share the same fundamental frequency as the carrier sentence, the same ITD as the carrier sentence, both fundamental frequency and ITD, or neither fundamental frequency nor ITD. The results demonstrated that listeners were more likely to select the correct target word if it shared an ITD with the carrier sentence than if it shared a fundamental frequency. This result is surprising in light of the evidence that listeners are unable to segregate two concurrent vowels on the basis of ITD cues alone. Darwin and Hukin concluded that listeners attend to perceived auditory objects at a subjective location, rather than attending explicitly to frequency components which share a common ITD. If attention is indeed directed to auditory objects rather than to a location or frequency, cues to location, frequency, and both location and frequency together, should prove equally effective.

Effect of stimulus competition

According to Duncan et al.'s (e.g. Duncan et al., 1997) conceptualisation of attention, enhancement of the attended stimulus is achieved through competitive brain activity. This implies that attention will have a greater effect in the presence of distracting stimuli. Evidence from visual attention (Motter, 2000) shows that attention can operate to attenuate the response to a distracting stimulus, allowing the response to the target to be at the same level as if the distractor were not present. In this way, attention is able to isolate the attended object from distracting items, thereby reducing the influence of distractors. It might therefore be expected that greater benefits from orienting attention will be obtained when a target is presented in the presence of distractors than when presented alone. Presenting more than one auditory stimulus simultaneously is complicated by the possibility that the two sounds will fuse together so that they are heard as a single percept. One solution is to use synthesised vowel sounds with different fundamental frequencies, and different ITDs. These simplified speech tokens are heard as two distinct auditory objects, arising from distinct locations (Summerfield & Akeroyd, 1998). This allows a target sound and a distractor sound to be presented simultaneously yet still be perceptually segregated and heard as distinct objects with different lateralisations.

The Vowels experiment

The Vowels experiment used a cueing task to investigate the attentional skills of alerting, orienting to a location, orienting to a pitch, and orienting to both a location and a pitch. The SOA was varied to evaluate the time course of these attentional effects. The experimental design permitted the investigation of two further effects: i) an auditory spatial conflict effect, and ii) the benefits of auditory perceptual pop-out. These effects are described in detail later in the chapter.

Method

Subjects

Subjects who took part in the auditory and visual ANTs and colour-word Stroop task reported in Chapter 2 also took part in this study during the same testing session. Forty volunteers (19 male, mean age 23.7 years, range 16 - 42) participated in the study. Participants gave informed consent prior to the study, and were paid £25 for participation in the battery of tests. Subjects all spoke English as their native language and had normal or corrected-to-normal vision, as tested using a Snellen chart. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed 34 subjects with normal hearing (thresholds below 20 dB HL) and six subjects with thresholds no greater than 25 dB HL, which was considered acceptable for this study. Three additional participants were rejected: two for unacceptable audiograms, and one who consistently scored more than three standard deviations from the mean on the discrimination tests.

Apparatus and Stimuli

Behavioural testing was conducted in a sound-attenuating chamber. Stimuli were presented under the control of Visual Basic programmes implemented at the MRC Institute of Hearing Research on an IBM-compatible personal computer running Windows 98. Auditory stimuli were presented via Sennheisser HD 480II headphones. Instructions and possible responses were presented on a 15-inch flatscreen monitor. Subjects responded by pressing buttons on a response box situated on the desk in front of them. The buttons were arranged in a left to right array.

Table 4.1: Formant frequencies of vowel stimuli (Culling, Summerfield, & Marshall, 1994)

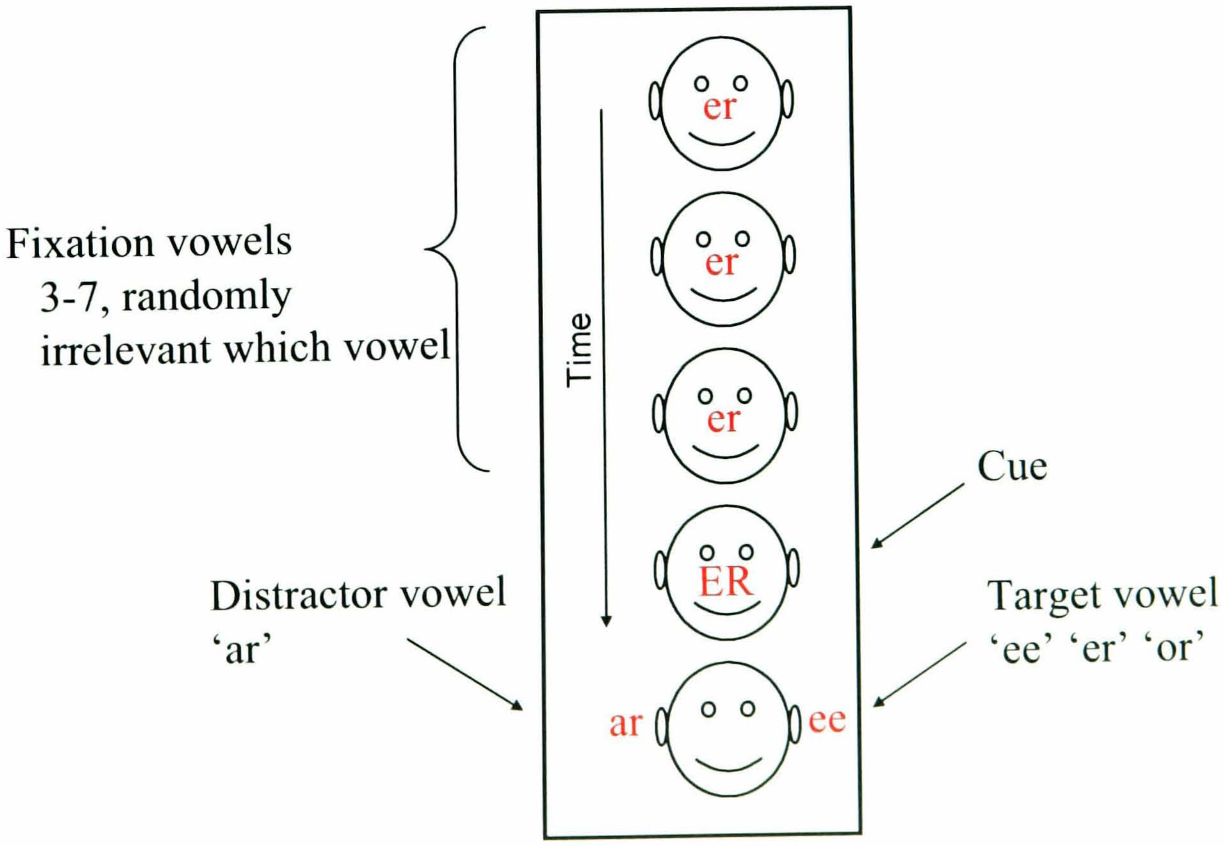
			Formant Frequencies (Hz)				
		Fundamental Frequency	F1	F2	F3	F4	F5
Vowel	Sound						
/a/	‘ar’	75 or 100 or 133	639	1016	2987	3429	4500
/i/	‘ee’	75 or 100 or 133	261	2032	3174	3630	4500
/ɜ/	‘er’	75 or 100 or 133	508	1240	2547	3272	4500
/ɔ/	‘or’	75 or 100 or 133	385	657	2929	3787	4500

The British-English monophthongs /a/, /i/, /ɜ/ and /ɔ/ (‘ar’, ‘ee’, ‘er’, and ‘or’) were created digitally (10,000 samples per second, 16-bit amplitude quantisation) by summing sine waves with amplitudes and phases appropriate for a cascade-formant synthesiser (Klatt, 1980). Table 4.1 details the formant frequencies (Culling et al., 1994), which did not vary over the duration of the stimulus. Stimuli were 50 ms long, gated with a 10 ms cosine window at the start and end, and otherwise of constant amplitude.

Design and Procedure

A single trial of the Vowels task (Figure 4.2) began with a series of identical single-vowel sounds. These *fixation* vowels were followed by a *cue* vowel, and then by the target to which the subjects responded. The target stimulus comprised a pair of vowels, one of which was the *target*, and one the *distractor*. The subjects' task was to identify the *target* vowel.

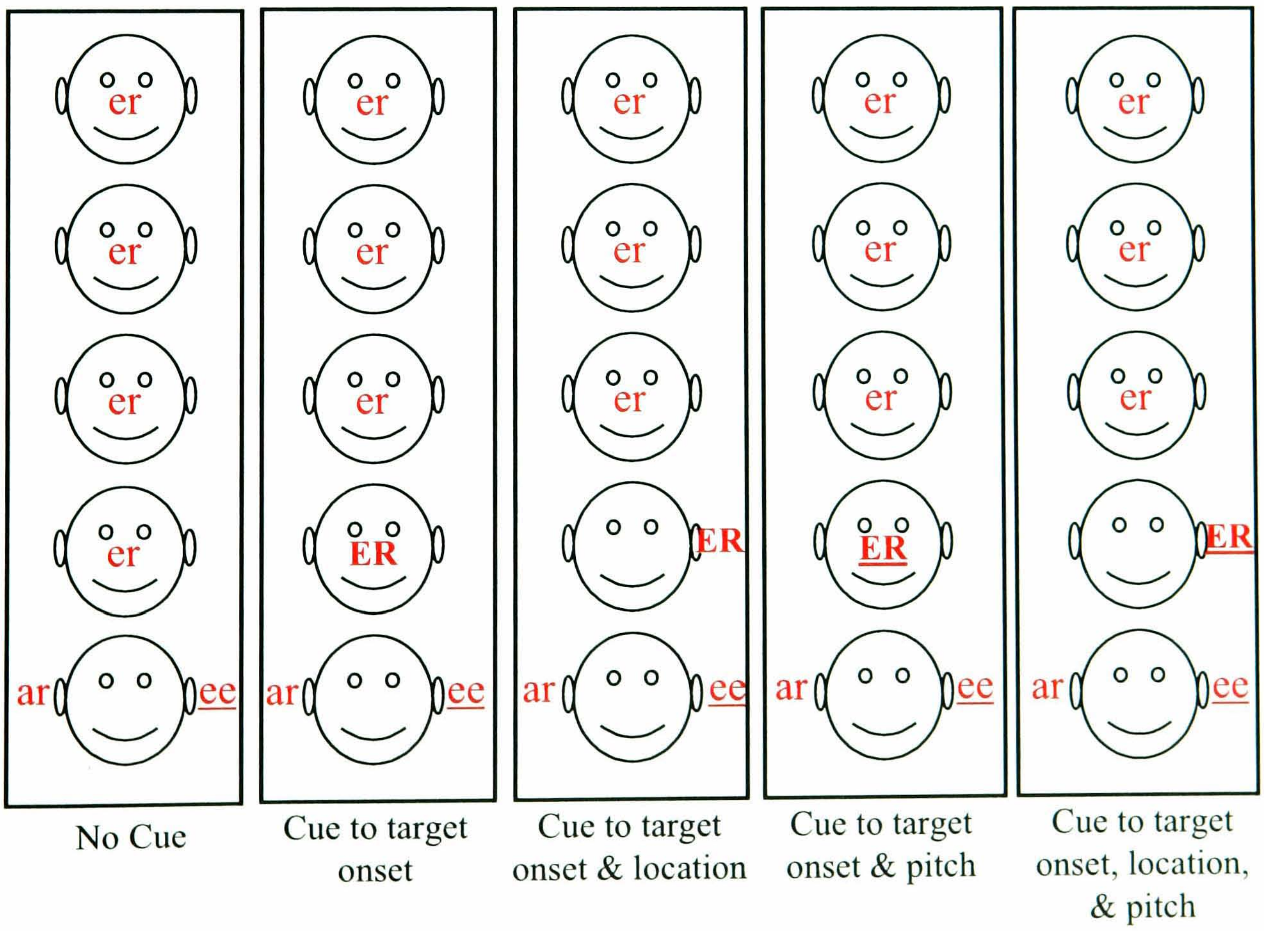
Figure 4.2: Illustration of a single trial on the 'Vowels' task.



Target stimuli were pairs of synthesised vowels presented simultaneously to listeners through headphones ('concurrent vowels'). One member of the pair had an interaural time difference (ITD) of +600 μs so that it appeared lateralised to the right ear. The other member of the pair had an ITD of -600 μs so that it appeared lateralised to the left ear. The members of the pair differed in fundamental frequency (f_0) by 10 semitones. The lower of the pair had an f_0 of 75 Hz (5 semitones below 100 Hz) and the higher of the pair had an f_0 of 133 Hz (5 semitones above 100 Hz). One of the vowels was always 'ar' (as in 'hard'). This vowel was the *distractor*. The other vowel was either 'ee' (as in 'heed'), 'er' (as in 'heard'), or 'or' (as in 'hoard'). This vowel was the *target*. Subjects were instructed to identify the target vowel.

Prior to each pair of concurrent vowels, subjects heard a single, randomly selected vowel repeated a random number of times to introduce uncertainty about target onset. These *fixation* vowels had a fundamental frequency (f_0) of 100 Hz and an ITD of zero, so that they were heard in the middle of the head. The parameters of the last *fixation* vowel were manipulated to provide cues. The *cue* vowel therefore had the same phonetic identity as the *fixation* vowel, but contained additional information to aid the subject in identifying the *target* vowel (Figure 4.3). A *no cue* condition provided a baseline. An *onset cue* was 6 dB louder than the other fixation vowels, and indicated that the target stimulus would be the next sound presented. Orienting cues were also 6 dB louder, but could additionally have the same ITD as the target vowel, the same f_0 as the target vowel, or both the same ITD and the same f_0. In this way the cue vowel could provide a cue to the *location* of the target vowel, or a cue to its *pitch*, or cues to both *location and pitch*. Since orienting cues also acted as alerting cues, performance with orienting cues was contrasted with performance with an alerting cue alone (i.e. in the *onset cue* condition) to obtain specific measures of listeners' ability to orient to location, to pitch, and to both location and pitch. The cue and target vowels were separated by one of three different SOAs (150 ms, 450 ms, or 1050 ms) to investigate whether the benefit obtained from the cues varied as a function of the time available to alert or to orient prior to target onset. The same SOA was used to separate the fixation and cue vowels.

Figure 4.3: Illustration of trial types on the Vowels task. Upper case indicates a 6 dB increase in level. Underlining indicates a higher fundamental frequency.



Participants were trained to identify concurrent vowels through a series of training sessions. Initially, subjects were presented with 20 100-ms single vowel sounds presented diotically (ITD=0 μs), with a fundamental frequency of 100 Hz, and asked to identify them. Subjects then identified 32 100-ms vowels which were presented to the right or the left ear (using an ITD of +/- 600 μs) and which were at a fundamental frequency of 75 Hz or 133 Hz. Subjects then heard 48 target stimuli (concurrent pairs of vowels), one of which was 'ar', and identified the vowel which was not an 'ar'. If subjects experienced difficulty with this practice task it was repeated until they demonstrated that they were able to do the task. Subjects then had a 20-trial practice session using the complete experimental procedure. If subjects experienced difficulty, they repeated this practice session until they could perform the task reliably.

During the experimental phase, subjects completed 720 trials (3 SOAs x 3 target vowels x 4 fixation vowels x 5 cue types x 2 locations x 2 fundamental frequencies) split across three separate occasions during the testing session. At the start of each new occasion, a practice session (with feedback) of twenty trials was completed, followed by twenty warm-up trials, the data from which were discarded. Each testing occasion took approximately 30 minutes, and these were interspersed with testing on the other experiments described in Chapter 2.

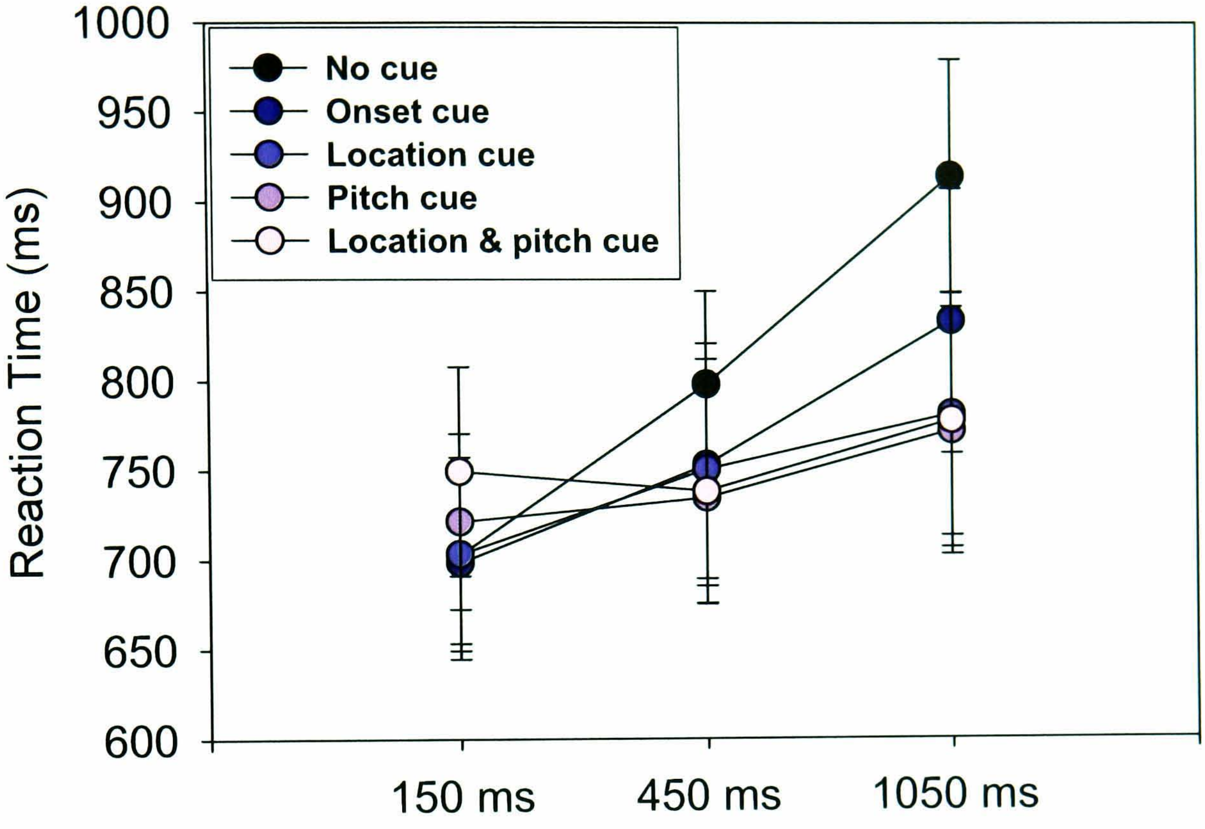
Results

Median reaction times (RTs) and accuracy were analysed using within-subjects ANOVAs and planned contrasts. RTs were taken from correct trials only, and trimmed to exclude responses faster than 100 ms and slower than 3000 ms. Trimming resulted in the removal of 1.19% of trials. Where Mauchley's test of sphericity indicated that sphericity could not be assumed, a Greenhouse-Geisser correction was applied. This is evident from non-integer degrees of freedom. Planned contrasts were conducted using one-way ANOVAs and t-tests with a Bonferroni correction (dividing the critical p value by the number of comparisons being made).

Effects of cue type and SOA

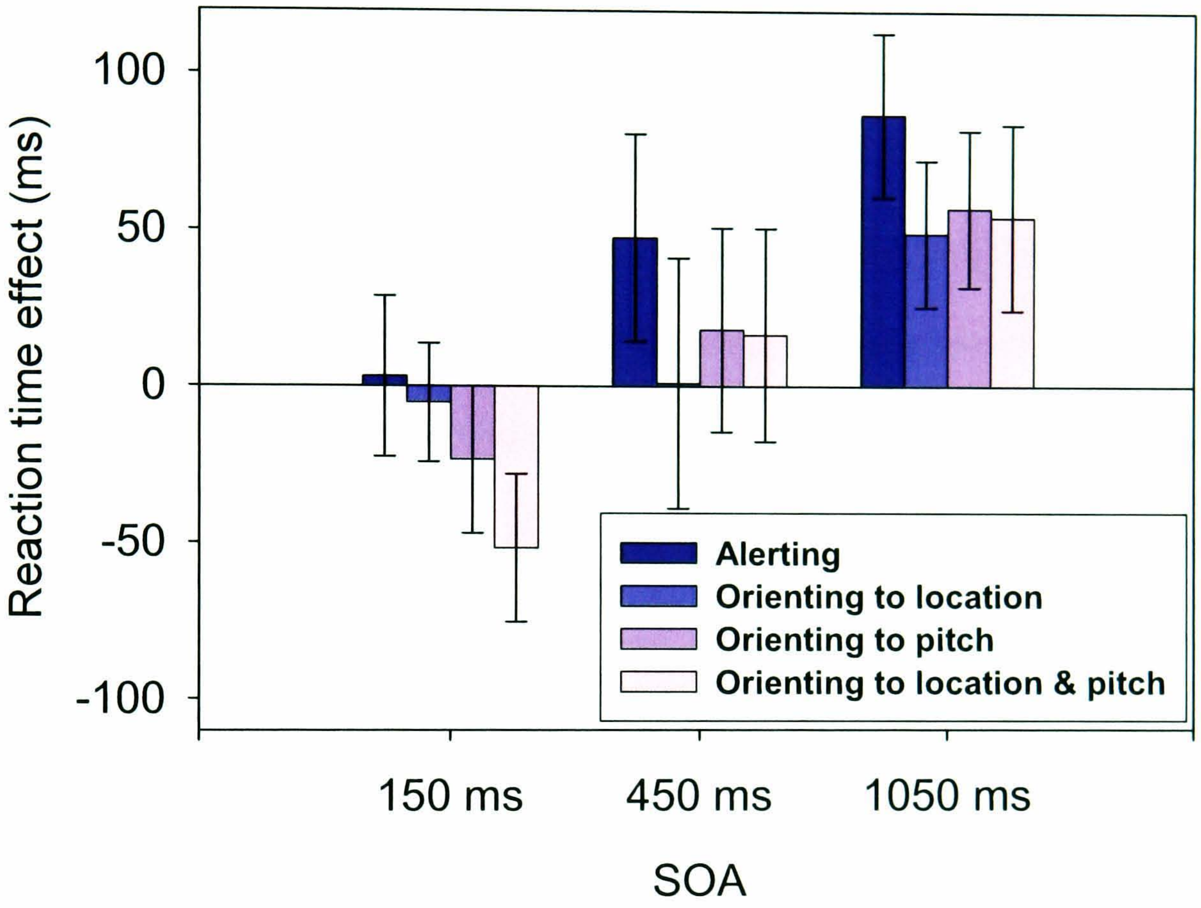
Data were collapsed across vowel identities, locations, and f_0s, since these were randomised across conditions and not relevant to the main hypotheses. RT data for the five cueing conditions at each of the SOAs can be seen in Figure 4.4. A two-way 3 (SOA = 150 ms; 450 ms; 1050 ms) x 5 (Cues = none; onset; location; pitch; location & pitch) within-subjects ANOVA was used to analyse the median RT data. This revealed a significant main effect of SOA ($F_{1.4,53.6}=36.856$, $p<0.001$), a significant main effect of cue type ($F_{4,156}=17.289$, $p<0.001$) and a significant interaction ($F_{5.3,207.3}=15.814$, $p<0.001$). RTs increased with increases in the SOA, so that the longest RTs were with an SOA of 1050 ms.

Figure 4.4: Average reaction times at the three SOAs and with the five cue types. Error bars show 95% confidence intervals.



To explore the cueing effects and the form of the interaction further, one-way ANOVAs were used to compare performance with the different cue types at each level of the SOA. A Bonferroni correction gave a critical p value of 0.05 / 3 = 0.0167. There were significant differences between cue types at the 150-ms SOA ($F_{3.3,128.5}$=5.485, $p<0.01$), the 450-ms SOA ($F_{2.7,106.8}$=6.773, $p<0.001$), and the 1050-ms SOA ($F_{3.3,128.5}$=33.676, $p<0.001$). Since the effects of interest (benefits from alerting and orienting cues) are calculated by comparing performance with different cue types, they should not be affected by overall differences in RTs across SOAs. To investigate these effects, planned contrasts were conducted at each level of the SOA using t-tests with a Bonferroni correction ($p_{critical}$=0.05/4 = 0.0125). These tests evaluate the effects of alerting (no cue - onset cue), orienting to a location (onset cue - location cue), orienting to pitch (onset cue - pitch cue), and orienting to location and pitch (onset cue - location & pitch cue). RT effects are shown in Figure 4.5.

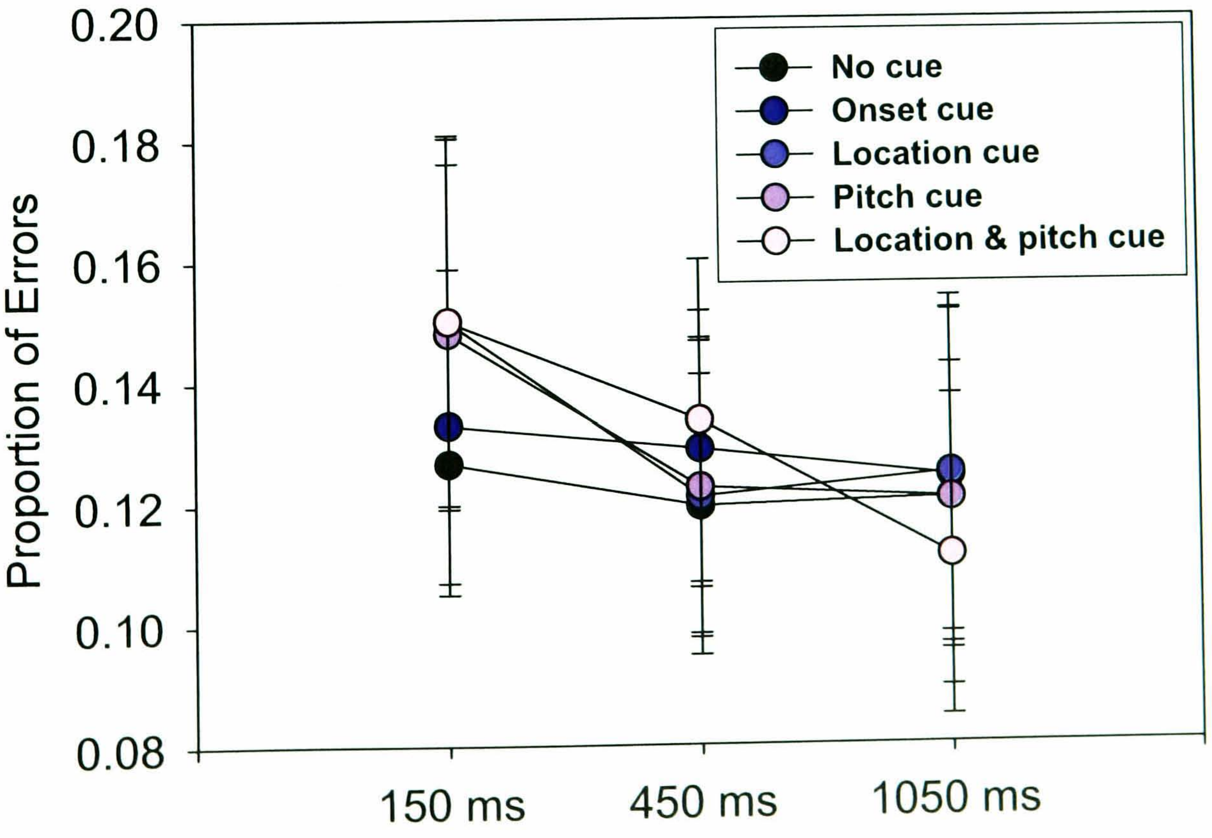
Figure 4.5: Alerting and orienting effects at each of the three SOAs. Error bars show 95% confidence intervals, uncorrected for multiple comparisons.



At the 150-ms SOA there was a significant cost associated with a cue to location & pitch (t_{39}=-4.425, p<0.001), but no other significant alerting or orienting effects. At the 450-ms SOA there was a significant alerting benefit (t_{39}=2.901, p<0.01), but no orienting benefits. At the 1050-ms SOA there were significant effects of alerting (t_{39}=6.704, p<0.001), orienting to a location (t_{39}=4.211, p<0.001), orienting to a pitch (t_{39}=4.598, p<0.001), and orienting to both location and pitch (t_{39}=3.676, p<0.01). Therefore subjects were able to gain a significant benefit from all three types of orienting cue at the 1050-ms SOA. The amount of benefit obtained is of a similar magnitude across orienting conditions, with no additive benefit from having cues to both location and pitch. This result is compatible with the conclusion that attention is being oriented towards an auditory object comprising both location and pitch information.

A two-way (SOA x cue type) ANOVA conducted on the error rates (Figure 4.6) revealed a significant main effect of SOA ($F_{2,78}$=9.520, p<0.001), but no significant main effect of cue type ($F_{4,156}$=0.962, p=0.430) and no significant interaction ($F_{8,312}$=1.439, p=0.179). Subjects made more errors at the shortest SOA, which in combination with the RT data suggests a speed-accuracy trade-off.

Figure 4.6: Proportion of errors at the three SOAs and with the five cue types. Error bars show 95% confidence intervals.



Correlations between RT alerting and orienting effects (Table 4.2) show significant positive correlations between the three orienting measures (orienting to location, pitch, and location & pitch) at the 450-ms and 1050-ms SOAs.

At the 450-ms SOA, the orienting measures correlated negatively with the alerting measure, indicating that the effects are not independent. The subtraction analysis evaluates orienting benefits over and above any alerting benefits conferred by the orienting cues. However, this analysis depends on the alerting and orienting measures being independent. The negative correlation indicates an interaction between the two, which suggests that the orienting measure is not reliable in this instance. At the 1050-ms SOA, alerting was not correlated with any of the orienting measures indicating that at this longest SOA the two effects may be independent. This increases confidence in the reliability of the subtracted values at this SOA.

Spatial Conflict

While the main aim of the Vowels study was to evaluate alerting and orienting, the design of the experiment also allows a measure of executive control to be evaluated, through an analysis of spatial conflict. When a left-hand response must be made to a stimulus lateralised to the right, subjects are typically slower to respond than if the stimulus and response are located on the same side (Craft & Simon, 1970).

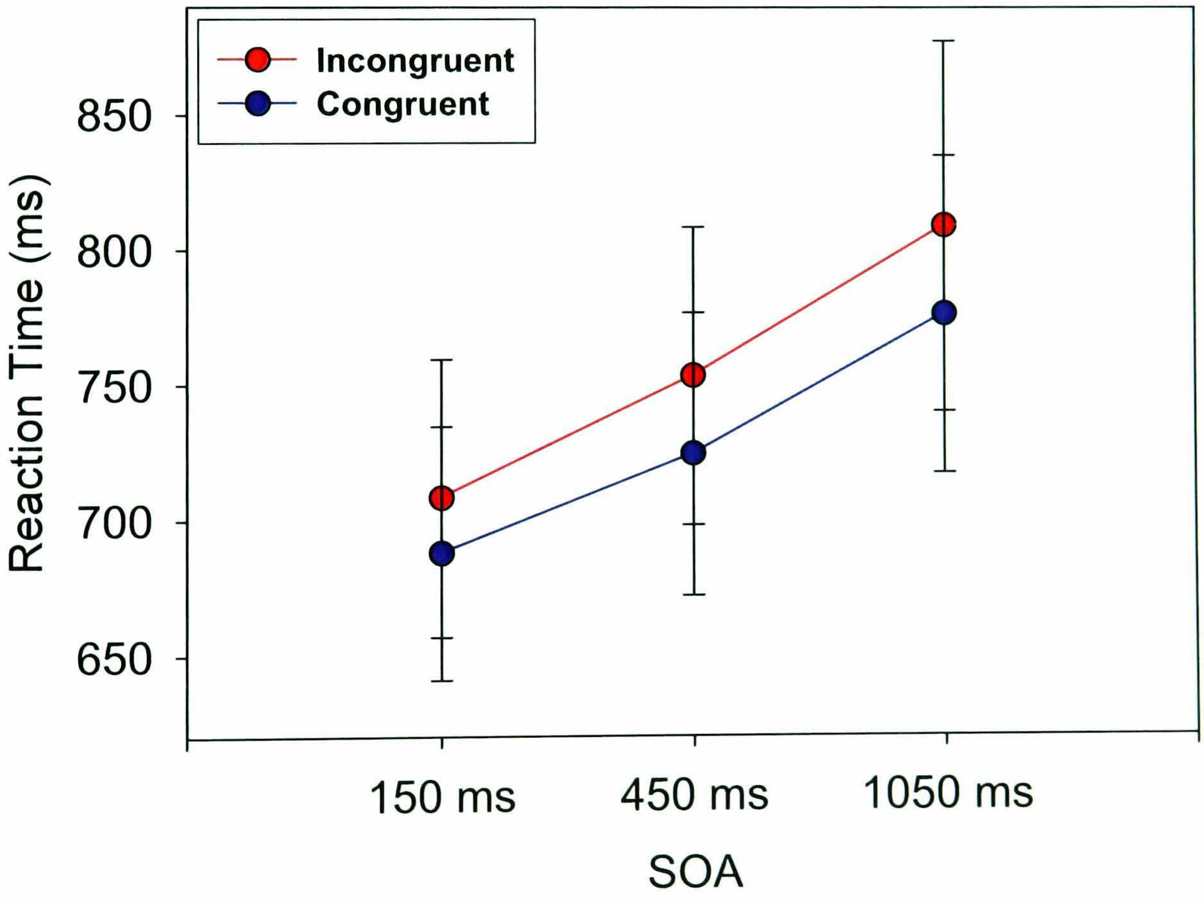
Table 4.2: Correlations between RT alerting and orienting measures at each of the three SOAs. *p<0.05, **p<0.0.1

		SOA: 150 ms				SOA: 450 ms				SOA: 1050 ms			
		Alert	Location	Pitch	Both	Alert	Location	Pitch	Both	Alert	Location	Pitch	Both
SOA: 150 ms	Alert	---											
	Location	-.049	---										
	Pitch	-.430**	.461**	---									
	Both	-.155	.116	.251	---								
SOA: 450 ms	Alert	.093	-.231	-.411**	.237	---							
	Location	-.004	.222	.306	-.424**	-.767**	---						
	Pitch	.153	.206	.306	-.394*	-.627**	.768**	---					
	Both	.103	.268	.228	-.482**	-.689**	.881**	.819**	---				
SOA: 1050 ms	Alert	.004	-.050	.158	-.080	.090	.142	.141	.133	---			
	Location	.065	.188	.179	.011	-.295	.298	.206	.230	-.038	---		
	Pitch	-.083	-.019	-.087	.196	.208	.018	-.061	-.105	-.134	.321*	---	
	Both	.186	-.036	-.204	-.192	-.184	.376*	.305	.354*	-.071	.334*	.471**	---

During the Vowels experiment, targets were lateralised to the left or to the right. Responses were made using a button box in which the response buttons were arranged in a left-to-right horizontal arrangement, in the order ‘ee’, ‘er’, ‘or’, from left to right. Spatial conflict was present when a left-sided response (‘ee’) was made to a stimulus lateralised to the right ear, or a right-sided response (‘or’) was made to a stimulus lateralised to the left ear. These two situations formed an incongruent condition. In the congruent condition, the stimulus and response were located to the same side (an ‘ee’ lateralised to the left, or an ‘or’ lateralised to the right). A neutral condition would have been provided by the vowel ‘er’ at either ear, since the response button was located centrally. However, RT data suggest that response latencies differed between the vowels due to their perceptual saliency, irrespective of their presented location. Average RTs were 671 ms to ‘ee’ (95% CI = 629 - 714 ms), 765 ms to ‘er’ (95% CI = 705 - 824 ms), and 822 ms to ‘or’ (95% CI = 753 - 891 ms). Because of these differences, responses to the vowel ‘er’ would not provide an unbiased neutral condition. However, a simple comparison between the congruent and incongruent conditions is possible because the vowels ‘ee’ and ‘or’ occur equally often in each condition.

A two-way ANOVA (SOA x condition) showed significant RT main effects of SOA ($F_{1.4,53.4}=33.677$, $p<0.001$) and condition ($F_{1,39}=12.951$, $p<0.01$) but no interaction. Overall RTs increased with increases in the SOA. RTs were slower to incongruent than to congruent stimuli at all SOAs (20 ms slower at the 150-ms SOA, 29 ms at the 450-ms SOA, and 33 ms at the 1050-ms SOA). The same analysis on the accuracy data also showed significant main effects of SOA ($F_{2,78}=3.389$, $p<0.05$) and condition ($F_{1,39}=6.540$, $p<0.05$), but no interaction. Error rates decreased with increases in the SOA, again suggesting a speed-accuracy trade-off. Accuracy was worse with incongruent stimuli than congruent stimuli at all SOAs (1.2% worse at the 150-ms SOA, and 1.5% worse at the 450- and 1050-ms SOAs). The experiment therefore included a successful auditory spatial conflict component, in which both RTs and accuracy were worse in the incongruent condition, irrespective of SOA.

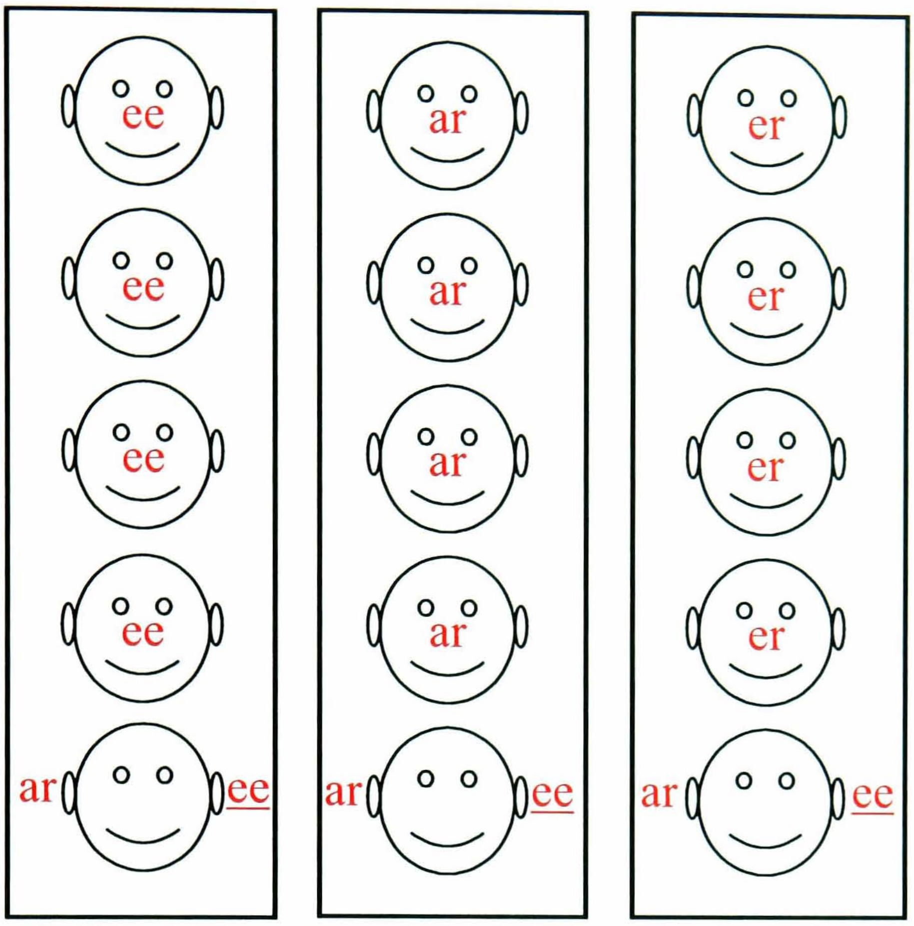
Figure 4.7: Reaction times on incongruent and congruent trials, at each of the three SOAs. Error bars show 95% confidence intervals.



Auditory perceptual 'pop-out'

Prior to target presentation, subjects heard one of the vowels repeated a random number of times (*fixation* vowels), in order to introduce uncertainty regarding target onset. The last of these fixation vowels was varied to provide cueing information (the *cue* vowel). The fixation vowel identity was selected at random, and could therefore be the same as the target vowel, the same as the distractor vowel, or different from both the target and the distractor. Electrophysiological investigations have demonstrated that the auditory system is highly sensitive to changes in a regular sequence of sounds. The mismatch negativity (MMN) is an event-related potential elicited by occasional 'deviant' sounds presented amid a larger number of 'standard' sounds (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). In this experiment, a behavioural correlate of the MMN is expected since subjects hear a sequence of vowel sounds. It is hypothesised that when either the target or distractor vowel identity differs from the fixation vowel identity, the deviant sound will 'pop out'. This automatic capture of attention by the deviant sound should lead to slower, less accurate, responses when the fixation and target vowels have the same identity, since the distractor vowel will pop out. Correspondingly, when the fixation and distractor vowels have the same identity, the target vowel is expected to pop out, leading to faster, more accurate, responses.

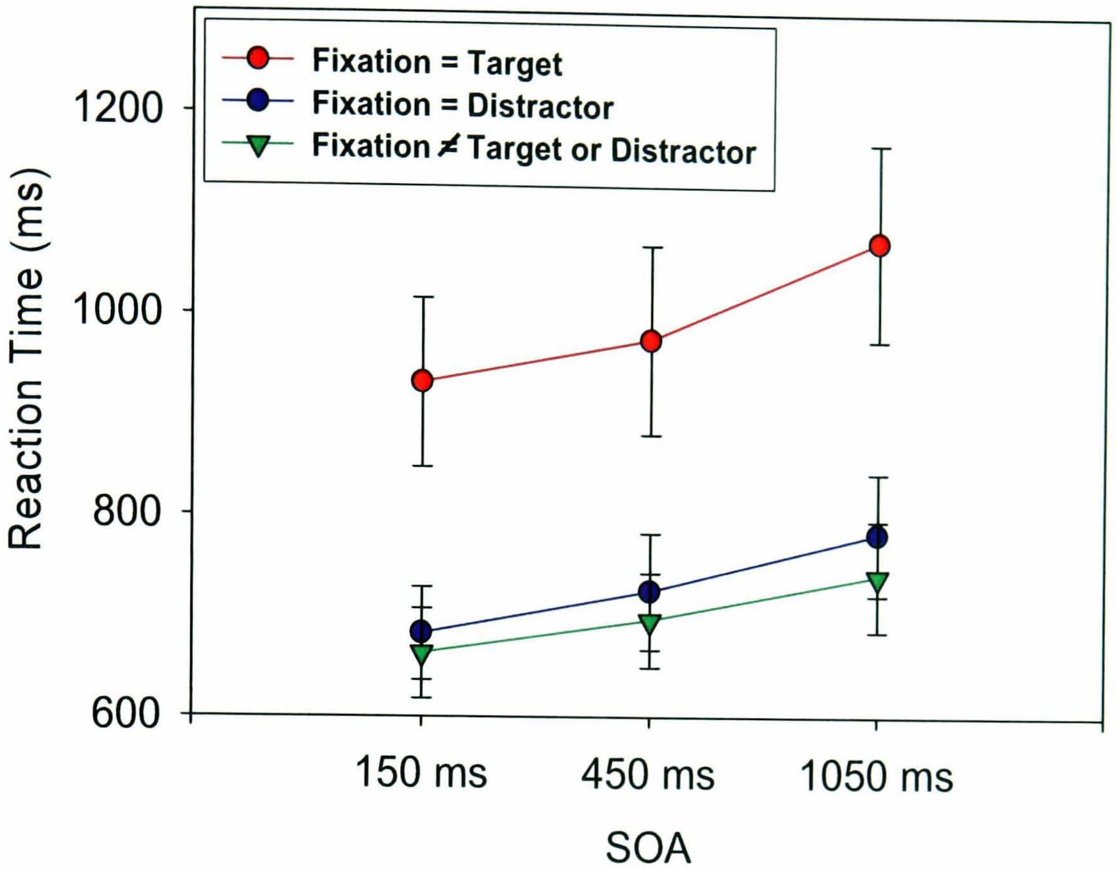
Figure 4.8: Illustration of fixation-target similarity conditions.



Fixation = Target **Fixation = Distractor** **Fixation ≠ Target or Distractor**

Figure 4.8 illustrates the three different types of relationship between the fixation and target. A two-way 3 (SOA: 150 ms, 450 ms, 1050 ms) x 3 (fixation-target similarity: fixation = target, fixation = distractor, fixation ≠ target or distractor) ANOVA was conducted on the RT data, which are plotted in Figure 4.9. The ANOVA revealed a significant main effect of SOA ($F_{1.4,54.0}=28.175$, $p<0.001$), and a significant main effect of fixation-target similarity ($F_{1.1,43.7}=110.143$, $p<0.001$). The interaction was not significant. As in the previous analyses, the main effect of SOA is due to longer RTs with longer SOAs. However, the most striking observation was the slow RTs for trials in which the fixation vowels had the same phonetic identity as the target vowel.

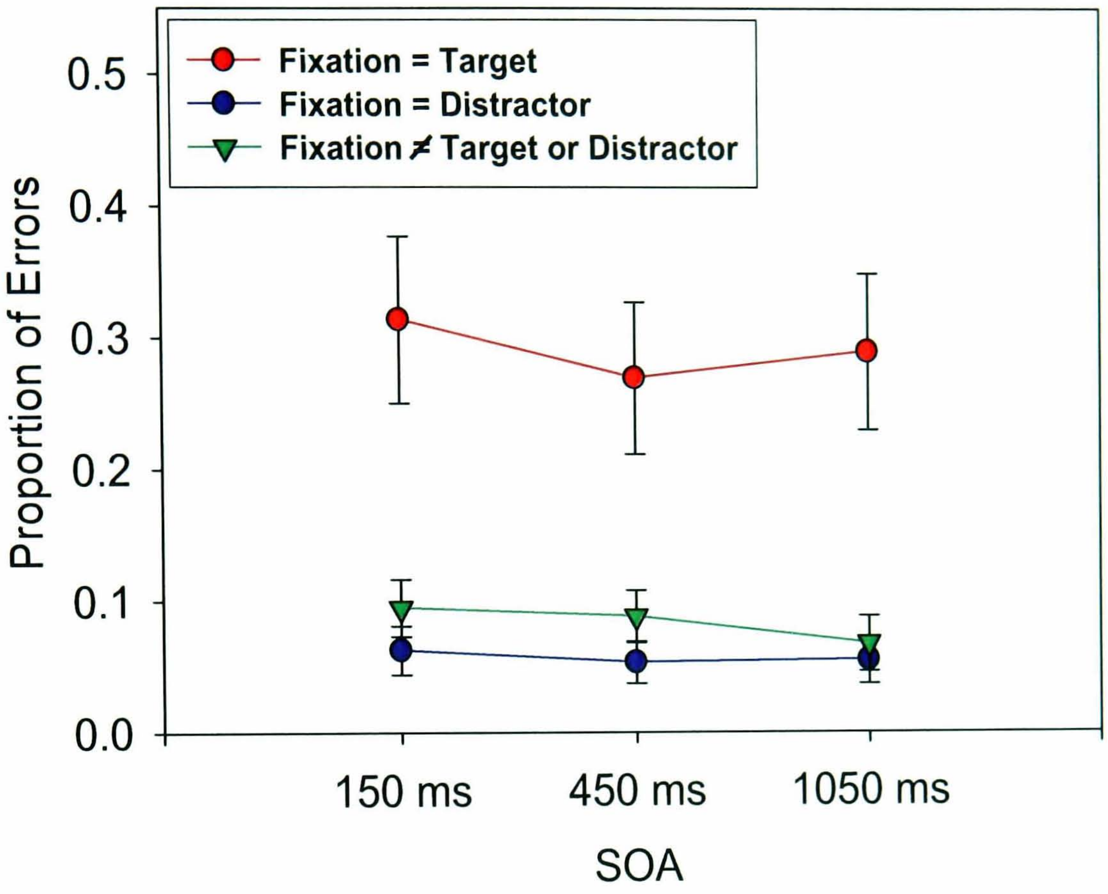
Figure 4.9: RTs to targets with the same vowel identity as the fixation vowels, the same vowel identity as the distractor, or a different vowel identity to the target and distractor. Error bars show 95% confidence intervals.



Since the interaction was not significant, RTs were collapsed across SOAs and a one-way ANOVA was conducted to compare responses across fixation-target similarity conditions. The ANOVA confirmed significant differences across conditions ($F_{1.2,139.3}$=252.096, p<0.001). Planned contrasts using t-tests and a Bonferroni correction ($p_{critical}$=0.05/3 = 0.0167) showed that subjects were significantly slower to respond when the fixation vowel was the same as the target vowel, compared with when the fixation vowel was the same as the distractor (t_{119}= 15.324, p<0.001) or when the fixation was different from both the target and distractor (t_{119}= 17.169, p<0.001). In addition, subjects were slower to respond when the fixation vowel was the same as the distractor, compared with when the fixation vowel was different from both the target and distract (t_{119}= 5.241, p<0.001).

The accuracy data (Figure 4.10) showed a significant main effect of SOA ($F_{2,78}$=7.365, $p<0.01$), and a significant main effect of fixation-target similarity ($F_{1.1,42.2}$=65.502, $p<0.001$). There was also a significant interaction ($F_{2.4,93.2}$=3.273, $p<0.05$). Error rates were higher at the shortest SOA, indicating a possible speed-accuracy trade-off. One-way ANOVAs comparing fixation-target similarity conditions at each level of the SOA ($p_{critical} = 0.05/3 = 0.0167$) showed significant differences with SOAs of 150 ms ($F_{1.1,43.1}$=60.025, $p<0.001$), 450 ms ($F_{1.1,44.8}$=47.475, $p<0.001$), and 1050 ms ($F_{1.1,42.9}$=68.195, $p<0.001$). T-tests with $p_{critical} = 0.05/3 = 0.0167$) showed significant differences between the three conditions at each level of SOA. An exception was at the 1050-ms SOA, where subjects made as many errors when the fixation vowel was the same as the distractor, as when the fixation vowel was different from both the target and distractor. This pattern of error rates differs from that found with RTs. In both the RT and accuracy data, subjects performed significantly worse (slower and less accurate) when the fixation vowel was the same as the target. However, when the fixation vowel was the same as the distractor, subjects were both slower and more accurate than when the fixation vowel was different from both the target and distractor. This suggests a speed-accuracy trade-off in one or both of these conditions.

Figure 4.10: Accuracy to targets with the same vowel identity as the fixation vowels, the same vowel identity as the distractor, or a different vowel identity to the target and distractor. Error bars show 95% confidence intervals.



Discussion

The design of the Vowels study was successful in eliciting auditory alerting effects at the 450-ms and 1050-ms SOAs, and spatial orienting effects at the 1050-ms SOA. Since the Vowels study differs from the auditory ANT in several respects it is not possible to identify the critical changes. However, three possibilities are: i) the SOA used during the auditory ANT (650 ms) was too short to allow participants to orient attention successfully; ii) subjects were able to alert and to orient to the onset of the lexical stimuli in the auditory ANT, attenuating cue-related alerting and orienting benefits; and iii) a competing stimulus may have been necessary in order to enhance spatial orienting benefits. Any one or combination of these possibilities could underlie the positive outcome in the Vowels study.

Effect of SOA

Overall, reaction times increased with increases in SOA. This appears to be primarily due to increased response latencies in the no-cue and onset-cue conditions. Since orienting benefits also increase with increases in SOA, overall RTs to trials with orienting cues do not appear to vary as a function of SOA. Since alerting and orienting effects are calculated by subtracting performance under different cueing conditions, they should not be influenced by overall changes in performance across SOAs. Error rates were higher at the shortest SOA, decreasing with increases in SOA. This pattern of results indicates a speed-accuracy trade-off, with subjects responding at the shortest SOA before they are confident that they have correctly identified the target vowel. There are two plausible causes of the speed-accuracy trade-off. First, subjects' speed of response may reflect the pace of the trial, with faster, but less accurate, responses when the pace of the trial is very rapid. Since the SOA was used to separate not only the cue and target stimuli, but also the fixation stimuli, a short SOA led to a rapidly-paced trial, while trials with the long SOA had a slow pace. Second, subjects may adopt a strategy of using the silence following the target to inform them of when to respond. At the shortest SOA they are able to make this judgment more quickly than at longer SOAs. The data from the current experiment do not allow us to distinguish between these two possibilities.

Subjects did not benefit from an *onset* cue at the shortest SOA. However, with longer SOAs they were significantly faster to respond when they were cued to target onset, compared with when they did not know when the target would occur. The size of this effect was greater at the 1050-ms SOA (87 ms) than at the 450-ms SOA (48 ms), suggesting that subjects were able to make better use of the alerting cue when they had more time between stimuli to do so. Posner (1978) found a different pattern of results. An auditory warning cue caused a greater reduction in RTs at a 150-ms SOA than at SOAs of 500 ms and 1000 ms, at which the alerting effects were of a similar magnitude. However, the large RT benefit at the 150-ms SOA observed by Posner was also associated with an increase in error rates. It therefore seems that the influence of warning cues does vary as a function of SOA, but that task demands and the subjects' desired level of accuracy also influence the size of this effect.

The pattern of orienting effects across SOAs is surprising, at least for the orienting-to-location measure. Typically, when a cue is presented at a target location it automatically draws attention for a short period of time (exogenous orienting). This should be reflected in cueing benefits at the 150-ms SOA (Spence & Driver, 1994). At longer SOAs, when subjects are provided with informative cues to target location, they benefit from voluntarily orienting their attention to the cued location (endogenous orienting) (Spence & Driver, 1994). While Mondor and Bregman (1994) investigated orienting to frequency across a range of SOAs, the shortest of these was 500 ms, and so it is not clear whether the same pattern of results would be expected in the orienting-to-frequency measure. There is no obvious explanation for the lack of exogenous orienting benefits at the 150-ms SOA. However, the speed-accuracy trade-off at this SOA does suggest that subjects were sometimes responding before reliably identifying the target vowel, and therefore that they may have been responding before benefiting from automatic cueing effects. Alternatively, the rapid rate of stimulus presentation at this SOA made the trial highly complex. Stimulus complexity has been shown to affect measures of attention (Lavie & Tsal, 1994).

Orienting to location, pitch, and location & pitch

At the longest SOA (1050 ms) subjects were quicker to respond with all three types of orienting cue than with a cue to target onset alone. The RT benefits from cues to location (49 ms) and pitch (57 ms) were of a similar magnitude, and there was no additive benefit from being cued to both location and pitch together (54 ms). Furthermore, all three orienting measures were positively correlated with each other. This suggests that subjects were attending to an auditory object comprising both location and pitch information, rather than attending to stimulus features independently.

At the 450-ms SOA there were no benefits from orienting cues over and above the benefit received from a cue to target onset. However, while all the orienting measures correlated positively with each other, they correlated negatively with the alerting measure. Since orienting cues also acted as alerting cues, specific measures of orienting benefits were obtained by subtracting out the benefit obtained from an alerting cue alone. However, this form of subtractive analysis relies on the two measures being independent. The negative correlations between the alerting and orienting measures indicate that at the 450-ms SOA this assumption of independence is violated. Since greater alerting benefits are associated with smaller orienting benefits, it is not possible to dissociate the two measures and determine the relative contribution of the alerting and orienting components. While there were no significant correlations between the alerting and orienting measures obtained in the ANT study (Chapter 2), there was a significant interaction between cueing and conflict conditions in the visual ANT (although not the auditory ANT). The SOAs were 500 ms in the visual ANT and 650 ms in the auditory ANT. The results of the ANT study combined with those of the Vowels experiment suggest that the relationship between alerting and orienting may vary as a function of the time between cue and target, with interactions between the two at SOAs of around 500 ms, which are no longer present at SOAs of around 1 second. However, further studies would be required to accurately assess this hypothesis.

At the 150-ms SOA cues to location and pitch separately did not significantly influence performance. However, there was a significant RT cost associated with a cue to both location and pitch. While the exact reason for this cost is unclear, a cue to both location and pitch would have sounded distinctly different from the fixation vowels. This may have been sufficient to introduce uncertainty over whether a target had been presented, even in the absence of the distractor vowel. This uncertainty would be reflected in RTs, particularly at the shortest SOA where there is less time to adapt.

Spatial conflict

The task contained a spatial conflict component, in which subjects were both slower to respond, and less accurate, when the stimulus appeared on the opposite side to the correct response button, compared with when the stimulus and response button were located on the same side. RT effects were modest (around 20 to 33 ms), but significant. Unfortunately, the neutral condition was not comparable to the incongruent and congruent conditions, since RTs to the neutral stimulus differed from those to the congruent and incongruent stimuli. However, this is a successful demonstration of an auditory spatial conflict task, which can be easily adapted to include stimuli matched for difficulty of identification in order to include a better-controlled neutral condition.

Fixation-target similarity

As hypothesised, subjects were substantially slower and less accurate when the fixation and target vowels had the same identity than when they did not. This effect was far larger (around 280 ms and 18% accuracy) than any of the orienting benefits (around 50 ms and 1.5% accuracy), and reflects an automatic effect whereby novel stimuli are highly salient when presented amid a sequence of repeated stimuli. Since the novel stimulus in this condition is the distractor sound, subjects are disadvantaged by its increased salience. The expected benefit when the fixation and distractor vowels had the same identity was not present. While there is a small increase in accuracy relative to the condition where the fixation vowel has a different identity from both the target and distractor vowels, this is also associated with a small RT cost. This pattern of results is difficult to interpret. However, it may be a function of the task demands, which required subjects to identify both the target and distractor in order to respond accurately. When the target was more salient, subjects may have been less sure that they had identified both of the concurrently presented vowels.

Future studies

The Vowels study measured a number of effects simultaneously. While it has revealed some interesting trends, additional experiments would be valuable for further investigating some of these trends. Of particular interest is the interaction between alerting and orienting effects at the 450 ms-SOA. The design of the Vowels study is easily adapted to investigate the interaction between alerting and orienting effects. Presenting a number of cue vowels, rather than fixation vowels followed by a single cue vowel, would enable the influence of orienting cues to be evaluated without the need to subtract alerting benefits from alerting-and-orienting benefits. By using a range of SOAs it would be possible to investigate how the relationship between attentional effects varies depending on the time available between cue and target.

The auditory pop-out effect was large and highly significant. However, the pattern of results was not quite as predicted. There was a substantial cost from distractor pop-out, but no clear benefit from target pop-out. A simpler experiment investigating this effect in isolation would be useful to establish whether this result reflects the need for subjects to identify both target and distractor vowels. A task in which the target is identified by its location or pitch, rather than by its difference from a distractor vowel might elicit different effects. Additionally, while the Vowels task incorporated a successful auditory spatial conflict task, this was flawed due to the lack of neutral condition. A specific auditory spatial-conflict task could include a valid neutral condition, which would allow interference and facilitation effects to be evaluated independently.

Conclusions

The main finding in the Vowels task revealed that subjects are able to benefit from cues to target onset, target location, target pitch, and both target location and pitch, at least at SOAs of around 1 second. This is an experimental paradigm that differed from the ANT, and in which it was possible to elicit reliable effects of spatial auditory attention. However, it also demonstrated some interdependence between alerting and orienting effects at the 450-ms SOA. The study incorporated an auditory spatial conflict task, which demonstrated both RT and accuracy costs for the incongruent condition relative to the congruent condition. Further, an auditory pop-out effect was present when the fixation vowel had the same identity as the target vowel: performance was substantially worse in this condition due to the increased salience of the distractor vowel. This reflects the salience of novel auditory stimuli.

Chapter 5: Spatial orienting of auditory attention: Effect of different cueing strategies

Introduction

Orienting attention to a spatial location facilitates processing of visual stimuli presented at that location (Fan et al., 2002; Posner, 1978). However, spatial orienting studies in the auditory modality produce variable results. Some experiments find that subjects are faster to process targets presented at a cued location, while other do not (McDonald & Ward, 1999; Spence & Driver, 1994). The auditory attention network test (ANT) reported in Chapter 2 showed that subjects received no benefit from a valid cue to spatial location. The Vowels task (Chapter 4) showed that subjects were able to benefit from a cue to spatial location, but only when the time between cue and target onsets was relatively long (1050 ms). This chapter investigates factors influencing spatial orienting of auditory attention through a review of the literature and a series of cueing experiments.

Visual information is coded spatiotopically, both on the retinae and in the cortex. This implies a special role for spatial location in visual information processing. Tootell et al. (1998) used fMRI to investigate the correspondence between regions of visual cortex which were active when visual stimuli were presented at a particular spatial location, and regions of the cortex which were active when covert visual attention was directed to the same spatial location. The results showed that there was indeed a correspondence between the two, indicating that attention operates on the spatial representation of the stimulus. Attention has been shown to operate via a number of mechanisms, including increasing the firing rate of neurons tuned to the attended location, influencing the size and shape of the cell's receptive field, increasing the specificity of the receptive field, and maintaining a constant response to an attended stimulus in the presence of distractors (see Motter (2000) for a review). Auditory information is coded tonotopically, and at present evidence suggests that while the superior colliculus contains a spatiotopic map of auditory space, there are no spatiotopic maps of auditory space in the cortex (Middlebrooks, 2000). Neurons have been found which are broadly tuned for space (Tian et al., 2001), but the evidence suggests that accurate spatial localisation is achieved through one of two mechanisms. One possible mechanism is that of two broadly tuned hemispheric channels, whereby localisation occurs based on the relative activity in these two channels (Boehnke & Phillips, 1999; McAlpine, Jiang, & Palmer, 2001). An alternative mechanism is that accurate localisation arises from the spike rate and firing pattern of cells broadly tuned for space (Middlebrooks, 2000). It is difficult to understand how auditory spatial attention might operate alongside either of these mechanisms in a manner comparable to that of visual attention.

While the exact manner in which auditory spatial attention may operate is unclear, subjective experience and results from early dichotic listening tasks (e.g. Cherry, 1953) strongly suggest that auditory attention can be directed to a spatial location. Since auditory information is initially coded tonotopically, and spatial location must be calculated from interaural and spectral cues, it has long been hypothesised that covert orienting of auditory attention can only occur when a spatially encoded auditory representation has been generated, such as during a localisation task (Rhodes, 1987). This hypothesis receives some support from a meta-analysis of fMRI studies comparing regions active during spatial and non-spatial auditory tasks (Arnott et al., 2004). Studies involving a spatial task were more likely to produce activation in the inferior parietal lobe (10 out of 11 studies) than those involving a non-spatial task (only 11 out of 27 studies), and conversely, non-spatial tasks were more likely to produce activation in the inferior frontal lobe (15 out of 27 studies) than spatial tasks (1 out of 11 studies). In addition, spatial tasks were associated with activation in the posterior temporal lobe, while non-spatial tasks produced activation that was distributed across the anterior and posterior temporal lobe. Arnott et al. (2004) suggested that this is evidence for ventral and dorsal 'what' and 'where' streams of processing, analogous to those found in the visual system (Mishkin et al., 1983). While the meta-analysis does not show a sharp segregation of activation into spatial and non-spatial regions, it does suggest some differentiation in processing depending on task demands. If attention were only able to operate upon spatial representations of the auditory stimuli, then spatial tasks might be necessary in order for spatial-orienting benefits to be obtained.

The suggestion that spatial encoding is necessary in order for auditory spatial orienting to occur was formalised by McDonald and Ward (1999) into the spatial relevance hypothesis. Based primarily on their own experiments, and those of Spence and Driver (1994), the spatial relevance hypothesis states that "the spatial location of an auditory target must be relevant to the accomplishment of the listener's task for ... spatial orienting to occur in audition" (p. 1236). The hypothesis extends Rhodes' (1987) original supposition by adding another condition under which space might be relevant to the task. With a non-spatial task, McDonald and Ward (1999) hypothesise that spatial relevance can be established by providing informative spatial cues which invite listeners to voluntarily orient their attention to a spatial location. In other words, by encouraging listeners to encode task stimuli spatially, spatial orienting benefits could be found in non-localisation tasks. Table 5.1 summarises the conditions under which spatial orienting benefits will and will not be found, according to the spatial relevance hypothesis.

Table 5.1: Conditions under which auditory spatial orienting benefits will (green ticks) and will not (red crosses) be found, according to the spatial relevance hypothesis (McDonald & Ward, 1999).

	Uninformative Cues	Informative Cues
Spatial Task	✓	✓
Non-Spatial Task	✗	✓

While the spatial relevance hypothesis accounts for much of the variation in the literature (reviewed below), it does not account for the failure of the auditory ANT to elicit spatial orienting benefits. The auditory ANT used informative spatial cues with a non-spatial task: conditions under which the spatial relevance hypothesis would have predicted spatial orienting benefits. The experiments reported in this chapter are designed to investigate this inconsistency. First, the existing literature on auditory spatial orienting is reviewed. Experiments which provide uninformative and informative cues, using spatial and non-spatial tasks, are considered first. There is then an overview of studies which investigate auditory attention using methodologies other than the cueing paradigm. Finally, the influence of stimulus presentation method (freefield or headphone presentation) on auditory spatial orienting effects is addressed.

Uninformative Cueing: Exogenous Orienting & Inhibition of Return

Uninformative cues are typically presented at one of the possible target locations. Following uninformative cues, the target is equally likely to occur at the cued and uncued locations (i.e. the cues are 50% valid, and 50% invalid). Uninformative cueing leads to a characteristic time course of visual attention (Klein, 2000). With short stimulus onset asynchronies (SOAs) of around 100 ms, subjects gain an advantage when the target appears at the cued location. This is exogenous orienting of attention, reflecting an automatic process by which attention is drawn to a cued location. At longer SOAs (typically greater than around 300 ms), subjects experience a cost when the target appears at the cued location. This is known as inhibition of return (IOR) (Klein, 2000; Prime & Ward, 2002). I will mostly restrict my review to the exogenous orienting effect, since it is the effect of interest in these studies. Facilitatory and inhibitory cueing effects are calculated by comparing performance following valid and invalid cues. It is worth noting that the difference between the valid and invalid cue conditions includes not only benefits obtained from orienting to the correct location, but also costs obtained from orienting to the wrong location. Since these two effects are only separable if a neutral cue has also been presented (and even then it is difficult to reliably calculate the relative contribution of each (Wright, Richard, & McDonald, 1995)), I will use 'benefit' to refer to improved performance following a valid cue, and 'cost' to refer to worsened performance following a valid cue, leaving it implicit that, relative to a neutral baseline, both benefits and costs are being measured.

Two seminal papers investigate spatial orienting of auditory attention: Spence & Driver (1994) and McDonald and Ward (1999). Each paper reports several experiments using different tasks and cueing conditions. Both papers include studies in which cues are informative and uninformative, and in which tasks are spatial and non-spatial. For their spatial discrimination task, Spence and Driver (1994) used an orthogonal cueing paradigm. In Experiment 1, the subject was seated between two horizontal arrays of three speakers, one to their left and one to their right. The central speaker in each array was located directly opposite each ear (+/- 90° azimuth). The remaining two speakers were located 28 cm in front or behind the central speaker. Subjects were cued (uninformatively) to their left or to their right by a cue presented from the central speaker of one of the arrays. A target sound was then presented from one of the front or back speakers, either on the cued side or on the uncued side. Subjects made a front / back discrimination. In this way, the cued direction was orthogonal to the task direction, avoiding a response-priming confound in which the cue also cues the response. With an SOA of 100 ms, subjects responded significantly faster (23 ms on average) if the target was presented from the same side as the cue (valid-cue trials), compared to when the target was presented from the opposite side (invalid-cue trials). Subjects did not receive a benefit or cost from valid cues at longer SOAs of 400 or 1000 ms. This pattern shows an auditory exogenous orienting effect, but no evidence of auditory IOR. Spence and Driver were concerned that the cue might act as a comparison point (or landmark) which might facilitate the front / back discrimination on valid-cue trials. To remove this confound, the task was changed to an up / down discrimination for Experiment 2. The speakers were arranged in a vertical column, rather than a horizontal array. Cues were 2000-Hz pure tones, which are difficult to localise in the vertical direction. Spence and Driver hypothesised that the difficulty judging the elevation of the cues would prevent them from providing a useful comparison point for judging target location. Targets were white noise, which is more easily localised. The new task revealed the same pattern of results as Experiment 1. Subjects gained a 26-ms advantage with valid cues at the 100-ms SOA, and no benefit or cost at SOAs of 400 or 1000 ms. In both experiments the results were reasonably consistent across subjects. At the 100-ms SOA, 19 out of 24 subjects were faster with valid cues in Experiment 1, and 10 out of 12 were faster with valid cues in Experiment 2.

Having successfully found exogenous spatial orienting effects with a spatial discrimination task, Spence and Driver then investigated whether it was also possible to find spatial orienting effects with a non-spatial discrimination task. In Experiment 3, subjects were cued to their left or right as in Experiments 1 and 2, but instead of the spatial discrimination task, they performed a frequency discrimination task. With this non-spatial task, subjects did not gain any benefits or costs from valid cues at any of the SOAs. In Experiment 7, the frequency discrimination task was made more difficult in order to equate overall RTs in the spatial and non-spatial tasks, but the experiment still gave rise to the same pattern of results: subjects did not receive benefits or costs from valid spatial cues in a non-spatial task, at any SOA. There are therefore two main findings from Spence and Driver's experiments with uninformative cues. First, exogenous spatial orienting effects were found with auditory spatial discrimination tasks, but not with non-spatial discrimination tasks. Second, IOR was not found with either spatial or non-spatial auditory discrimination tasks.

McDonald and Ward (1999) also used a cueing paradigm, but with a go/no-go task in which subjects either responded, or inhibited a response, depending on the target event. The go/no-go task was used instead of a straightforward discrimination task because McDonald and Ward had reservations about the orthogonal cueing paradigm used by Spence and Driver. Despite Spence and Driver's attempts to avoid the problem of the cue acting as a landmark on valid-cue trials, McDonald and Ward were still concerned that it might confer some advantage. In addition, in Spence and Driver's studies the cue was not presented from the same location as the targets. The cue was therefore cueing a hemifield, rather than a specific location. The mechanism by which attention is directed to a hemifield is not necessarily the same as that by which it is directed to a spatial location. McDonald and Ward therefore had three criteria for their spatial discrimination task: 1) that subjects would respond based on spatial location; 2) that cues and targets were presented from the same locations; and 3) that cues did not also cue a possible response. The go/no-go task, in which subjects make (or withhold) a single response based on spatial criteria, meets these three requirements.

In Experiment 1, subjects made an implicit spatial discrimination, responding to targets from peripheral speakers, but not to targets from a central speaker. The central speaker was located directly in front of the subject, while two peripheral speakers were located 37° to the left and right of the central speaker. Uninformative cues were presented 100, 300, or 700 ms prior to targets. Subjects responded 31 ms faster following a valid cue at the 100-ms SOA, and 23 ms slower following a valid cue at the 700-ms SOA, with no difference between performance with valid and invalid cues at the 300-ms SOA. This pattern demonstrates both exogenous orienting benefits and IOR. In Experiment 3, McDonald and Ward repeated this experiment, but used slightly longer SOAs, and monitored eye movements to ensure central fixation throughout the study. They found the same pattern of results: subjects were significantly faster to respond following valid cues at 100- and 500-ms SOAs (49 ms and 18 ms respectively), but 18 ms slower at the 900-ms SOA. To investigate whether spatial orienting effects would also be present with a non-spatial task and uninformative cues, in Experiments 2 and 5 subjects responded or withheld responses based on target frequency. Both experiments revealed no exogenous cueing benefits and no evidence of IOR. Experiment 7 used a different type of non-spatial task: a detection task. Again, there were no facilitatory or inhibitory spatial orienting effects. To exclude the possibility that the variation in frequency influenced results in the frequency-discrimination tasks, Experiment 4 used the same spatial discrimination task as Experiments 1 and 2, but with targets which randomly varied in frequency. Subjects were significantly faster to respond with valid cues at the 100- and 500-ms SOAs (47 and 22 ms respectively), but experienced a non-significant 12-ms cost at the 900-ms SOA. McDonald and Ward hypothesised that the lack of significant IOR might be caused by involuntary shifts of attention to the frequency of the targets, interfering with spatial orienting effects.

In summary, McDonald and Ward found exogenous orienting effects and IOR when subjects performed an implicit spatial discrimination task, but not when subjects performed non-spatial or detection tasks. In addition, IOR was not present when there was irrelevant variation in the frequency of the target, suggesting that automatic attention to frequency might influence spatial orienting effects.

A number of other studies have also investigated auditory spatial orienting with uninformative cues. Quinlan and Bailey (1995) presented cues and targets monaurally over headphones. Subjects performed a spatial discrimination task, pressing a left button to a left-ear target, and a right button to a right-ear target. When the cues were uninformative about target location, subjects were around 99 ms faster to respond with valid cues than with invalid cues at the 100-ms SOA. However, it should be noted that this design contains a response-priming confound, in which both the target and response are cued. Under these conditions, it is impossible to determine whether the cue influenced stimulus processing, or response preparation. Using a non-spatial task, Mondor and Breau (1999) presented cues and targets from speakers located directly in front of the subject, or 45° to the left or right of the central speaker. Subjects were asked to make a judgment about the rise time of the target stimulus (volume either increased slowly and then decreased sharply, or increased sharply and then decreased slowly). Cues were presented prior to targets, and were uninformative about target location. When cues and targets were presented from the same speaker, subjects were 28 ms faster to respond at the 150-ms SOA, and 26 ms slower at the 750-ms SOA, demonstrating both exogenous orienting and IOR in a non-spatial task with uninformative cues.

Crossmodal studies, in which auditory and visual spatial cues are used with auditory and visual targets, can provide useful information about auditory spatial orienting. These studies typically include an auditory cue / auditory target condition, which adds to the single-modality literature. In addition, crossmodal studies can provide information about subjects' ability to direct attention to the source of an auditory cue, and to process an auditory target at attended and unattended locations. Ward (1994) conducted a crossmodal study using auditory and visual cues and targets. The task was a spatial discrimination with a response-priming confound, similar to that used by Quinlan and Bailey. Uninformative visual cues produced both exogenous orienting and IOR effects with visual targets, but only exogenous orienting effects with auditory targets. Auditory cues produced no significant effects with visual targets, but both exogenous orienting and IOR effects with auditory targets. The key finding here is that auditory cues were unable to draw attention to visual target locations. This is in contrast to a crossmodal study by Spence and Driver (1997) (replicated by Schmitt, Postma, and De Haan (2000), Experiment 5) in which auditory cues drew attention to both auditory and visual targets, whereas visual cues drew attention to visual, but not auditory, targets. For clarity, Table 5.2 shows a summary of crossmodal cueing effects in these studies, and those discussed below.

Table 5.2: Summary of within- and cross-modality exogenous cueing effects found in crossmodal studies.

	Cue-Target Condition			
Study	Aud-Aud	Vis-Vis	Aud-Vis	Vis-Aud
Ward (1994) (spatial discrimination task)	✓	✓	✗	✓
Spence and Driver (1997) (spatial discrimination task)	✓	✓	✓	✗
Schmitt et al. (2000) (Exp 5: spatial discrimination task)	✓	✓	✓	✗
Ward et al. (2000) (spatial discrimination task)	✓	✗	✗	✓
Schmitt et al. (2000) (Exp 2: spatial discrimination task)	✓	✓	✓	✓
Mondor and Amirault (1998) (non-spatial discrimination task)	✓	✓	✗	✗
Schmitt et al. (2000) (Exp. 1: detection task)	✗	✓	✓	✗

In a later study, Ward, McDonald, and Lin (2000) used McDonald and Ward's implicit spatial discrimination (go/no-go) task in order to avoid a response-priming confound. In this study they replicated Ward's (1994) earlier finding that while visual cues could improve processing of auditory targets, auditory cues did not influence processing of visual targets. Surprisingly, they found within-modality effects in the auditory but not visual conditions, although they did find significant IOR effects in the visual conditions, and suggest that the early onset of IOR (the shortest SOA was 100 ms) might account for the lack of exogenous cueing effects. Ward et al. (2000) propose reasons for both their crossmodal results and those of Spence and Driver. In their own experiments the cueing environment was complex (since both cues and targets could be auditory or visual, and cues could be valid or invalid). Under these circumstances, Ward et al. suggest that the spatial location of the auditory cue is not fully processed, enabling the cue to act as an alerting cue, but not a spatial-orienting cue. This position is somewhat supported by Schmitt et al.'s (2000) Experiment 2, which used a very similar design, but with a less complex cueing environment. Auditory and visual cues and targets were presented, and subjects performed a spatial discrimination task with a response-priming confound. However, unlike Ward et al.'s studies, the trials were blocked so that each block contained only one combination of cue and target modalities. Under these (less complex) conditions, Schmitt et al. found exogenous cueing benefits with all four cue-target combinations. With respect to Spence and Driver's (1997) results, Ward et al. share McDonald and Ward's reservations regarding the cue and target being presented from different locations. Since the cue is presented from a central speaker, and targets are presented from speakers located above and below the cue location, a narrow focus of attention to the cued location would not be beneficial. While narrowly focused visual attention might not influence processing of distant auditory targets, auditory cues might engage distributed attention which would facilitate processing of distant visual targets.

Two further crossmodal studies investigated the effect of uninformative auditory and visual cues on non-spatial task performance. Mondor and Amirault (1998) used non-spatial discrimination tasks in which subjects either determined whether the target sound increased or decreased in frequency (an upward or downward frequency glide), or discriminated the colour of visual targets. The study revealed significant exogenous cueing effects for the within-modality conditions (auditory cues with auditory targets, and visual cues with visual targets), but no effects in the crossmodal conditions (auditory cues with visual targets, and visual cues with auditory targets). Schmitt et al. (2000) (Experiment 1) investigated within- and cross-modality cueing with a detection task. Subjects gained an advantage from valid auditory and visual cues when required to detect visual targets, but no advantage when detecting auditory targets.

To summarise, the pattern of results found with uninformative cues is less consistent than might be anticipated given the spatial relevance hypothesis. In some respects, the hypothesis seems accurate: spatial orienting benefits were reliably found when auditory cues preceded auditory targets in spatial tasks (McDonald & Ward, 1999; Quinlan & Bailey, 1995; Schmitt et al., 2000; Spence & Driver, 1994, , 1997; Ward, 1994; Ward et al., 2000). However, spatial orienting effects were also found when uninformative auditory cues were used with non-spatial auditory tasks (Mondor & Amirault, 1998; Mondor

& Breau, 1999), although not in all studies (McDonald & Ward, 1999; Spence & Driver, 1994). It therefore seems that while the spatial relevance hypothesis accounts for the spatial-task results, it does not seem to fully account for the results found with non-spatial tasks. The crossmodal literature can provide important information about whether attention can be directed to auditory cue locations, and whether spatial attention can facilitate processing of auditory stimuli presented at cued locations. However, the studies reported here do not provide consistent results. Auditory cues were found to facilitate processing of visual targets in some studies (Schmitt et al., 2000; Spence & Driver, 1997), but not others (Mondor & Amirault, 1998; Ward, 1994; Ward et al., 2000). Similarly, some studies found that visual cues facilitated performance on auditory tasks (Schmitt et al., 2000; Ward, 1994; Ward et al., 2000) while others did not (Mondor & Amirault, 1998; Schmitt et al., 2000; Spence & Driver, 1997).

Informative Cueing: Endogenous Orienting

Endogenous orienting occurs when subjects voluntarily orient their attention to a spatial location. Voluntary orienting would only be expected when the cue is informative about target location. This is typically achieved by increasing the probability that the target will occur at the cued location, from 50% in uninformative cueing studies to around 80% in informative cueing studies. Voluntary orienting of attention is a slower process than automatic orienting of attention, so effects are typically not found at very short SOAs (less than around 200 ms). However, because attention is under top-down control, spatial orienting benefits are found over much longer time courses than with exogenous orienting. Inhibition of return is not found when subjects voluntarily orient to cued locations. Endogenous orienting effects can be investigated independently of exogenous orienting effects by presenting symbolic cues, such as arrows, which indicate probable target location without being presented at that target location.

Spatial Tasks

Using their orthogonal cueing paradigm with a front/back discrimination task, Spence and Driver (1994) presented subjects with cues which were 75% valid (i.e. the target appeared on the cued side on 75% of trials, and on the uncued side on the remaining 25% of trials). The cue drew exogenous attention at the 100-ms SOA, at which subjects were on average 62 ms faster to respond following a valid cue compared with an invalid cue. At longer SOAs endogenous orienting effects were found. Subjects were 49 ms faster to respond at the 400-ms SOA, and 34 ms faster at the 1000-ms SOA. As with the uninformative cueing studies, most, but not all, subjects were faster to respond following a valid cue. Eleven out of twelve subjects were faster at the 100- and 400-ms SOAs, and ten out of twelve at the 1000-ms SOA. To investigate endogenous orienting effects independently of exogenous orienting effects, the next experiment presented cues which were 75% invalid. Here, subjects knew that the target was more likely to occur on uncued side than on the cued side. Surprisingly, subjects were able to prevent attention being automatically drawn to the cued location, and received no benefit or cost at the 100-ms SOA. At longer SOAs, responses were faster at the uncued location, where the target was most likely to occur. The effect was of a comparable magnitude to the 75% valid condition: 58 ms at the 400-ms SOA (in 9 out of 12 subjects), and 78 ms at the 1000-ms SOA (in 11 of 12 subjects).

Endogenous spatial orienting benefits have been found in a number of other studies. Bédard et al. (1993) used symbolic visual cues to indicate probable target location. Performance on an auditory spatial discrimination task (which included a response-priming confound) was improved following valid cues with an SOA of 500 ms. Quinlan and Bailey (1995) and Golob et al. (2002) used auditory symbolic cues and a spatial discrimination task (also with a response-priming confound), and found spatial orienting effects at SOAs of between 280 and 1800 ms. With a similar task, Schröger and Eimer (1996) found that informative peripheral cues improved performance at validly-cued locations with SOAs of 400, 600, and 800 ms.

Sach, Hill, and Bailey (2000) presented cue and target stimuli over headphones, lateralised using interaural time differences (ITDs). When sounds are lateralised using ITD cues alone, the perception is of sounds arising from inside the head, from positions approximately on an imaginary line which runs between the two ears. Since there is a limited impression of space, these sounds are said to be 'lateralised', rather than 'localised'. Sach et al. used ITDs to present stimuli which were heard at three lateralisations on each side of the head. This enabled them to design a task in which the response direction was orthogonal to the cue direction. Cues were 80% valid, and were presented from the centre position on each side of the head (cueing left or right). Two targets were then presented sequentially: from one or both of the remaining two locations on either the cued side (valid-cue trials) or uncued side (invalid-cue trials). Subjects made a same / different discrimination, based on whether the two targets were presented from the same location or different locations. Subjects underwent several training sessions in order to select ITDs which would lead to the percept of stimuli presented at three discriminable lateralisations on each side. Four subjects took part in the first experiment. Two of the subjects found the discrimination more difficult, and were allocated a larger ITD difference between the possible target positions than the other two subjects (200 μs rather than 150 μs). The results showed that the two subjects who performed poorly in training, and therefore had the larger ITD difference, did not gain an advantage on valid-cue trials compared with invalid-cue trials, while the remaining two subjects performed significantly better on valid-cue trials. Sach et al. hypothesised that the listeners who needed a larger ITD difference "were not using the information from the cue to full advantage" (p. 720). Across all subjects, the valid-cue benefit reached significance.

Sharing Spence and Driver's concerns about the cue acting as a landmark in valid cue trials, Sach et al. then changed the experimental design, and instead directed voluntary attention to one side by presenting 80% of targets to that side. All listeners gained an advantage on the expected side relative to the unexpected side. The expected side was switched every 25 trials, but there is still some possibility that the results are influenced by a practice effect on the attended side. In a final experiment (Sach et al., 2000), 80% valid symbolic visual cues (arrows) were presented. This design avoids both landmark and practice confounds. In this study, the overall effect of cue validity was significant. However, only 5 out of 9 subjects gained an advantage on valid-cue trials. Sach et al. offered the following explanation: “Because orienting was exclusively under voluntary control, it is possible that not all listeners successfully directed their attention to the lateral position indicated by the arrow, despite being instructed to do so. The lack of background noise or distractor signals renders this especially plausible: Except for adherence to the experimenter’s instructions, there was little need for a narrow focus of attention, and certain listeners may instead have adopted a wide-band listening strategy.” (p. 725).

Schmitt et al. (2000) presented informative (80% valid) auditory and visual cues, followed by auditory or visual targets, with SOAs of 125, 175, 225, and 575 ms. In the first of these experiments, a spatial discrimination task was used which included a response-priming confound. Orienting benefits were found with all cue-target combinations at SOAs of 125, 175, and 225 ms. However, orienting benefits were only found at the 575-ms SOA when visual cues were presented prior to auditory targets. This pattern of results suggests that exogenous orienting occurred with all cue-target combinations, but that subjects were only maintaining voluntary attention at the cued location in the visual cue / auditory target condition. In a later experiment, Spence and Driver’s orthogonal cueing paradigm (with an up/down discrimination) was used to avoid a response-priming confound. In this study, spatial orienting benefits were found at all SOAs in the visual task (with both auditory and visual cues). However, benefits were only found at the 125-ms SOA when auditory cues were used to cue auditory targets, and only at the 575-ms SOA when visual cues were used to cue auditory targets. There are two interesting things to note here. First, Schmitt et al. used a blocked design. The auditory cue / auditory target block was therefore very similar to Spence and Driver's (1994) Experiment 4 (informative cues; spatial task), in which significant cue benefits were found at SOAs of 100, 400, and 1000 ms. In Schmitt et al.'s study, an effect was only found at the 125-ms SOA, and not at SOAs of 175, 225, or 575 ms. Second, the results from Schmitt et al.'s informative cueing study are consistent with the results of Schmitt et al.'s uninformative cueing studies: exogenous orienting effects (at short SOAs) were found in all conditions except the visual cue / auditory target condition.

The spatial relevance hypothesis would strongly predict that experiments with informative cues and spatial discrimination tasks would find spatial orienting benefits. Space is made relevant by both the informative cues and the spatial task. While the studies described above broadly support the hypothesis, two results suggest that spatial orienting benefits are not highly robust. In two of Sach et al.'s experiments, spatial orienting benefits were only found in around half of the subjects, suggesting that spatial orienting effects can be variable across subjects. Further, using the orthogonal cueing paradigm with informative cues, Spence and Driver (1994) found cueing benefits, while Schmitt et al. (2000) did not. This suggests that the inter-subject variability might influence group results, even when the experimental paradigms are similar across studies.

Non-spatial Tasks

McDonald and Ward (1999) reported only one study with informative spatial cues: a go/no-go task in which subjects responded on the basis of target frequency. The cues were 80% valid, and produced spatial orienting benefits of 41 ms at the 100-ms SOA, 28 ms at the 400-ms SOA, and 34 ms at the 700-ms SOA. Schröger and Eimer (1997) used the frequency-based go/no-go task with informative symbolic visual cues and also found spatial-cue benefits. Posner (1978) also used predictive symbolic visual cues with a non-spatial task in which subjects only responded to 'weaker' stimuli. However, unlike the previous two studies, subjects gained no benefit from valid spatial cues, although spatial orienting effects were found in a comparable visual task.

Two experiments from the Spence and Driver (1994) paper investigated auditory spatial orienting with informative cues and non-spatial tasks. The first involved a frequency discrimination task with spatial cues which were 75% valid. Subjects were significantly faster to respond following valid cues, but the effects were smaller than those found with spatial tasks. On average, subjects were 22 ms faster at the 100-ms SOA (in 12 out of 16 subjects), 34 ms faster at the 400-ms SOA (in all 16 subjects), and 18 ms faster at the 1000-ms SOA (in 12 of 16 subjects). The remaining study used informative cues with a detection task, and found no significant benefits from valid spatial cues.

Detection tasks may constitute a special type of non-spatial task. Reaction times on detection tasks are substantially shorter than those on discrimination tasks, suggesting that subjects might be responding based on an early, non-spatial representation of the stimulus (Spence & Driver, 1994). Results from detection tasks with informative cues are particularly inconsistent. As stated above, Spence and Driver (1994) found no spatial orienting effects with informative cues on a detection task. Posner (1978) also found that while informative symbolic visual cues led to spatial orienting benefits in a visual detection task, there were no such benefits in an auditory detection task. Similarly, Hugdahl and Nordby (1994) presented 80% valid cues with an SOA of 800 ms in two detection tasks, one with visual cues and targets, and one with auditory cues and targets. While subjects were faster to respond following valid cues in the visual task, there were no orienting effects in the auditory task. Buchtel and Butter (1988) also presented visual or auditory cues which were 80% valid. With visual targets, both visual and auditory cues led to faster detection at SOAs of between 50 and 1000 ms. However, with auditory targets no effects were found with either visual or auditory cues, at any of the SOAs. Similarly, Schmitt et al. (2000) presented 80% valid auditory and visual cues with auditory and visual targets. In a detection task, a valid-cue advantage was only found in the visual cue / visual target condition.

While the studies reported above show spatial-orienting benefits in visual detection tasks, but not auditory detection tasks, other studies have found spatial-orienting benefits in auditory detection tasks. In an auditory task with only a short (100 ms) SOA, Quinlan and Bailey (1995) found a significant 9-

ms benefit from valid cues compared to invalid cues. Using a longer SOA (500 ms) that would reflect endogenous orienting, Bédard et al. (1993) found that informative symbolic visual cues led to response costs (slower responses on invalid trials compared with neutral trials), but not response benefits (faster responses on valid trials compared with neutral trials). Buchtel, Butter, and Ayvasik (1996) investigated conditions under which auditory spatial orienting effects might be found. They hypothesised that covert orienting effects might be found in situations where subjects were motivated to move one ear closer to the sound source, such as with low-intensity stimuli or with monaurally presented stimuli. Their results support this premise. With 80% valid cues and a detection task, they found spatial orienting benefits at short SOAs (less than 200 ms) with both high and low intensity stimuli presented monaurally, but only with low intensity stimuli presented in freefield (over speakers). They did not find spatial orienting effects in any of their conditions at longer SOAs (400 or 1000 ms), and suggest that "the endogenous component was weak in our tasks" (p. 984).

The spatial relevance hypothesis would predict that spatial orienting benefits would be found with non-spatial tasks, as long as informative cues are presented which make space relevant to the task. The hypothesis is supported by the majority of studies which presented a non-spatial discrimination task (McDonald & Ward, 1999; Schröger & Eimer, 1997; Spence & Driver, 1994). However, the results from Posner (1978) are not consistent with the hypothesis. Cued visual detection tasks consistently elicit spatial orienting benefits (Buchtel & Butter, 1988; Hugdahl & Nordby, 1994; Posner, 1978; Schmitt et al., 2000). However, cued auditory detection tasks sometimes elicited spatial orienting benefits (Bédard et al., 1993; Buchtel et al., 1996; Quinlan & Bailey, 1995), and sometimes did not (Buchtel & Butter, 1988; Buchtel et al., 1996; Hugdahl & Nordby, 1994; Posner, 1978; Schmitt et al., 2000; Spence & Driver, 1994). Since detection might be based on early, non-spatial representations, the spatial relevance hypothesis would not necessarily predict that spatial orienting benefits would be found.

Studies Using Different Methodologies

Some studies have investigated spatial orienting of auditory attention using methodologies other than the cueing paradigm. In general, these studies show benefits from orienting attention to a spatial location, but are designed to investigate the nature of auditory spatial attention, rather than to simply determine its presence or absence.

Darwin and Hukin (1999) used a variation of the dichotic listening task, and investigated whether listeners use interaural time difference (ITD) or fundamental frequency (f_0) information to attend to one of two concurrently presented sentences and ignore the other. Listeners were presented with two carrier sentences and two target words. For example, "Could you please write the word *bird* down now" and "You will also hear the sound *dog* this time", where italics indicate target words. The two sentences were presented together, with a delay to the shorter sentence so that the target words were presented simultaneously. Subjects were instructed to attend to one of the sentences. The carrier sentences could have the same or different f_0s and ITDs. The target words could share one, none, or both characteristics with the carrier sentence. Subjects were much more likely to select the correct target word if it shared an ITD with the carrier sentence, demonstrating that subjects were attending to ITD information. However, follow-up experiments demonstrated that subjects were unable to track components sharing a common ITD over time. The auditory system groups components based on other factors, such as common onset time and harmonicity, and then determines the location of the auditory object (Culling & Summerfield, 1995; Summerfield & Akeroyd, 1998). Darwin and Hukin concluded that subjects must be attending to objects which shared a perceived spatial location, rather than to a common ITD. In a later paper using the same design (Darwin & Hukin, 2000), other cues, such as prosody, pitch and vocal-tract size, were shown to influence listeners' decisions about which target word was associated with the attended carrier sentence. However, large differences in these other cues were required to overcome relatively small ITD differences.

Sach and Bailey (2004) also investigated the level at which auditory selection might operate. Interaural time and level cues are processed separately in the midbrain, and then combined at higher levels of processing. By combining ITD and interaural level difference (ILD) cues, it is possible to generate stimuli which share a common ITD or ILD, but are perceived at different spatial locations because of differences in the other dimension. If attention operates upon a low-level representation, then attention can be directed towards an ITD or ILD. If attention operates at a higher level, then it will be directed towards a representation comprising both interaural cues. Darwin and Hukin's results are consistent with the latter possibility, since subjects could not attend to frequency components which shared a common ITD, but instead attended to auditory objects which shared a perceived spatial location. Sach and Bailey (2004) used a 'rhythmic masking release' paradigm, in which subjects identified which of two target rhythms was being presented. The target rhythm was always presented with an ITD of 0, but could have an ILD of 0 dB (perceived at the centre of the head) or 4 dB (lateralised to the right). Irrelevant masker tones were presented with an ILD of 0 dB, but with a range of ITDs, so that they varied in perceived lateralisation. The results showed that performance was poor when the target rhythm and maskers shared the same spatial location (i.e. when both targets and maskers were perceived in the centre, or when both were perceived lateralised to the right). However, performance was improved when the target rhythm and maskers shared a common ITD, but different perceived location. This pattern of results, with more interference when the target and maskers share a perceived location than when they share a common ITD, indicates that it is the location that is being attended, not the ITD, and that therefore attention operates upon a higher-order representation of the sound.

Rhodes (1987) investigated whether auditory attention moves in an analogue manner. Specifically, she addressed the question of whether increased distance between successive attended locations was reflected in increased time required to make the attentional shift. Subjects were seated in a circular array of speakers, and asked to localise tones. On around a third of trials consecutive tones were presented from the same speaker, to encourage subjects to maintain their attention at the previous target location. Consistent with the analogue model, reaction times and accuracy worsened with increasing distance between the current and previous target locations. However, Rhodes' study has been subject to criticism for two reasons. First, localisation was performed by naming the speaker number from which the stimulus was presented. Since responding with one number would prime neighbouring, but not distant, numbers, there could be a response-priming confound (cf. Spence & Driver, 1994). Second, localisation performance is optimal for targets located directly in front of the subject, and declines with increasing azimuthal angle. Rhodes attempted to control for this both statistically and experimentally, but these attempts may not have been entirely successful (Mondor & Zatorre, 1995). Mondor and Zatorre (1995) evaluated Rhodes' conclusions using a non-spatial discrimination task, in which subjects decided whether they had heard a pure tone or a complex tone presented from speakers arranged in a semicircular array in front of them. Mondor and Zatorre argued that since auditory acuity for spectral judgments is unrelated to spatial position, this task would avoid the confound with localisation performance. A 'fixation sequence', in which subjects had to detect a drop in intensity in a pure tone, was presented prior to each trial to control the direction of attention. A 100% valid cue was then presented from one of the speakers, followed by the target tone. SOAs of 150, 600, 1050, and 1500 ms were used. The authors hypothesised that if auditory attention were moved in an analogue manner, as proposed by Rhodes, then there would be an interaction between the distance from the fixation speaker to the target speaker, and the time available between cue and target in which to move attention. In fact there was no interaction, which was taken as evidence that auditory attention does not move in an analogue manner.

Further experiments (Mondor & Zatorre, 1995) provided evidence for a gradient model of auditory attention, in which attention is most effective at the focus of attention, and declines gradually with distance from that point. Targets were preceded by spatial cues which were either valid or invalid. Following invalid cues, targets could be presented at locations close to the cued location, or far from the cued location. Subjects gained a significant advantage from valid spatial cues. In addition, performance on invalid-cue trials was better for targets presented close to the cued location than for targets presented further from the cued location, supporting a gradient model of auditory attention. Using a different methodology, Arbogast and Kidd (2000) failed to find evidence to support a gradient model of auditory attention. Their study used the probe-signal method, in which the majority of targets are presented from one location, with less frequent targets from different locations. It is difficult to separate attentional benefits from priming and practice effects using this methodology. However, subjects identified sounds presented at expected locations faster and more accurately than those presented at unexpected locations. Performance did not decline with increasing distance from the expected location, although Arbogast and Kidd suggest that this may be because the angular distance between their speakers (30°) was too large to detect a gradient of attention.

Influence of Stimulus Presentation Method

The majority of studies described above presented stimuli from speakers (in freefield), while the remaining studies presented stimuli over headphones, lateralised using monaural presentation or by varying ITDs and/or ILDs. When sounds are presented to one ear using monaural headphone presentation, it is possible that attention can be oriented to that ear, rather than to a spatial location. This might also be possible with freefield presentation in which sounds are presented from speakers located directly opposite each ear (+/-90° azimuth). Spatial orienting effects might be more readily found when a 'better ear' strategy is possible than when a genuine spatial location must be attended, given that auditory information is not coded spatiotopically in the cortex.

When stimuli are presented in freefield, ITD, ILD and spectral cues are available with which to localise the sound. Similarly, stimuli can be presented over headphones using ITD and/or ILD cues, and additionally, spectral cues can be simulated using head-related transfer functions (HRTFs). Under these circumstances, when localisation is achieved using binaural information, a 'better ear' strategy would be less effective. However, a greater wealth of cues are available with which to localise the target stimulus, and studies using this type of presentation can investigate attention to a spatial location, rather than to an ear. An attention system primarily aimed at orienting to visual locations might be better able to operate on this type of auditory representation.

Buchtel et al. (1996) conducted the only study to compare spatial-orienting effects with different presentation methods (although in different groups of subjects). Buchtel et al. hypothesised that covert auditory orienting might be related to overt auditory orienting. Visual overt orienting is achieved by moving the head or eyes to the attended location, and it has been shown that overt and covert visual orienting operate using common cortical regions (Corbetta et al., 1998). The mechanism by which overt auditory orienting might occur is less clear. There are essentially two options available when orienting towards a sound source: 1) to move the head towards the sound source in order to fixate it visually, and 2) to move one of the ears closer to the sound source. Buchtel et al. proposed that since two contradictory options are available, covert orienting could only occur when one of the options was clearly advantageous. They then tested two scenarios in which they proposed that the clear overt orienting movement would be to move the closest ear nearer to the stimulus. The first scenario was that in which stimuli were presented at a low intensity, and therefore only audible at one ear, an example of which might be attempting to hear whether a watch is ticking. The second scenario was that in which stimuli were presented monaurally, irrespective of their intensity. The rationale for this was that natural sounds arriving at one ear only were always of low intensity, and therefore the overt movement described in the first scenario would still be applicable. The results from their studies support the hypothesis: spatial orienting benefits were found when stimuli of high and low intensity were presented monaurally, but only when low-intensity stimuli were presented in freefield.

The limited number of auditory-attention studies makes it difficult to evaluate the influence of stimulus presentation method as so many other factors also vary between studies. The majority of studies have presented stimuli in freefield: some from speakers directly opposite each ear (e.g. Spence & Driver, 1994, , 1997), others from speakers at less than 45° azimuthal angle (e.g. McDonald & Ward, 1999; Mondor & Breau, 1999). Other studies have presented stimuli over headphones, usually monaurally (e.g. Bédard et al., 1993; Quinlan & Bailey, 1995), but also using ITDs to lateralise the sounds (Sach et al., 2000). There is no clear indication that any one methodology is more effective than the others.

Summary and Proposed Experiments

The auditory ANT (Chapter 2) failed to elicit spatial orienting benefits. While this might have related to methodological issues, such as subjects being able to alert and orient to the onset of the linguistic stimuli, or the SOA not being optimal, a review of the literature reveals several other inconsistencies across studies. The spatial relevance hypothesis states that auditory spatial orienting benefits will be found when a spatial representation of the task is available to orient towards, and that for a spatial representation to be generated space has to be relevant to task performance. Space can be relevant either through a spatial discrimination task, or through informative spatial cueing, which encourages participants to voluntarily orient their attention to a target location. The suggestion therefore is that uninformative cues will only elicit auditory spatial orienting benefits in a spatial discrimination task, while informative cues will elicit orienting benefits in both spatial and non-spatial discrimination tasks. Spatial orienting benefits were reliably found with uninformative cues and spatial tasks, and found in the majority of studies which presented informative cues and spatial tasks, although there was some suggestion of high inter-subject variability. The majority of studies presenting informative cues and non-spatial tasks did find spatial orienting effects. However, while no orienting effects would be predicted with uninformative cues and non-spatial tasks, in fact some studies did find spatial orienting benefits, while others did not. Evidence from crossmodal studies using auditory and visual cues and targets is currently highly variable. Further, spatial orienting effects are inconsistently found in auditory detection tasks. Overall, the spatial relevance hypothesis accounts for a great deal of the variability in the literature. However, it does not appear to be entirely comprehensive as it is unable to account for all experimental outcomes. In particular, it does not explain the high variability sometimes found between subjects and between studies.

In order to investigate spatial orienting of auditory attention, and to identify reasons for the null result in the auditory ANT (Chapter 2), a series of cueing experiments was conducted. The aim of these studies was to determine factors which influence auditory spatial orienting effects in non-spatial tasks, and to directly compare different stimulus presentation methods. The first two experiments specifically address the null result in the auditory ANT - the first uses the auditory ANT stimuli and cueing protocol, but with three different SOAs to investigate the time course of cueing effects. The second experiment uses brief steady-state stimuli with the ANT cueing protocol to investigate whether the linguistic stimuli influenced the null cueing effects found in the auditory ANT. Experiment 3 increases the proportion of trials in which a spatial cue is provided, in order to increase the salience of the spatial nature of the task. Experiment 4 uses invalid cues rather than neutral cues as a baseline, to match the cueing protocols of Spence and Driver (1994) and McDonald and Ward (1999). Experiments 5 and 6 separate exogenous and endogenous cueing effects. Experiment 5 presents uninformative cues, to elicit exogenous, but not endogenous, spatial orienting. Conversely, Experiment 6 uses centrally-presented symbolic cues to elicit endogenous, but not exogenous, spatial orienting. Each of these experiments is reported separately, but graphs showing results from the first six experiments can be found in the interim discussion on page 205. Finally, Experiment 7 compares spatial orienting effects with monaural, freefield, and binaural ITD presentation methods.

Experiment 1: ANT stimuli and cueing protocol

Experiment 1 addressed some of the potential methodological problems with the auditory ANT. While keeping all other factors constant, the fixation tone was removed, to eliminate the possibility that it was distracting or re-orienting subjects following cues. Additionally, performance was tested at three different SOAs (in separate blocks) to investigate the timecourse of spatial orienting. By making these changes but in all other ways replicating the auditory ANT, it is possible to determine whether the fixation tone had an adverse effect on spatial orienting effects, and whether the SOA used during the auditory ANT (650 ms) was inappropriate for assessing auditory orienting benefits. This might be the case if only exogenous attention effects were present (detectable only at shorter SOAs) or if endogenous attention effects took longer to build (detectable at longer SOAs). Experiment 1 therefore contained four cue types, in equal proportions: no cue; alerting (centre) cues; alerting (double) cues; and 100% valid spatial cues. Target stimuli were the words ‘high’, ‘low’, and ‘day’ spoken on a high and low pitch, and the task was to identify the pitch of the voice whilst ignoring word meaning. Subjects participated in three experimental blocks: one at each SOA (150, 450, and 750 ms).

Method

Subjects

Twelve volunteers (3 male, mean age 21.8 years, range 18 - 41) participated in Experiment 1. Participants all spoke English as their native language, gave informed consent prior to the study, and received £5 compensation. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that all subjects had normal hearing (thresholds below 20 dB HL).

Apparatus and Stimuli

Testing was conducted in a sound-attenuating chamber. Stimuli were presented under the control of a Visual Basic programme implemented at the MRC Institute of Hearing Research on an IBM-compatible personal computer running Windows 98. Stimuli were presented via Sennheisser HD 480II headphones, at a clearly audible volume in the range 75 to 80 dB(A). Subjects responded by pressing buttons on a response box. The buttons were arranged in a front to back vertical array to allow subjects to press the higher (further) button for 'high' and the lower (nearer) button for 'low'. Cue and target stimuli were those used for the auditory ANT, as reported in Chapter 2. Cues were speech-shaped noise cues, presented diotically (heard in the centre of the head), dichotically (heard to both sides, or as a diffuse sound source), or monaurally to the left or right. Target stimuli were the words 'high', 'low', and 'day' spoken on a high or low pitch.

Procedure

Subjects heard the target words 'high', 'low', and 'day' spoken on a high or low pitch and were asked to ignore the semantic meaning of the word and respond to the pitch of the voice. Target words were presented to the left or right ear, monaurally. Prior to the target word, subjects heard one of four cue types. In the *no cue* condition subjects only heard the target word. In the *centre cue* and *double cue* conditions, subjects heard cues which alerted them to target onset. A *spatial cue* both alerted subjects and indicated the side on which the target would be presented (100% valid). The stimulus onset asynchrony (SOA) was either 150 ms, 450 ms, or 750 ms. Between trials there was a random period of silence, of between 2400 and 3600 ms.

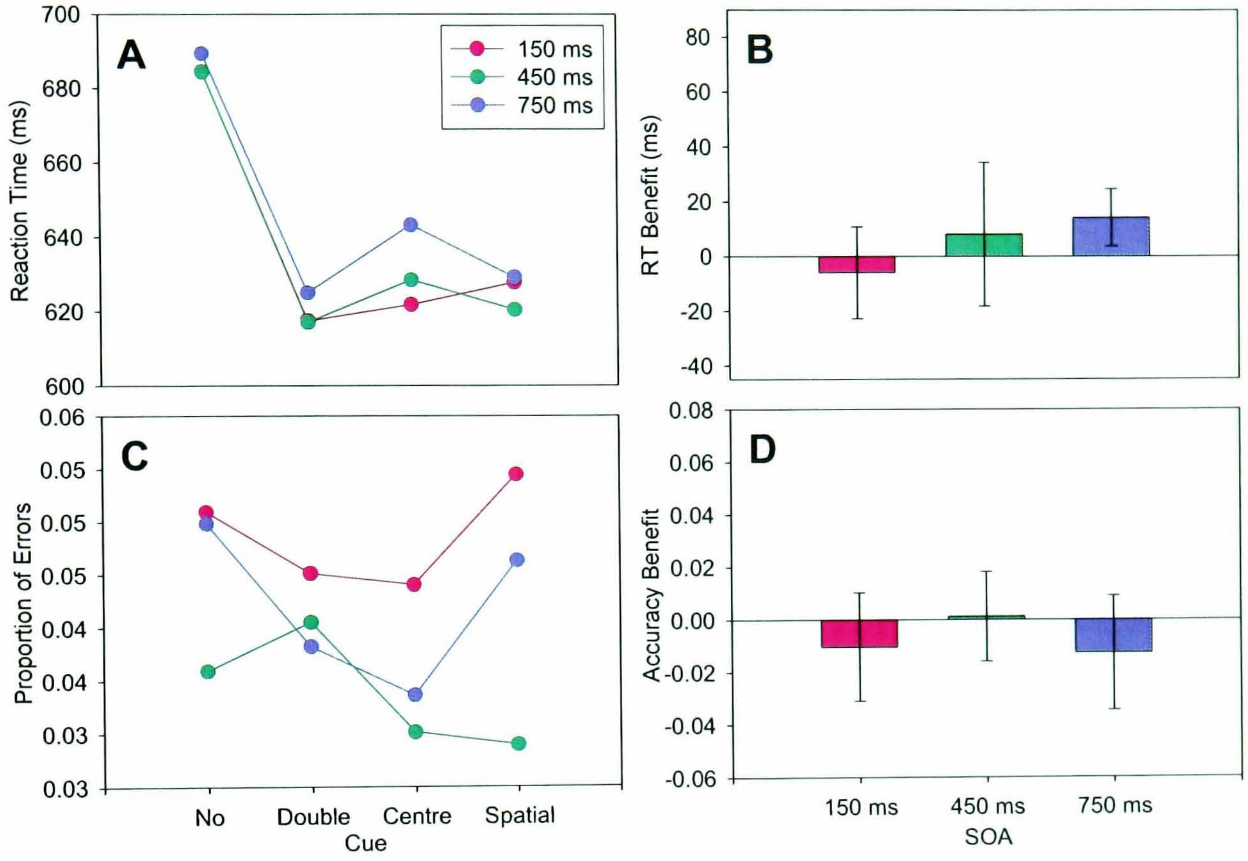
Subjects were initially presented with a 24-trial practice session in which they responded to the target words without any cues. Feedback was provided on whether they responded correctly. Subjects then had a 48-trial practice session which included cues, still with feedback. If subjects experienced difficulty with the practice sessions they were able to repeat them until they felt confident that they could do the task. Following the practice sessions, subjects took part in three experimental blocks, one with an SOA of 150 ms, one with an SOA of 450 ms, and one with an SOA of 750 ms. Each block contained 288 trials (3 words x 3 examples of each word x 2 pitches x 2 locations x 4 cue types x 2 repeats). The order of the blocks was counterbalanced across subjects. During the experimental blocks no feedback was provided. Subjects were instructed to respond as quickly and as accurately as possible. Each experimental block lasted approximately 17 minutes.

Results

RTs from correct trials were trimmed to exclude responses faster than 100 ms and slower than 2500 ms, resulting in the removal of 1.28% of responses. Median values were then computed for each subject, for each condition. Two analyses were performed on the RT data and error rates (Figure 5.1, panels A and C). First, t-tests were used to evaluate alerting and orienting effects at each level of the SOA, to enable a comparison with the auditory ANT experiment reported in Chapter 2. Second, a two-way two by three ANOVA was performed to investigate spatial cueing effects (centre cue, spatial cue) at each level of the SOA (150, 450, 750 ms). Where Mauchley's test of sphericity indicated that sphericity could not be assumed a Greenhouse-Geisser correction was applied. This is evident from non-integer degrees of freedom. Planned contrasts were performed using t-tests with a Bonferroni correction: contrasting centre-cue and spatial-cue performance at each level of the SOA. There were therefore 3 planned contrasts, giving a critical p value of $0.05 / 3 = 0.0167$. The ANOVA analysis was performed to enable a comparison with the experiments reported later in this chapter, and with the results reported by Spence and Driver (1994). Benefits from a spatial cue compared with a neutral (centre) cue are shown in panels B and D of Figure 5.1. All graphs showing spatial-cue benefits in this chapter show 95% confidence intervals which are not corrected for multiple comparisons. This is to provide maximum transparency given that different methods for correcting for multiple comparisons vary in their stringency. RT benefits from the first 6 experiments reported in this chapter are shown in Figure 5.8 (group data) and Figure 5.9 (individual subject data).

Subjects were significantly faster to respond following an alerting (double) cue than with no cue at all three SOAs (150 ms: $t_{11}=4.062$, $p<0.01$; 450 ms: $t_{11}=9.435$, $p<0.001$; 750 ms: $t_{11}=5.909$, $p<0.001$). The effect was of a similar magnitude across SOAs: 67 ms at the shorter SOAs and 64 ms at the 750-ms SOA. Subjects were also significantly faster to respond following a spatial cue than with a neutral (centre) cue at the 750-ms SOA ($t_{11}=2.951$, $p<0.05$), but not the other SOAs (150 ms: $t_{11}=-0.779$, $p=0.452$; 450 ms: $t_{11}=0.674$, $p=0.514$). Although the effect at the 750-ms SOA was significant, on average subjects were only 14 ms quicker to respond. There were no significant alerting or orienting effects in the accuracy data.

Figure 5.1: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a spatial cue compared with a neutral (centre) cue. Error bars show 95% confidence intervals.



The two-way 3 (SOA: 150 ms; 450 ms; 750 ms) by 2 (cue: centre cue; spatial cue) ANOVA on the RT data did not show any significant effects (Table 5.3), but planned comparisons revealed a significant valid-cue benefit at the 750-ms SOA (t_{11}=2.951, $p<0.0167$). In the accuracy data there was a significant main effect of SOA, but no other significant effects. Error rates were highest at the 150-ms SOA (4.92%), followed by the 750-ms SOA (3.99%), and lowest at the 450-ms SOA (2.95%).

Table 5.3: ANOVA results from the RT and accuracy data. The table shows main effects of SOA (150, 450, 750 ms) and cue type (centre cue, spatial cue), and the interaction between SOA and cue type.

	SOA	Cue	SOA x Cue
RT	$F_{2,22}=0.144$, *n.s.*	$F_{1,11}=0.903$, *n.s.*	$F_{1.2,13.5}=1.663$, *n.s.*
Accuracy	$F_{2,22}=3.464$, $p<0.05$	$F_{1,11}=3.897$, *n.s.*	$F_{2,22}=0.542$, *n.s.*

Discussion

The auditory ANT (Chapter 2) revealed a 36-ms alerting benefit and a non-significant 10-ms orienting benefit, with an SOA of 650 ms. This study, which used the same stimuli and cueing conditions, but without a fixation tone and at SOAs of 150, 450, and 750 ms, revealed significant alerting benefits at all three SOAs (67, 67, and 64 ms, respectively), and orienting benefits of -6, 8, and 14 ms, respectively. The orienting benefit was significant at the 750-ms SOA, but not at the shorter SOAs. In addition, there was no overall effect of cue type in the cue type by SOA ANOVA analysis.

Compared with the auditory ANT, the present study produced a much larger alerting effect, but a similar sized orienting effect. This suggests that the fixation tone might have attenuated the alerting benefit of the double cue, but did not have a substantial re-orienting effect following a spatial cue. The alerting benefit was of a consistent magnitude across SOAs, while the orienting effect increased slightly across SOAs, but in both cases there is no evidence that the SOA selected for the auditory ANT was inappropriate. Given the increasing size of the orienting benefit, it is possible that larger spatial orienting effects might have been found at longer SOAs, but other cueing studies (e.g. McDonald & Ward, 1999) were able to find spatial orienting effects at similar SOAs.

Unlike the auditory ANT, the present study did reveal significant RT benefits from a cue to spatial location. However, on average subjects only gained a 14-ms benefit - not substantially larger than the non-significant 10-ms benefit found in the auditory ANT. Interestingly, RTs on spatial-cue trials were similar to those on double-cue trials at the 450- and 750-ms SOAs. The difference between double-cue and spatial-cue response times was -3 ms with the 450-ms SOA, and -4 ms with the 750-ms SOA, while responses following a centre cue were 11 and 18 ms slower than those following a double cue, at the 450- and 750-ms SOAs respectively. This pattern suggests that the centre cue was drawing attention away from the target locations, while the double cue was keeping attention relatively diffuse across the two target locations. The cues did not appear to function in this way in the original auditory ANT, in which RTs with no cue, double cue, single cue, and spatial cue were 686, 650, 650, and 641 ms, respectively. Similarly, RTs in the visual task were 592, 552, 559, and 510 ms, respectively. The implication is that in this experiment, without the fixation tone, subjects were able to divide their attentional focus across the two possible target locations, or diffusely across target locations. However, when a centre cue was presented, it drew attention to a non-target location, slowing responses relative to the double-cue and spatial-cue conditions.

Experiment 2: Steady-state stimuli, ANT cueing

Experiment 1 found a small but significant benefit from spatial cues. Additionally, there were large and significant benefits from an alerting cue. Experiment 2 investigates the possibility that alerting and orienting benefits were attenuated in both Experiment 1 and the auditory ANT by the use of linguistic stimuli. Since the target stimuli vary in their spectro-temporal structure, it is possible that subjects were able to alert and to orient to the onset of the target word, before sufficient information was available on which to respond. Experiment 2 presents steady-state target stimuli (which have a constant spectro-temporal structure). The task in this experiment is to decide whether the target stimulus was harmonic or inharmonic. Harmonic stimuli contain harmonically-related frequency components, and therefore have a clear pitch. Inharmonic stimuli contain harmonically-related frequency components, with one component which is mistuned so that it is no longer harmonically related. Inharmonic stimuli sound as though two sounds with different pitches are being presented concurrently. The harmonic/inharmonic discrimination was selected because it is a non-spatial task in which a cue cannot aid performance by acting as a landmark. In all other respects, Experiment 2 is a replication of Experiment 1.

Method

Subjects

Twelve volunteers (3 male, mean age 22.4 years, range 18 - 33) participated in Experiment 2. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that all subjects had normal hearing (thresholds below 20 dB HL). All participants spoke English as their native language, gave informed consent prior to the study, and received £5 compensation.

Apparatus and Stimuli

The experimental set-up was the same as for Experiment 1, but with changes to the task and stimuli. Subjects performed a two-alternative forced-choice task with brief steady-state stimuli. The four cueing conditions were the same as in Experiment 1. Target stimuli were harmonic and inharmonic complex tones. Subjects responded by pressing buttons on a response box containing buttons arranged in a left to right horizontal array. Subjects pressed the button on the left if they heard a harmonic complex tone, and the button on the right if they heard an inharmonic complex tone.

Target stimuli were 100-ms complex tones. To prevent subjects listening out for a particular pitch or fundamental frequency, rather than making a general judgment about the harmonicity of the complex tones, five complex tones were generated, each comprising ten integer multiples of a different fundamental frequency. The fundamental frequencies were 190, 195, 200, 205, and 210 Hz. Harmonic complex tones contained equally-spaced components. For example, the 200-Hz harmonic complex tone contained components at 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, and 2000 Hz. Inharmonic complex tones contained one mistuned component. This was the third component, which was altered by +/-79 Hz, so that, for example, the 200-Hz inharmonic complex tone contained components at 200, 400, 679, 800, 1000, 1200, 1400, 1600, 1800, and 2000 Hz. The inharmonic component is heard as a distinct component, so that the resulting stimulus sounds like a complex tone and a single frequency presented concurrently. Stimuli were gated with a 10-ms cosine window at the start and end, and were otherwise of constant amplitude. Target stimuli were presented monaurally to the left or right ear. Cue stimuli were created from the same speech-shaped noise samples as those used in Experiment 1, but were bandpass filtered between 190 and 2100 Hz so that they encompassed approximately the same frequency spectrum as the target stimuli.

Procedure

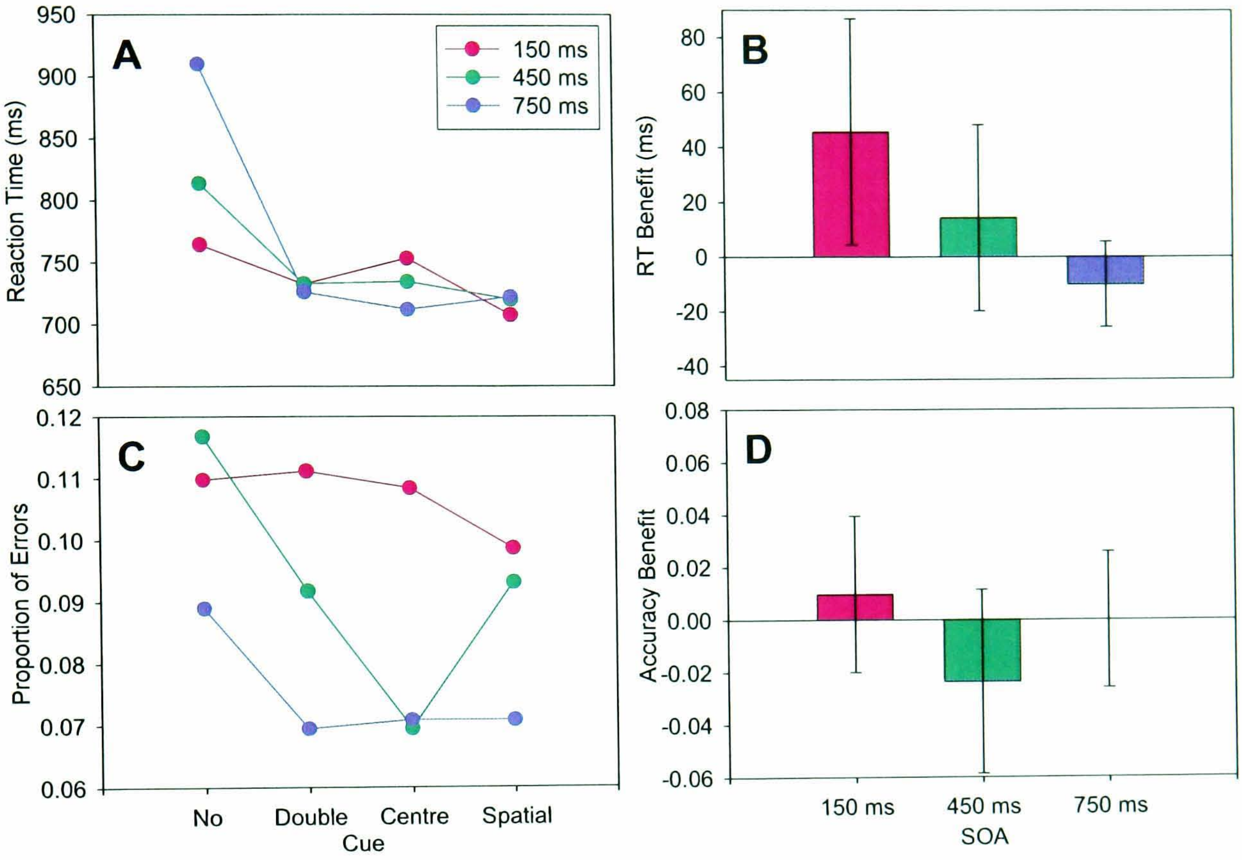
The experimental procedure was the same as for Experiment 1, with the exceptions that the initial practice session only contained 20 trials, and that target stimuli were harmonic and inharmonic complex tones. This also reduced the number of trials per block to 240 (2 targets x 5 fundamental frequencies x 2 locations x 4 cue types x 3 repeats). If subjects experienced difficulty with the practice sessions they were able to repeat them until they felt confident that they could do the task.

Results

Overall RTs and accuracy suggest that the harmonicity task was more difficult than the linguistic task. Average RTs were 745 ms in the harmonicity task, compared with 642 ms in Experiment 1, while average accuracy was 91% in the harmonicity task compared with 96% in Experiment 1. Accuracy was still relatively high though, indicating that subjects were able to perform the task competently.

Data were analysed in the same way as in Experiment 1. Trimming resulted in the removal of 0.82% of responses. Subjects were significantly faster to respond following an alerting (double) cue than with no cue at the 450-ms ($t_{11}=2.384$, $p<0.05$) and 750-ms ($t_{11}=2.942$, $p<0.05$) SOAs, but not at the 150-ms SOA ($t_{11}=1.236$, $p=0.242$). The benefit was 184 ms at the 750-ms SOA, much larger than that found at the 450-ms SOA (81 ms) and 150-ms SOA (32 ms), or in Experiment 1 with linguistic stimuli (between 64 and 67 ms). RTs, error rates and spatial-cue benefits are shown in Figure 5.2.

Figure 5.2: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a spatial cue compared with a neutral (centre) cue. Error bars show 95% confidence intervals.



Subjects were significantly faster to respond following a spatial cue than with a neutral (centre) cue, but only at the shortest SOA ($t_{11}=2.441$, $p<0.05$). At SOAs of 450 ms ($t_{11}=0.917$, $p=0.379$) and 750 ms ($t_{11}=-1.433$, $p=0.180$) subjects gained no benefit from a cue to spatial location. There were no significant alerting or orienting accuracy effects.

The ANOVA (Table 5.4) conducted on the RT data showed no significant main effects of SOA or cue type. There was a significant interaction, but planned comparisons with a Bonferroni correction revealed that the spatial orienting effect did not reach significance at any level of the SOA. However, there was a near-significant 46-ms benefit at the 150-ms SOA ($t_{11}=2.441$, p=0.033). ANOVA analysis did not reveal any significant effects in the accuracy data.

Table 5.4: ANOVA results from the RT and accuracy data. The table shows main effects of SOA (150, 450, 750 ms) and cue type (centre cue, spatial cue), and the interaction between SOA and cue type.

	SOA	Cue	SOA x Cue
RT	$F_{2,22}=0.089$, *n.s*	$F_{1,11}=2.514$, *n.s.*	$F_{2,22}=5.008$, $p<0.05$
Accuracy	$F_{1.3,13.9}=1.641$, *n.s.*	$F_{1,11}=0.248$, *n.s.*	$F_{2,22}=1.854$, *n.s.*

Discussion

As in Experiment 1, substantial alerting effects were found at all SOAs, although they were only significant at the 450- and 750-ms SOAs. The benefit from an alerting cue was particularly large at the 750-ms SOA (184 ms), which was partly due to one subject responding 850 ms faster with a double cue than with no cue. Without this subject the average alerting benefit was 124 ms. A significant spatial cue benefit of 46 ms was found at the 150-ms SOA, while small, non-significant spatial cue effects were found at the 450- and 750-ms SOAs (14 and -10 ms, respectively). The large orienting benefit found at the 150-ms SOA in this study, but not in Experiment 1, suggests that subjects were able to orient to the start of the linguistic stimuli, before sufficient information was available on which to respond. Similarly, the increased magnitude of the alerting benefit at the 750-ms SOA suggests that the linguistic stimuli conferred some alerting benefit.

The pattern of alerting benefits obtained in this experiment is difficult to reconcile with that found in the previous experiment. Experiment 1 revealed significant alerting benefits which did not vary with SOA (67, 67, and 64 ms at the 150-, 450-, and 750-ms SOAs, respectively). This experiment elicited alerting benefits which increased with increasing SOA (32, 81, and 184 ms, respectively). It appears that the alerting benefit obtained from the linguistic stimuli, and the alerting benefit obtained from the alerting cues, do not vary with SOA in the same manner: that while alerting to the onset of linguistic stimuli is rapid and sustained, alerting in response to alerting cues is slower to occur. However, these experiments are unable to address this question adequately since they were not specifically designed to do so.

The spatial-cue benefits found at the 450- and 750-ms SOAs were of a similar magnitude to those found in the auditory ANT and Experiment 1. However, in this experiment a valid-cue cost, rather than benefit, was found at the 750-ms SOA. As with Experiment 1, response times with a spatial cue are similar to those with a double cue, although in this experiment responses on spatial-cue trials were slightly faster than on double-cue trials (25, 12, and 4 ms at the 150-, 450-, and 750-ms SOAs, respectively).

Since Experiment 2 also failed to elicit robust endogenous spatial orienting benefits, the following two experiments were conducted. These studies aimed to increase subjects' motivation to generate a spatial representation of the non-spatial task.

Experiment 3: Steady-state stimuli, higher spatial salience

According to the spatial relevance hypothesis, spatial orienting benefits can be obtained on non-spatial tasks in which cues are informative about target location. The previous two experiments provided 100% valid spatial cues, yet these only elicited a small (14 ms) spatial orienting effect at the 750-ms SOA in Experiment 1, and a 46-ms cueing benefit at the 150-ms SOA in Experiment 2. Since the role of informative spatial cues is to encourage the listener to generate a spatial representation of the task, it is possible that the relatively infrequent occurrence of spatial-cue trials was insufficient for this to occur. Because the ANT uses the cueing paradigm to investigate both alerting and orienting, spatial cues are only presented on 25% of trials: a low proportion compared to studies which only investigate spatial orienting effects. To increase the salience of the spatial nature of the task, in Experiment 3 only neutral cues and spatial cues were presented. This increased the proportion of trials on which a spatial cue was presented to 50%. Spatial cues were still 100% valid (i.e. they always accurately predicted target location).

In the ANT, a measure of spatial orienting is obtained by comparing performance on spatial-cue trials with performance on centre-cue trials. In designing the ANT, Fan et al. (2002) reasoned that a double cue would keep attention distributed across both possible target locations, as it would be in the no-cue condition, and that a centre cue would narrowly focus attention, as would a spatial cue. Therefore alerting benefit was calculated by comparing no-cue and double-cue conditions, and orienting benefit was calculated by comparing centre-cue and spatial-cue conditions. While this reasoning appears valid for the visual task, the different effects found with double and centre cues in Experiments 1 and 2 suggest that the double cue might be a more appropriate neutral baseline in these experiments than the centre cue. Comparing spatial-cue performance with double-cue performance reveals the amount of benefit received from orienting to the correct spatial location compared with keeping attention diffuse across both target locations. In Experiment 3, therefore, the neutral cues will be double cues, rather than centre cues. Experiment 3 replicates Experiment 2, but removes the no-cue and centre-cue conditions in order to present double cues on 50% of trials, and spatial cues on the remaining 50% of trials.

Method

Subjects

Twelve volunteers (1 male, mean age 21.0 years, range 17 - 27) participated in Experiment 3. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that 11 subjects had normal hearing (thresholds below 20 dB HL). The remaining subject had thresholds no greater than 25 dB HL, which was considered acceptable for this study. Two further subjects were excluded: one due to overall accuracy of 59%, and one due to an overall average reaction time of 1497 ms (compared to a group average of 564 ms). All participants spoke English as their native language, gave informed consent prior to the study, and received £5 compensation.

Apparatus and Stimuli

Apparatus and stimuli were identical to those in Experiment 2.

Procedure

The experimental procedure was the same as for Experiment 2, with the exception that only two cue types were used. These were neutral (double) cues and 100% valid spatial cues. Each block still contained 240 trials (2 targets x 5 fundamental frequencies x 2 locations x 2 cue types x 6 repeats).

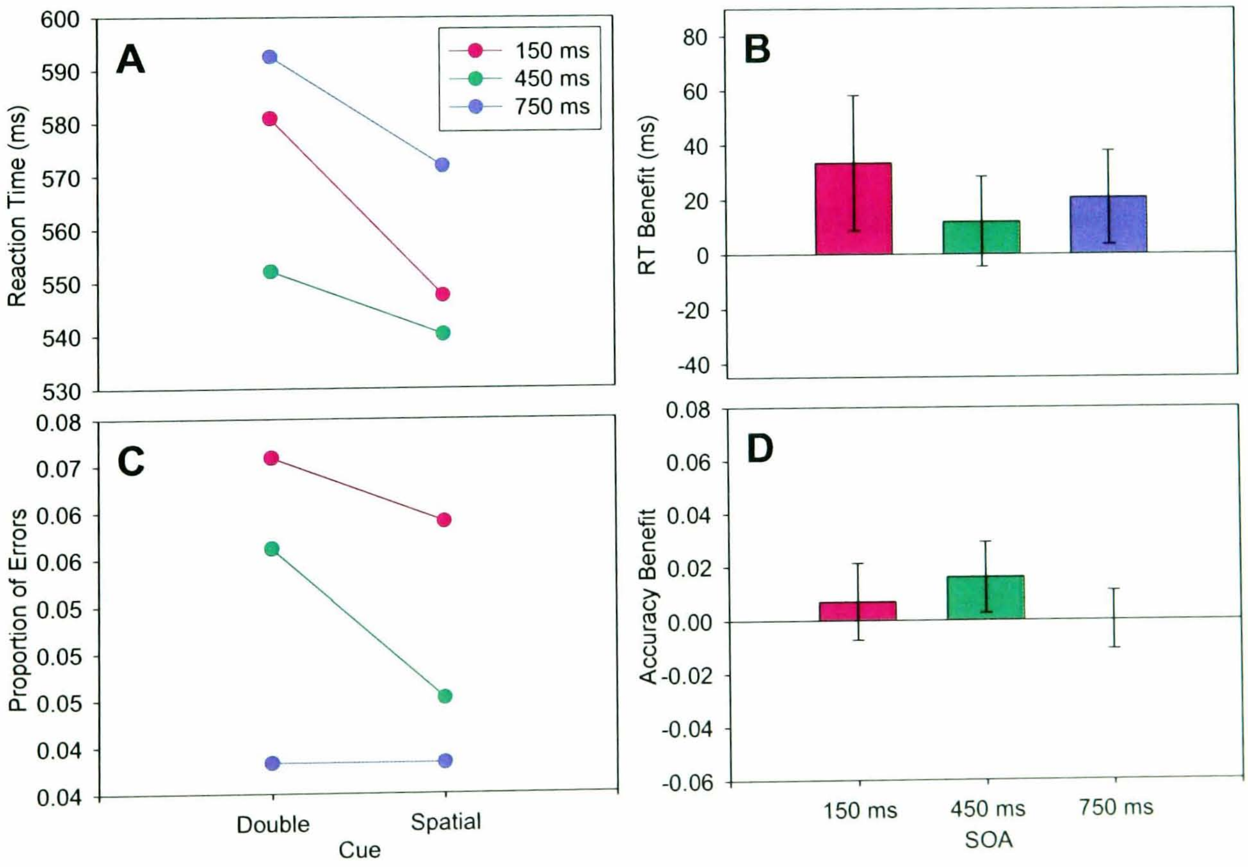
Results

Data were analysed using the ANOVA analysis described in Experiment 1, except that neutral (double) cues were used as the baseline instead of neutral (centre) cues. Trimming resulted in the removal of 0.86% of responses. The ANOVA (Table 5.5) showed significant RT effects of SOA and cue type, but no interaction. While responses were generally faster with a spatial cue than with a neutral cue, planned comparisons showed that the effect was only significant at the 150-ms SOA ($t_{11}=2.969$, $p<0.0167$), although there was a trend at the 750-ms SOA ($t_{11}=2.650$, $p=0.023$). In the accuracy data there were significant main effects of SOA and cue type, but no interaction. Planned comparisons showed that cue benefit for accuracy did not reach significance at any individual level of the SOA, although there was a trend at the 450-ms SOA ($t_{11}=2.653$, $p=0.022$). RTs, error rates and spatial-cue benefits are shown in Figure 5.3.

Table 5.5: ANOVA results from the RT and accuracy data. The table shows main effects of SOA (150, 450, 750 ms) and cue type (double cue, spatial cue), and the interaction between SOA and cue type.

	SOA	Cue	SOA x Cue
RT	$F_{2,22}=3.756$, $p<0.05$	$F_{1,11}=17.542$, $p<0.01$	$F_{1.3,14.8}=1.451$, *n.s.*
Accuracy	$F_{2,22}=3.597$, $p<0.05$	$F_{1,11}=9.576$, $p<0.05$	$F_{2,22}=11.507$, *n.s.*

Figure 5.3: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a spatial cue compared with a neutral (double) cue. Error bars show 95% confidence intervals.



Discussion

The results from Experiment 3 are broadly similar to those found in Experiment 2, despite the changes to the cueing protocol. Subjects gained a 33-ms orienting benefit at the 150-ms SOA, comparable to the 46-ms benefit found in Experiment 2. Similarly, at the 450-ms SOA both studies revealed non-significant effects: of 12 ms in the present experiment, compared with 14 ms in Experiment 2. The spatial cue benefit at the 750-ms SOA in the current experiment was slightly larger (21 ms) than that found in the earlier experiments (14 and -10 ms in Experiments 1 and 2, respectively) but did not quite reach significance. Increasing the salience of the spatial nature of the task did not seem to substantially alter the amount of benefit subjects were able to obtain from the spatial cues.

Experiment 4: Steady-state stimuli, invalid vs. valid

Many studies investigating auditory spatial orienting (e.g. Spence & Driver, 1994) contrast performance with valid spatial cues with performance with invalid spatial cues (which cue the wrong target location). Invalid cueing might theoretically increase the size of spatial orienting benefits in two ways. First, it adds costs of attending to the wrong location to benefits from attending to the correct location. Second, being cued to the wrong spatial location is subjectively a striking effect, which increases the salience of the spatial nature of the task. Experiment 4 replicates Experiment 3, but with 80% valid spatial cues (which occur at the correct target location), and 20% invalid spatial cues (which occur at the wrong target location).

Method

Subjects

Twelve volunteers (9 male, mean age 24.3 years, range 18 - 36) participated in Experiment 4. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that 11 subjects had normal hearing (thresholds below 20 dB HL). The remaining subject had thresholds no greater than 25 dB HL, which was considered acceptable for this study. Two further subjects were not tested: one who had thresholds up to 55 dB HL, and one who was unable to discriminate between the harmonic and inharmonic stimuli. All participants spoke English as their native language, gave informed consent prior to the study, and received £5 compensation.

Apparatus and Stimuli

Apparatus and stimuli were identical to those in Experiment 2.

Procedure

The experimental procedure was the same as for Experiment 3, apart from the cueing conditions. Subjects heard either a valid spatial cue, which correctly cued target location, or an invalid spatial cue, which cued the wrong target location (i.e. the opposite side to the target). Valid spatial cues were presented on 80% of trials. Invalid spatial cues were presented on the remaining 20% of trials. Each block contained 240 trials. To maintain this number of trials while cueing 80% of trials validly and 20% invalidly, and keeping the number of each type of target, on each side, equal, additional trials with a fundamental frequency of 200 Hz were included. The trials therefore comprised 200 trials (2 targets x 5 fundamental frequencies x 2 sides x 5 samples (4 validly-cued and 1 invalidly-cued) x 2 repeats), plus an additional 40 trials with targets at 200 Hz (2 targets x 2 sides x 2 cue types x 5 samples).

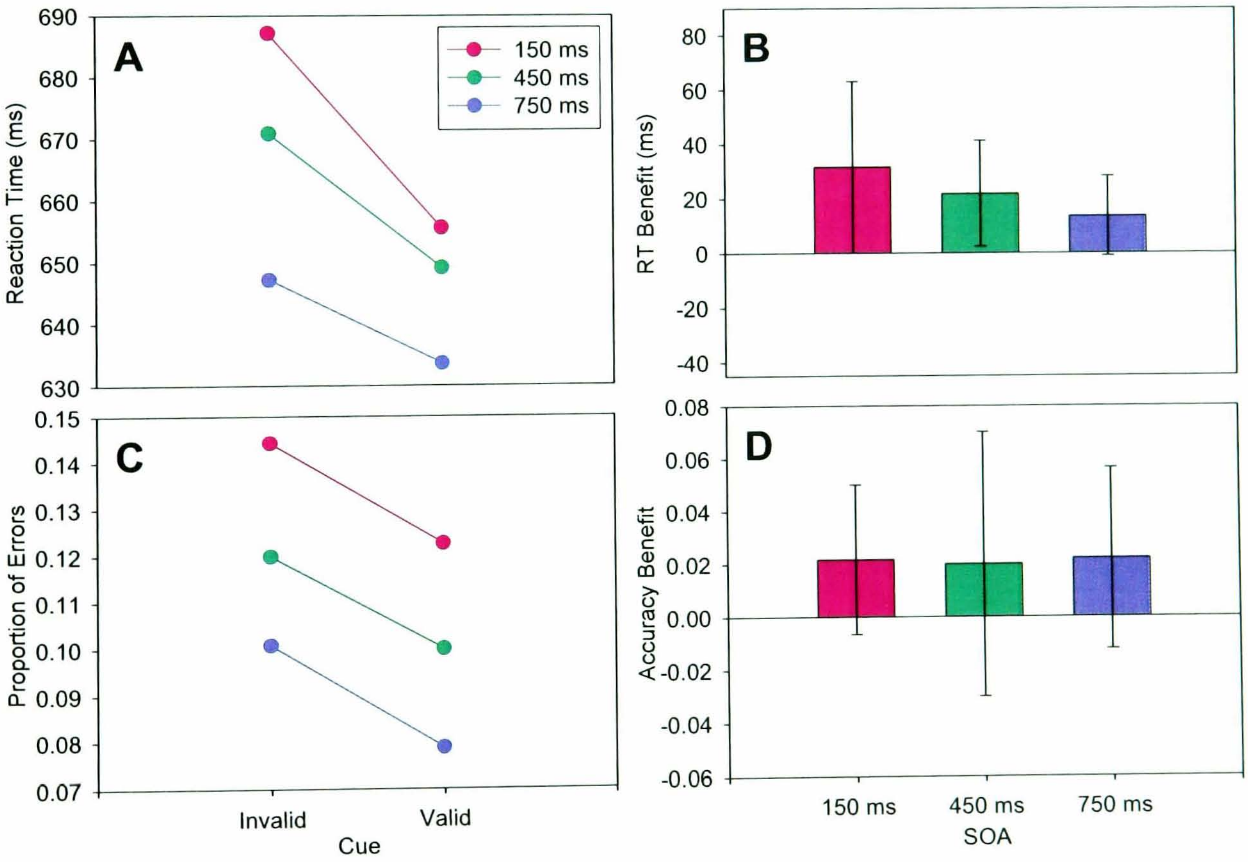
Results

Data were analysed using the ANOVA analysis described in Experiment 1. Trimming resulted in the removal of 1.28% of responses. The ANOVA (Table 5.6) showed a significant RT main effect of cue type, but no effect of SOA and no interaction. While responses were generally faster with a valid spatial cue than with an invalid spatial cue, planned comparisons showed that these effects were not significant at any individual level of the SOA, although there were trends at all three SOAs (150 ms: t_{11}=2.206, p=0.050; 450 ms: t_{11}=2.472, p=0.031; 750 ms: t_{11}=2.037, p=0.066). The ANOVA did not reveal any significant effects in the accuracy data. RTs, error rates and spatial-cue benefits are shown in Figure 5.4.

Table 5.6: ANOVA results from the RT and accuracy data. The table shows main effects of SOA (150, 450, 750 ms) and cue type (invalid cue, valid cue), and the interaction between SOA and cue type.

	SOA	Cue	SOA x Cue
RT	$F_{2,22}=0.365$, *n.s.*	$F_{1,11}=10.729$, $p<0.01$	$F_{1.1,11.6}=0.864$, *n.s.*
Accuracy	$F_{2,22}=2.388$, *n.s.*	$F_{1,11}=2.430$, *n.s.*	$F_{2,22}=0.007$, *n.s.*

Figure 5.4: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a valid spatial cue compared with an invalid spatial cue. Error bars show 95% confidence intervals.



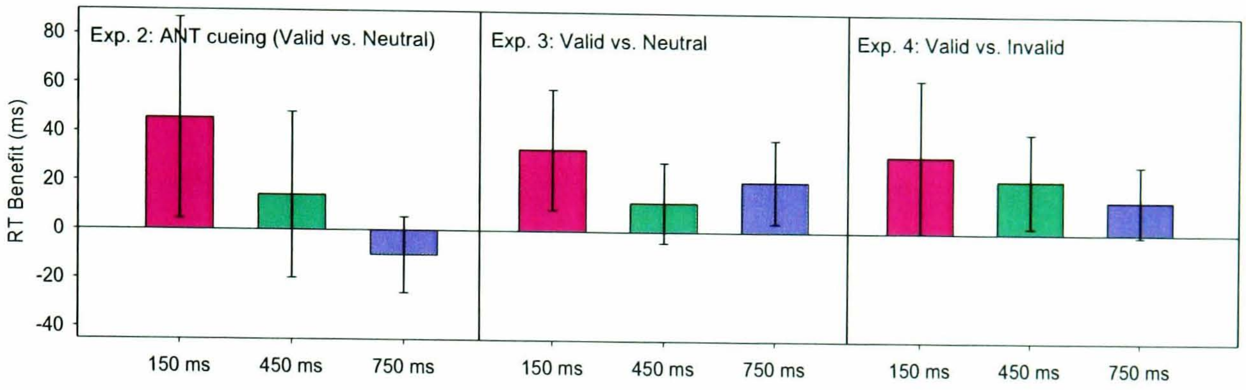
Discussion

In Experiment 4, performance with invalid and valid cues was compared to determine whether combined costs from orienting to the wrong location, and benefits from orienting to the correct location, might increase the size and reliability of the spatial orienting effect. In fact, the results from this experiment are comparable to those from Experiment 3, in which performance with valid spatial cues was compared with performance with neutral cues. The orienting effect at the 150-ms SOA was 33 ms in Experiment 3, and 32 ms in the present experiment, although the effect only reached significance in Experiment 3. At the 450-ms SOA a slightly larger valid-cue benefit was found in the current experiment (22 ms) than the previous one (12 ms), although neither effect reached significance. Similarly, both experiments produced non-significant effects at the 750-ms SOA, although the pattern was reversed, with a larger benefit in the previous experiment (21 ms) than in the current one (14 ms).

Interim Summary: Experiments 2 to 4

The results from the first three experiments using steady-state stimuli (Experiments 2, 3, and 4, summarised in Figure 5.5) demonstrate two key effects. First, the orienting effect at the 150-ms SOA is reasonably robust, being of a similar magnitude in all experiments (46, 33, and 32 ms, respectively), and either significant or near-significant in all cases. This effect is likely to reflect exogenous (reflexive) orienting to the cued location. Second, endogenous orienting effects at longer (450- and 750-ms) SOAs are less robust. In all three experiments, the effects are small and do not reach significance, although there are trends towards significance at the 750-ms SOA in Experiment 3, and at both the 450- and 750-ms SOAs in Experiment 4. There is certainly no indication of the robust spatial orienting effects found in visual studies (such as the visual ANT, Chapter 2).

Figure 5.5: Spatial cue RT benefits in Experiments 2, 3, and 4.



The pattern of results (a strong effect at the 150-ms SOA which is not present at longer SOAs) suggests that while the valid spatial cues are automatically capturing attention, subjects are not voluntarily orienting their attention to the cued location in an effective manner. Whether this is because they are ignoring the cues, or because they are unable to use the cues effectively, is not clear. To address this, the next two experiments separate out exogenous and endogenous orienting effects. Experiment 5 presents uninformative spatial cues (50% valid and 50% invalid) to investigate exogenous spatial orienting alone. If the pattern of results from this study resembles the pattern of results from the previous three studies, it would suggest that subjects may not be trying to orient their attention to the cued locations. Experiment 6 presents symbolic spatial cues to investigate endogenous spatial orienting alone. If the pattern of results at the longer SOAs resembles the pattern of results in the previous three studies, this might suggest that participants are attempting to voluntarily orient their attention to the cued location, but are not able to do so very effectively.

Experiment 5: Steady-state stimuli, uninformative cueing

Experiment 5 replicates Experiment 4, but with different proportions of valid-cue and invalid-cue trials. By altering the proportions to 50% valid and 50% invalid, cues become uninformative, allowing exogenous (automatic) spatial orienting effects to be investigated independently of endogenous (voluntary) spatial orienting effects. The spatial relevance hypothesis predicts that no spatial orienting benefits will be found in this experiment, since it involves a non-spatial task and uninformative cues. Spence and Driver (1994) and McDonald and Ward (1999) each presented two experiments using non-spatial tasks and uninformative cues. None of these experiments elicited exogenous (or endogenous) cueing benefits. However, other studies (Mondor & Amirault, 1998; Mondor & Breau, 1999) have demonstrated uninformative spatial-cue benefits at short SOAs in non-spatial tasks.

Method

Subjects

Twelve volunteers (6 male, mean age 23.4 years, range 18 - 30) participated in Experiment 5. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that all subjects had normal hearing (thresholds below 20 dB HL). All participants spoke English as their native language, gave informed consent prior to the study, and received £5 compensation.

Apparatus and Stimuli

Apparatus and stimuli were identical to those in Experiment 2.

Procedure

The experimental procedure was the same as for Experiment 4, except for the proportions of valid and invalid spatial cue trials. Valid spatial cues were presented on 50% of trials, and invalid spatial cues were presented on the remaining 50% of trials. Since the cue was equally likely to cue the incorrect side as the correct side, this is an uninformative cueing paradigm. Each block contained 240 trials (2 targets x 5 fundamental frequencies x 2 locations x 2 cue types x 6 repeats).

Results

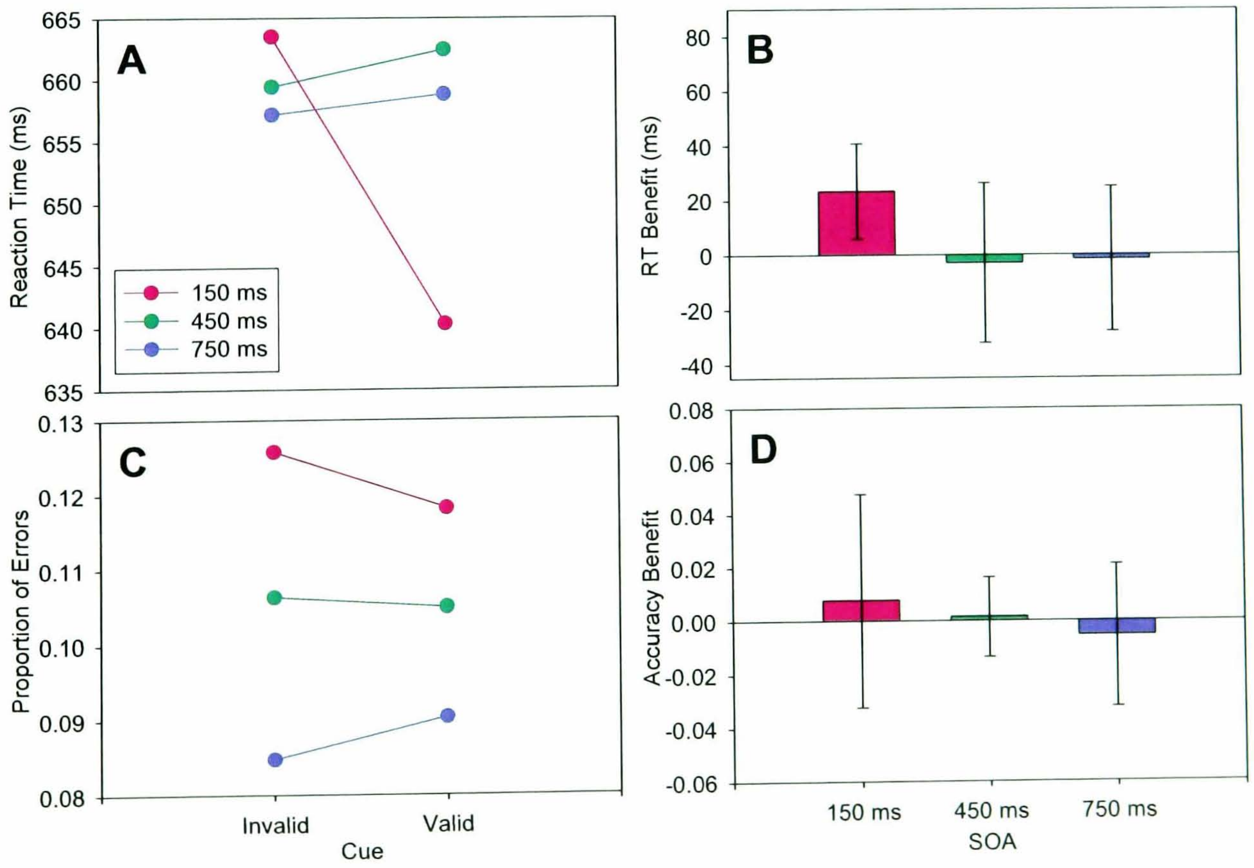
Data were analysed using the ANOVA analysis described in Experiment 1. Trimming resulted in the removal of 1.01% of responses. There were no significant main effects or interactions in either the RT or accuracy data (Table 5.7). Although the ANOVA did not produce significant results, planned comparisons revealed a significant valid-cue benefit at the 150-ms SOA

($t_{11}=2.927$, $p=0.014$). RTs, error rates and spatial-cue benefits are shown in Figure 5.6.

Table 5.7: ANOVA results from the RT and accuracy data. The table shows main effects of SOA (150, 450, 750 ms) and cue type (invalid cue, valid cue), and the interaction between SOA and cue type.

	SOA	Cue	SOA x Cue
RT	$F_{2,22}=0.082$, *n.s.*	$F_{1,11}=0.698$, *n.s.*	$F_{2,22}=1.995$, *n.s.*
Accuracy	$F_{2,22}=2.127$, *n.s.*	$F_{1,11}=0.013$, *n.s.*	$F_{1.3,14.8}=0.390$, *n.s.*

Figure 5.6: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a valid spatial cue compared with an invalid spatial cue. Error bars show 95% confidence intervals.



Discussion

Experiment 5 presented uninformative cues in order to assess exogenous orienting effects independently of endogenous orienting effects. The spatial relevance hypothesis predicted that no spatial orienting effects would be found with a non-spatial task and uninformative cues. However, in contrast to the spatial relevance hypothesis, but consistent with studies by Mondor and colleagues (Mondor & Amirault, 1998; Mondor & Breau, 1999), the results reveal an exogenous valid-cue benefit of 23 ms at the 150-ms SOA. As would be expected with uninformative cues, no endogenous orienting benefits were found at the longer SOAs. Unlike the previous three studies, which have produced small, non-significant, spatial orienting benefits, this study clearly demonstrates that performance is not altered by cue validity (-3 and -2 ms effects at the 450- and 750-ms SOAs, respectively). These results are consistent with the suggestion that subjects in the previous three informative-cueing studies were attempting to use the spatial cues, but were unable to do so effectively.

Experiment 6: Steady-state stimuli, symbolic cueing

Experiment 6 investigates endogenous (voluntary) orienting effects independently of exogenous (automatic) orienting effects. To achieve this, symbolic auditory cues were presented centrally, which directed subjects to the correct side on 80% of trials (valid cues), and to the wrong side on the remaining 20% of trials (invalid cues). Cues were a single 50-ms noise burst to direct attention to the left, and two consecutive 25-ms noise bursts to direct attention to the right. Subjects took part in a practice session prior to the experiment in which they familiarised themselves with the cues. In all other respects, Experiment 6 replicates Experiment 4. Providing central, symbolic cues allows a test of whether subjects in Experiments 2, 3, and 4 were voluntarily orienting their attention to the cued location, or whether attention was remaining at the cued location following automatic capture by the peripheral cues. Experiment 5, which presented uninformative cues and found no cue validity effects at the 450- and 750-ms SOAs suggests that this was not the case, but the result may have been due to subjects actively re-orienting their attention to the centre, or to both target locations, under uninformative cueing conditions.

Method

Subjects

Twelve volunteers (2 male, mean age 23.1 years, range 19 - 32) participated in Experiment 6. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that all subjects had normal hearing (thresholds below 20 dB HL). Two further subjects were excluded, due to overall accuracy of 66% and 68%. All participants spoke English as their native language, gave informed consent prior to the study, and received £5 compensation.

Apparatus and Stimuli

Apparatus and stimuli were identical to those in Experiment 2, except for the cue stimuli. Cue stimuli were designed to cue target location without being presented at target locations. The symbolic auditory cues were presented diotically, so that they were heard in the middle of the head. The symbolic cue directing subjects to their left comprised 50 ms of speech-shaped noise, bandpass filtered between 190 and 2100 Hz, and gated with a 10-ms cosine window at the start and end. The symbolic cue directing subjects to their right comprised two consecutive 25-ms bursts of the same noise, each cosine gated for 10 ms at the start and end. Cues were therefore matched for frequency and duration, but differed in their amplitude envelope.

Procedure

The experimental procedure was the same as for Experiment 4, but with an additional 40-trial practice session to accustom subjects to the symbolic cues. In the practice trials, subjects heard symbolic cues and responded with a left button-press if the cue directed them to their left, and with a right button-press if the cue directed them to their right. During the experiment subjects were cued to the correct target location on 80% of trials (valid spatial cues) and to the wrong target location on the remaining 20% of trials (invalid spatial cues). In the experimental trials, subjects responded to the identity of the target (harmonic or inharmonic tone). As with Experiment 4, 240 trials were presented in each block, including additional trials with targets with a 200-Hz fundamental frequency to create a balanced experiment.

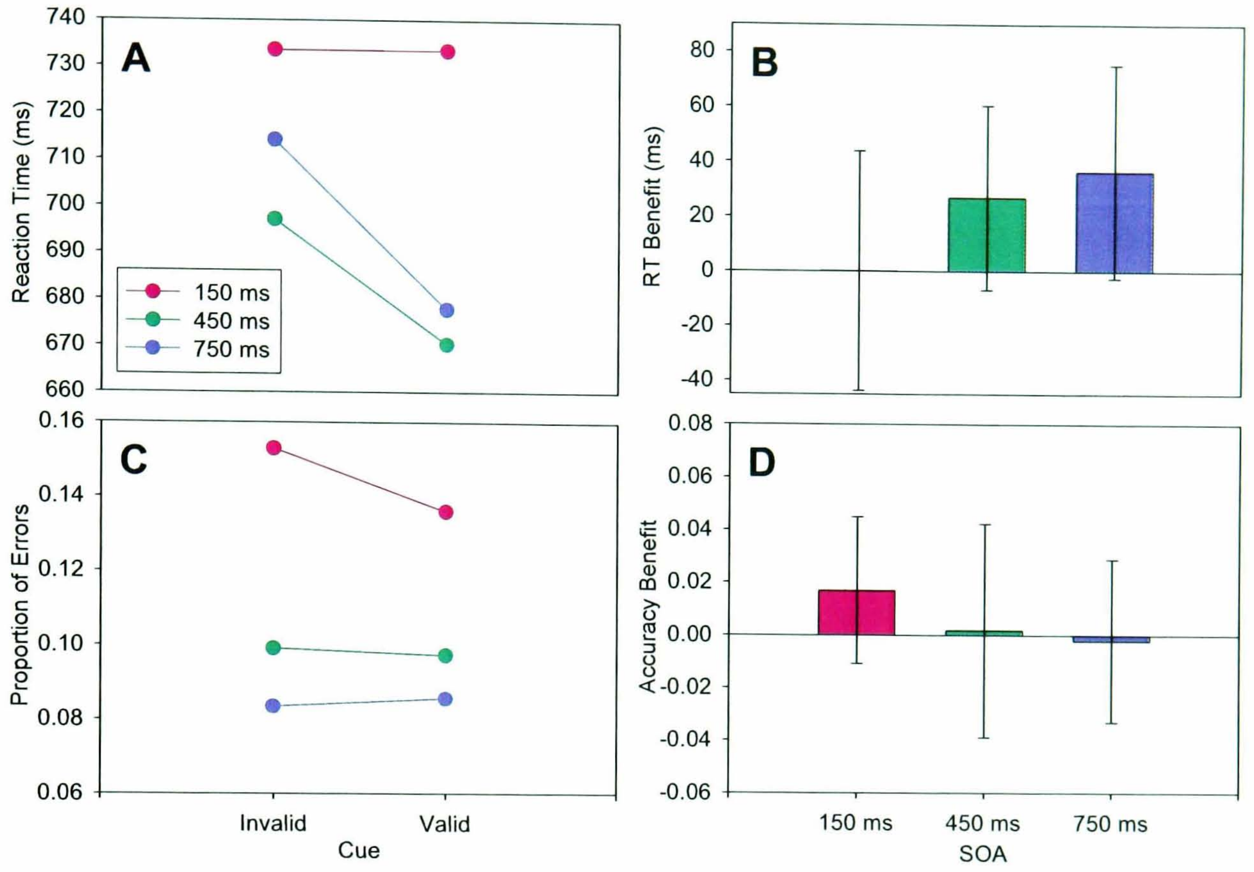
Results

Data were analysed using the ANOVA analysis described in Experiment 1. Trimming resulted in the removal of 2.01% of responses. There were no significant RT main effects, and there was no interaction (Table 5.8). Despite the non-significant main effects, there was a trend towards a significant spatial-cue benefit at the 750-ms SOA (t_{11}=2.059, p=0.064). There was a significant effect of SOA in the accuracy data, but no effect of cue type and no interaction. The highest error rates occurred with the 150-ms SOA, followed by the 450-ms and then 750-ms SOAs. RTs, error rates and spatial-cue benefits are shown in Figure 5.7.

Table 5.8: ANOVA results from the RT and accuracy data. The table shows main effects of SOA (150, 450, 750 ms) and cue type (invalid cue, valid cue), and the interaction between SOA and cue type.

	SOA	Cue	SOA x Cue
RT	$F_{2,22}=1.060$, *n.s.*	$F_{1,11}=3.016$, *n.s.*	$F_{1.3,14.7}=1.422$, *n.s.*
Accuracy	$F_{2,22}=12.941$, $p<0.01$	$F_{1,11}=0.365$, *n.s.*	$F_{2,22}=0.452$, *n.s.*

Figure 5.7: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a valid spatial cue compared with an invalid spatial cue. Error bars show 95% confidence intervals.



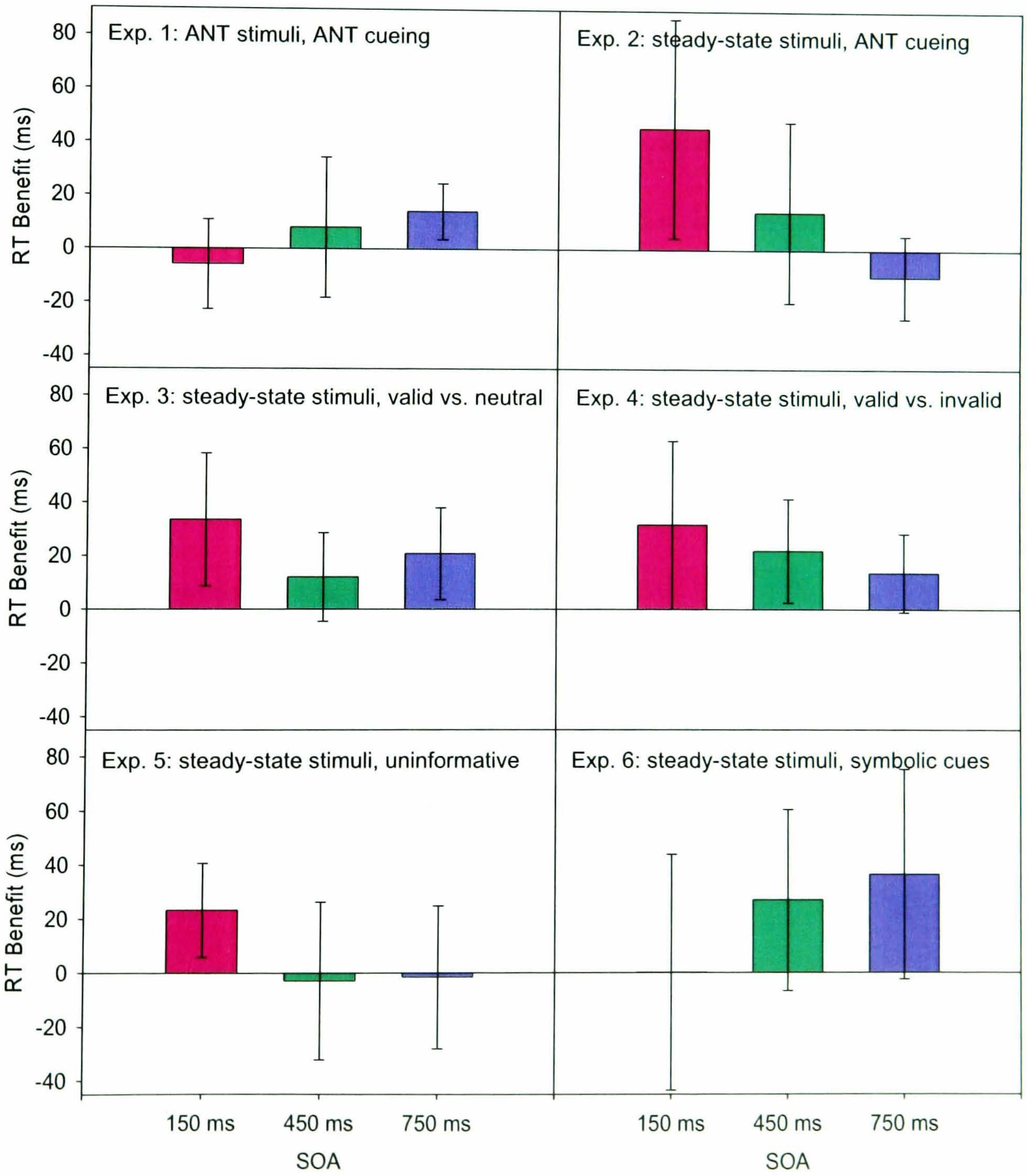
Discussion

As expected, subjects were unable to orient their attention to the cued location at the shortest (150 ms) SOA. However, at the 450-ms SOA subjects gained a 27-ms benefit, and at the 750-ms SOA subjects gained a 36-ms benefit. While both of these effects are large relative to those found in earlier experiments, substantial variability across subjects prevented either from reaching significance. However, the results do demonstrate that subjects are able to gain some benefit from cues to spatial location, and that the results at the longer (450- and 750-ms) SOAs in Experiments 2, 3, and 4 are not simply the result of attention remaining at cued locations after being automatically captured by peripheral cues.

Interim Discussion: Experiments 1 to 6

Experiments 1 to 6 investigated spatial cueing benefits under a range of different cueing protocols. Figure 5.8 shows the average spatial-cue benefits found in each experiment.

Figure 5.8: RT benefit from a valid spatial cue in each experiment. Error bars show 95% confidence intervals.

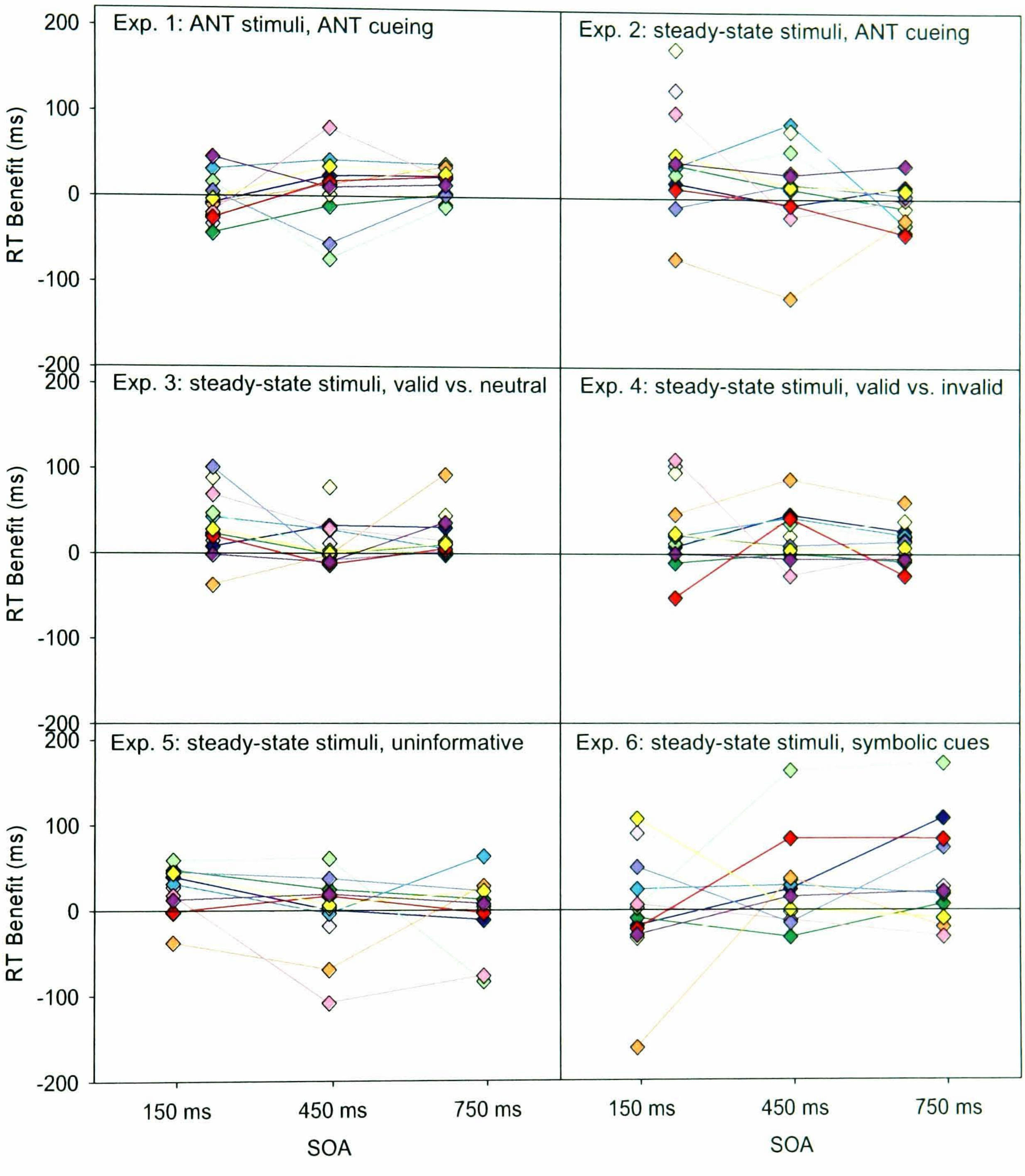


Experiment 1 used linguistic stimuli from the auditory ANT (Chapter 2), while Experiments 2 to 6 used steady-state harmonic and inharmonic complex tones. Both tasks were non-spatial: a pitch discrimination task in Experiment 1, and a harmonic vs. inharmonic discrimination in Experiments 2 to 6. Spatial-cue benefits were therefore only expected (according to the spatial relevance hypothesis) when informative spatial cues were provided, since only then would subjects generate a spatial representation of the task. All but one experiment which provided peripheral spatial cues elicited significant or near-significant spatial-cue benefits at the 150-ms SOA, including Experiment 5, which presented uninformative cues. Spatial-cue benefits at the 150-ms SOA reflect exogenous orienting, in which attention is automatically captured by the peripheral cue. The experiment which failed to elicit exogenous spatial orienting was Experiment 1, in which linguistic target stimuli were presented. The lack of exogenous orienting might result from subjects being able to orient their attention to the start of the target word, before sufficient information was available on which to respond.

Endogenous spatial orienting effects were much less robust than exogenous spatial orienting effects, but were present at least as a trend in most studies. Experiment 5 presented uninformative cues, and, as expected, did not elicit endogenous orienting benefits. All of the remaining studies showed an average spatial-cue benefit at the 450-ms SOA, although this was never significant, and only near-significant in Experiment 4, which presented 80% valid spatial cues, and 20% invalid spatial cues. At the 750-ms SOA four of the five informative-cueing studies produced an average spatial-cue benefit, of which one was significant (Experiment 1, with linguistic stimuli) and two were near-significant (Experiment 3, with 50% spatial cues and 50% neutral cues; and Experiment 4, with 80% valid spatial cues and 20% invalid spatial cues). However, effects in all three of these studies were small (between 14 and 21 ms). In Experiment 6, with informative symbolic cues, subjects gained an average benefit of 36 ms, but this did not reach significance. Experiment 2, which used the ANT cueing protocol with steady-state stimuli, produced a spatial-orienting cost of 10 ms.

The overall impression from Experiments 1 to 6 is that while endogenous spatial orienting does occur with non-spatial tasks, spatial-cue benefits are not very robust. All endogenous effects were small and/or non-significant, and there was considerable variation across subjects. Figure 5.9 shows spatial-cue benefits obtained by individual subjects. The striking impression from these graphs is that some subjects gained substantial spatial-cue benefits, while others gained substantial spatial-cue costs. For example, in Experiment 2 the subject depicted in turquoise gained an 89-ms advantage at the 450-ms SOA, while the subject depicted in orange experienced a 118-ms cost at the 450-ms SOA. In the symbolic cueing experiment (Experiment 6) in particular, it is possible to see that some subjects (such as those shown in light green and red) gained a large and consistent cue benefit, while others (such as those shown in pink and dark green) either received no benefit, or a slight cost. In contrast, some subjects (such as those depicted in dark yellow and dark blue) obtained a benefit at one or more SOAs, but not at others.

Figure 5.9: RT benefit obtained by individual subjects. Colours indicate different subjects, but each experiment used a different set of subjects, so colours are not consistent across experiments.

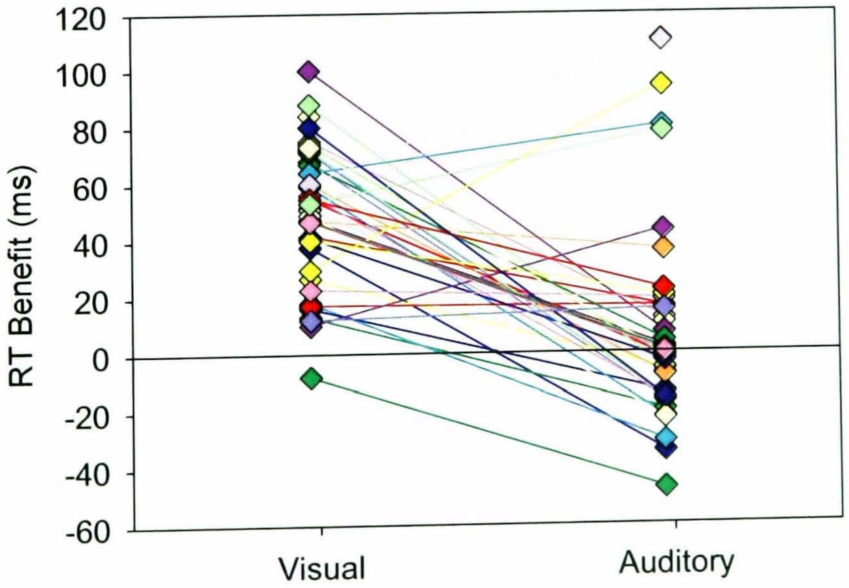


Looking at the first two sets of subjects, three explanations appear possible. First, certain subjects may be better-able to make use of the spatial cues than others. Second, subjects may adopt different listening strategies, some of which are more successful than others. Third, some subjects might simply choose to ignore the spatial cues, as they do not expect to obtain much benefit from them.

These accounts could also apply to subjects who obtain benefit from spatial cues at some SOAs but not others, although it would then be necessary to assume, for example, strategies which confer a benefit at some SOAs, but not others.

For comparison, Figure 5.10 shows individual cueing-benefits from the visual and auditory ANTs. While the visual ANT produced some variability in the amount of benefit subjects were able to obtain, with one subject obtaining a small cost from spatial cues, in general subjects are reasonably consistent. This is in contrast to the auditory ANT data, where some subjects were able to gain a great deal of benefit from the auditory cues, and others experienced substantial costs.

Figure 5.10: Individual spatial-orienting benefits from the visual and auditory ANTs (Chapter 2). Lines link scores from the same subjects, each colour is used for more than one subject.



Experiment 7: Influence of Stimulus Presentation Method

Experiment 7 investigated whether stimulus presentation method influenced the amount of benefit subjects obtained from a valid spatial cue. Three different presentation methods were contrasted: 1) freefield presentation, in which sounds were presented from speakers situated to the subject's right and left (+/-90° azimuth), at head height; 2) monaural presentation, in which sounds were presented monaurally to the left or right ear over headphones; and 3) binaural ITD presentation, in which sounds were presented binaurally over headphones, but with an interaural time difference of +/-600 μs, so that stimuli were lateralised to approximately 90° azimuth. If auditory spatial orienting is achieved by orienting attention to an ear, rather than to a genuine spatial location, then spatial-cue benefits would be expected in the monaural and freefield conditions (since both present sounds predominantly to one ear, and therefore favour a better-ear strategy), but not in the binaural ITD condition in which stimuli are presented at the same level to both ears. If however a genuine spatial location is attended, then all three conditions should be effective in eliciting spatial orienting benefits. In particular, the freefield condition offers the greatest wealth of cues (ITD, ILD, spectral and reverberation) with which to localise sounds, and to which to orient attention. Studies of spatial orienting of auditory attention have used all three presentation methods successfully (e.g. Spence and Driver (1994) used freefield presentation, Quinlan and Bailey (1995) used monaural presentation, and Sach et al. (2000) used binaural ITD presentation). In addition, Buchtel et al. (1996) compared performance with monaural and freefield presentation, and found similar results with low-intensity stimuli, but different results with high-intensity stimuli. However, this is the first direct comparison of three types of presentation method, using the same task and the same group of subjects. Since Experiment 4 (with 80% valid and 20% invalid peripheral cues) produced the most convincing spatial-orienting effects, the same cueing protocol is used in this study. The task is the harmonic vs. inharmonic discrimination used in Experiments 2 to 6. Subjects took part in three experimental blocks, one with each type of stimulus presentation method, but at only one SOA. The 750-ms SOA was selected to ensure that endogenous (voluntary) orienting effects were being measured.

Method

Subjects

Twelve volunteers (1 male, mean age 26.3 years, range 20 - 51) participated in the experiment. Participants gave informed consent prior to the study. Pure tone air-conduction audiometry at frequencies between 250 Hz and 8000 Hz, inclusive, revealed that all subjects had normal hearing (thresholds below 20 dB HL). All participants spoke English as their native language, gave informed consent prior to the study, and received £5 compensation. One subject had previously participated in Experiment 2.

Apparatus and Stimuli

Apparatus and stimuli were similar to those used in Experiment 4. Stimuli in the two headphone-listening conditions (monaural and binaural ITD) were presented via Sennheisser HD 480II headphones. Freefield stimuli were presented via Mordaunt Short MS10 speakers. During the freefield condition subjects were seated in a high-backed chair to discourage movement, and were asked to fixate a point on the opposite wall. Speakers were situated to the subject's left and right, 58 cm from the subject's midline, and at head-height. Sounds were presented from a single speaker at a clearly audible volume, in the range 75 to 85 dB(A).

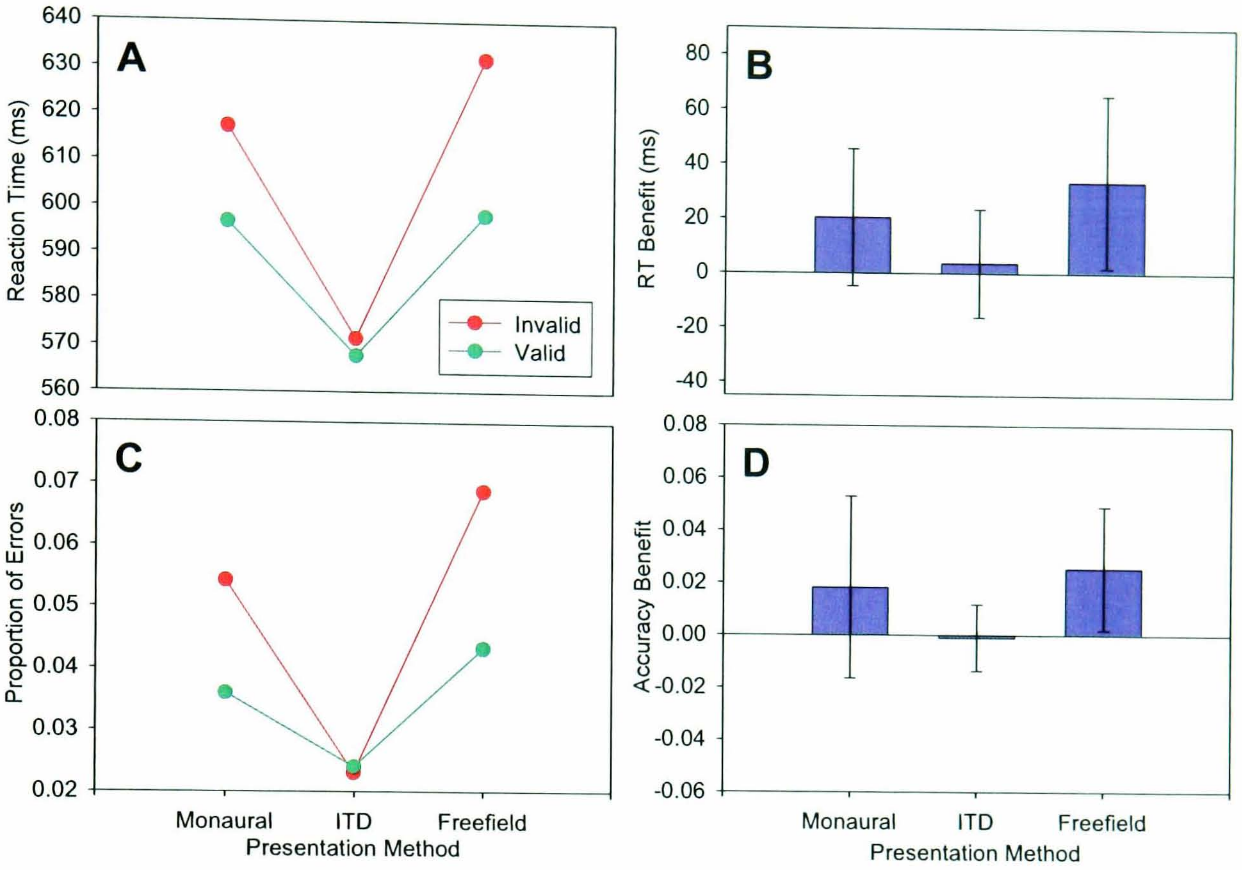
Procedure

The experimental procedure was the same as in Experiment 4, except that subjects participated in three blocks: one with each of the stimulus presentation methods, and all at an SOA of 750 ms. In the monaural condition, stimuli were lateralised to the left or to the right by presenting the sounds monaurally from one of the headphone channels. In the binaural ITD condition, stimuli were lateralised to the left or to the right by presenting the sounds from both headphone channels, but with one channel delayed by 600 μs. In the freefield condition, stimuli were localised to the left or to the right by presenting the sounds from one of the speakers located to the subject's left and right. The order of the blocks was counterbalanced across subjects. Each experimental block contained 200 trials (2 targets x 5 fundamental frequencies x 2 locations x 5 samples (4 validly-cued and 1-invalidly cued) x 2 repeats) and lasted approximately 15 minutes.

Results

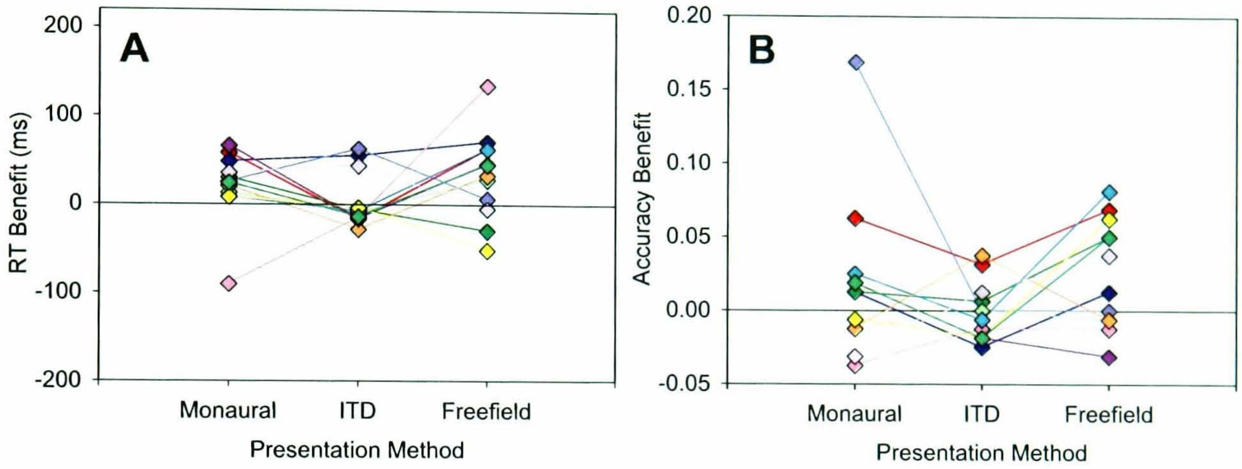
RTs from correct trials were trimmed to exclude responses faster than 100 ms and slower than 2500 ms, resulting in the removal of 0.42% of responses. Median RTs and error rates (panels A and C of Figure 5.11) were then analysed using a two-way 3 (presentation: monaural; binaural ITD; freefield) by 2 (cue type: valid; invalid) ANOVA. Where Mauchley's test of sphericity indicated that sphericity could not be assumed a Greenhouse-Geisser correction was applied. This is evident from non-integer degrees of freedom. Planned contrasts were performed using t-tests with a Bonferroni correction. Spatial orienting with each presentation method was investigated by contrasting performance with valid and invalid cues (i.e. there were three planned comparisons, giving a critical p value of 0.05/3 = 0.0167). RT and accuracy benefits from valid spatial cues compared with invalid spatial cues are shown in panels B and D of Figure 5.11, respectively. RT and accuracy data from individual subjects can be seen in Figure 5.12.

Figure 5.11: Panels A and C show overall RTs and error rates, respectively. Panels B and D show RT and accuracy benefits from a valid spatial cue compared with an invalid spatial cue. Error bars show 95% confidence intervals.



There was a significant main effect of cue type in the RT data ($F_{1,11}=12.051$, $p<0.01$), but no effect of presentation method ($F_{2,22}=3.292$, $p=0.056$) and no interaction ($F_{2,22}=1.381$, $p=0.272$). RTs were faster on valid-cue trials than on invalid-cue trials. Planned comparisons showed that the effect of cue validity did not reach significance with any one presentation method. However, there was a trend towards a significant benefit from a valid spatial cue in the monaural ($t_{11}=1.792$, $p=0.101$) and freefield ($t_{11}=2.337$, $p=0.039$) conditions, but not the ITD condition ($t_{11}=0.407$, $p=0.692$). One subject gained a substantial RT cost from a spatial cue in the monaural presentation condition (Figure 5.12). When this subject was excluded from the analysis, the main effect of cue validity was still significant ($F_{1,10}=11.198$, $p<0.01$). Removal of the outlying subject resulted in a significant benefit from a valid spatial cue when stimuli were presented monaurally (t_{10}=5.182, $p<0.001$), a trend towards a significant benefit when stimuli were presented in freefield (t_{10}=2.017, p=0.071), and no benefit when stimuli were lateralised using ITDs (t_{10}=0.549, p=0.595).

Figure 5.12: Individual RT (panel A) and accuracy (panel B) benefits from a valid spatial cue compared with an invalid spatial cue. Each colour indicates a different subject; subject colours are consistent across graphs.



There were significant differences in error rates across conditions ($F_{2,22}$=5.045, $p<0.05$), and a trend towards a significant effect of cue type ($F_{1,11}$=4.183, p=0.066), but no interaction ($F_{2,22}$=1.497, p=0.246). There was a trend towards significantly fewer errors with valid cues in the freefield condition (t_{11}=2.379, p=0.037), but no significant difference between cue conditions in the monaural (t_{11}=1.157, p=0.272) or ITD (t_{11}=-0.180, p=0.861) conditions.

Discussion

Subjects gained a near-significant benefit from valid spatial cues in the monaural and freefield conditions, but not in the binaural ITD condition. This pattern of results suggests that subjects may have been attending to an ear, rather than to a genuine spatial location, since spatial-orienting effects were found in conditions which favour a better-ear strategy. However, variability between subjects was again very high.

With the exception of the outlier, the monaural condition produced the most consistency across subjects, showing a significant 31-ms spatial-cue benefit when the outlier is removed. It is interesting to note that the monaural condition is a replication of Experiment 4 at the 750-ms SOA, and produces similar results (a near-significant 14-ms benefit in Experiment 4, and a near-significant 21-ms benefit in the present experiment). However, while only one subject in the present experiment failed to gain a spatial-cue benefit, only 8 out of 12 subjects gained a benefit in Experiment 4. Whether the weak spatial-cueing effects are the result of insufficient power can be addressed by combining data from the 750-ms SOA in Experiment 4, and the monaural condition from Experiment 7. This analysis produces a significant valid-cue benefit of 17 ms (t_{23}=2.631, p=0.015), which while significant is still not particularly large or robust.

While the spatial-cue benefits found in the freefield condition almost reached significance overall (with an average benefit of 34 ms), the individual spatial-cue effects are highly variable. Three subjects experienced a spatial-cue cost, while the remaining subjects gained between 7 and 135 ms of benefit.

The RT data from the binaural ITD condition appear to show two separate groups of subjects: a group of three who gain an average spatial-cue benefit of 54 ms, and a group of nine who gain an average spatial-cue cost of 9 ms. While the overall lack of spatial-orienting benefit found in this condition might reflect an inability to attend to a spatial location determined only by ITD cues, there are two alternative explanations for this effect. First, it is possible that at least some listeners did not gain a strong percept of a lateralised sound source. Without any training, ITD discrimination is fairly poor. In a study conducted at IHR using the same harmonic and inharmonic stimuli as were used in these experiments, naïve psychophysical listeners needed an ITD of 157 μs in order to discriminate a probe target from a reference target with an ITD of 430 μs with 79% accuracy. This large value suggests that listeners probably heard a more diffuse sound source in the ITD condition than in the monaural and freefield conditions. A weak percept of lateralisation may have influenced listeners' performance in two ways. Listeners may be less motivated to attend to the cued location when they perceive a diffuse sound source than when they are able to localise the sound source more accurately. Alternatively, listeners may gain less spatial-cue benefit when they are cued to a broad region of space than when they are cued to a precise location. In their study using sounds lateralised using ITD cues, Sach et al. (2000) selected different ITDs for each subject using a lengthy (5-9 hours) training session or an adaptive threshold procedure to determine the minimum ITD difference required between two targets in order for listeners to discriminate them with 79% accuracy. They found that subjects experienced more difficulty performing the spatial discrimination when the target stimulus varied randomly in frequency from trial to trial. Since stimuli in this experiment varied in fundamental frequency from trial to trial, and subjects were not trained to discriminate ITDs, it is possible that subjects did not obtain a strong percept of laterality in the ITD condition. Subjects who were able to benefit from the valid cues in the ITD condition may be those who obtained the strongest percept of laterality.

A second possible explanation for the null result in the ITD condition is that subjects were unable to orient attention to an internalised percept. While sounds presented from speakers or monaurally over headphones have an external sound source, sounds lateralised using ITDs appear to arise from inside the head. Since sounds arising from an internal sound source have low ecological validity, it is possible that subjects were unable to direct attention in the ITD condition as effectively as in the remaining conditions. Sounds lateralised using ILDs are also perceived intracranially, and therefore one way to test this hypothesis would be to compare performance with sounds lateralised by ITDs and ILDs.

In summary, the monaural and freefield conditions elicited the strongest spatial orienting effects, but these were not particularly large or robust. The results are consistent with the hypothesis that listeners attend to an ear, rather than to a genuine spatial location, but alternative explanations, such as difficulty lateralising stimuli in the ITD condition, and difficulty orienting to an internal location, might also be applied to the pattern of results.

General Discussion

The experiments reported in this chapter were designed with three aims: 1) to identify reasons for the null spatial orienting effect in the auditory ANT; 2) to determine which factors influence auditory spatial orienting in non-spatial tasks; and 3) to directly compare spatial orienting effects obtained using three different stimulus presentation methods. The first two aims were addressed in a series of six cueing experiments in which target stimuli and cueing protocols were varied. The third aim was addressed through a within-subjects experiment in which monaural, binaural ITD, and freefield presentation methods were contrasted.

Factors Influencing Auditory Spatial Orienting

The first two experiments were conducted to address possible methodological problems with the auditory ANT. Experiment 1 removed the fixation tone used in the auditory ANT and tested performance at three different SOAs. The results showed stronger alerting effects than were found in the auditory ANT, suggesting that the fixation tone may have attenuated the benefit obtained from the neutral (alerting) cues. However, while a significant spatial orienting effect was found, it was only slightly larger than that found in the auditory ANT, and only present at one SOA. Experiment 2 investigated whether subjects were able to alert and to orient to the onset of the linguistic target stimuli used in the auditory ANT and the first experiment of this chapter. Steady-state target stimuli were used in place of linguistic stimuli. The results showed a large and significant exogenous spatial orienting effect at the shortest SOA, which suggests that the onset of the linguistic stimuli may have been a useful orienting cue at the shortest SOA. The striking effect from both of these studies was that neither produced large, robust spatial-orienting benefits at the longer SOAs where endogenous (voluntary) orienting effects would be expected.

The spatial relevance hypothesis (McDonald & Ward, 1999) states that auditory spatial orienting should be possible in non-spatial tasks, as long as informative spatial cues are presented. Theoretically, informative spatial cues will make space relevant to task performance, and will therefore ensure that the auditory stimuli are encoded spatially. Once this has occurred, auditory spatial attention will be able to operate upon the spatially-encoded representation. The auditory ANT, and the first two experiments reported in this chapter, presented 100% valid spatial cues, which accurately predicted target location. Spatial orienting benefits would therefore have been expected according to the spatial relevance hypothesis. Experiments 3 and 4 attempted to increase subjects' motivation to process the task stimuli spatially, by making the spatial nature of the task more salient. Experiment 3 increased the proportion of trials on which a spatial cue was presented, and Experiment 4 contrasted performance with valid and invalid cues to target location. Neither experiment was particularly successful in generating large or robust endogenous spatial orienting benefits. Experiments 5 and 6 were designed to separate exogenous (automatic) and endogenous (voluntary) cueing effects. Experiment 5 presented uninformative cues, and found exogenous orienting benefits, but no endogenous orienting effects. The complete absence of orienting effects at the longer (450- and 750-

ms) SOAs demonstrated that subjects in the previous experiments were gaining some benefit from the spatial cues, even though the effects were small and weak. Experiment 6 presented informative central symbolic cues, and found endogenous orienting effects, but no exogenous orienting. The results from these two studies taken together demonstrate that the exogenous orienting effect is reliable in studies presenting peripheral cues (presented at target location), and that while the endogenous orienting effect is weak in all studies presenting informative spatial cues, overall subjects are able to obtain some degree of benefit from cues to target location. The results from all six experiments provide partial support for the spatial relevance hypothesis. There is some indication that subjects gain a spatial-cue benefit in non-spatial tasks with informative cues. However, exogenous orienting effects were also found in a non-spatial task with uninformative cues. In addition, the spatial relevance hypothesis does not explain why the endogenous orienting effects were small and weak. It therefore appears that while the spatial relevance hypothesis goes some way towards explaining the variability in auditory spatial cueing studies, it does not form a comprehensive theory which can address all findings.

Exogenous Orienting

Exogenous orienting benefits were found in all experiments which presented peripheral spatial cues, except for Experiment 1, in which the linguistic stimuli may have attenuated any spatial-cue effects. Exogenous orienting was particularly unexpected in Experiment 5, in which uninformative cues and a non-spatial task meant that there was no motivation for subjects to encode task stimuli spatially. The spatial relevance hypothesis would have predicted no benefits in this experiment, and is supported by the null results found by Spence and Driver (1994) and McDonald and Ward (1999). However, Mondor and Breau (1999) and Mondor and Amirault (1998) did find spatial orienting benefits in uninformative cueing studies using non-spatial tasks.

It is possible that robust exogenous auditory orienting effects result from involvement of the superior colliculus, which is associated with reflexive head and eye movements (Sparks, 2002). The superior colliculus contains spatiotopic maps of visual, tactile, and auditory space (Meredith & Stein, 1986). It is therefore possible that reflexive orienting to the source of auditory stimuli is mediated by the superior colliculus. This explanation would account for the robust exogenous orienting effects found in these experiments. The effect does not seem to be strongly altered by task demands, supporting a role for a reflexive mechanism which does not depend on optional localisation of task stimuli. Unfortunately there is no readily apparent explanation for Spence and Driver's (1994) and McDonald and Ward's (1999) failure to find exogenous orienting effects with their uninformative cueing studies using non-spatial tasks.

Endogenous Orienting

Endogenous orienting effects found in these experiments do not appear to be highly robust. All experiments which presented informative cues produced average endogenous orienting benefits at the 450-ms SOA, although none of these effects were large or robust. Similarly, all but one of these experiments produced endogenous orienting benefits at the 750-ms SOA, but again these effects were not particularly large or robust, and the remaining study produced a spatial-cue cost. However, compared with the lack of effects found at longer SOAs in the uninformative cueing paradigm (Experiment 5), it is clear that at least some subjects were able to benefit from informative cues to target location. Individual subject data show large differences between the amount of benefit obtained by different subjects, and sometimes between the amount of benefit an individual subject was able to obtain at different SOAs. Since the overall spatial-cue benefit was larger in the informative-cue experiments than in the uninformative-cue experiment, it seems likely that subjects were attempting to orient their attention to the cued location. The variation therefore appears to reflect difficulty in doing so effectively. The spatial relevance hypothesis offers no explanation as to why spatial orienting might be more variable in the auditory modality than in the visual modality. An obvious explanation relates to the coding of auditory and visual information. Visual information is coded spatiotopically, and the foveal-peripheral organisation of the visual system, whereby foveated (fixated) information is processed with higher spatial acuity, promotes orienting to spatial location. Auditory information is coded tonotopically, and while there is some variation in acuity with spatial location, it is a subtle effect relative to that found in the visual system. The auditory system is more suited to act as an early-warning system, in which detecting the presence of a stimulus is more important than identifying its location. The lack of spatiotopic representation of auditory stimuli in the cortex might be reflected in the difficulty subjects experience in trying to orient their attention to the source of auditory stimuli. Current evidence suggests that accurate auditory localisation is achieved through the rate or pattern of neural firing (Middlebrooks, 2000), or through the relative activity in two broadly-tuned hemispheric channels (Boehnke & Phillips, 1999; McAlpine et al., 2001). If attention is a supramodal mechanism, it is difficult to envisage how it could operate upon either type of representation as efficiently as it does upon visual spatiotopic representations. Similarly, if attention is a modality-specific resource, the exact manner by which it might enhance processing at attended locations is not clear.

An alternative interpretation of these results is that the reaction-time measure is simply inappropriate for detecting spatial orienting in the auditory modality. Schröger and Eimer (1996) recorded both RT and ERP correlates of auditory spatial orienting. The ERP data indicated that spatial orienting had occurred, while the RT data did not. Schröger and Eimer proposed that the orienting effect was present, but that differences in the sensitivity of the two measures led to divergent results. Sach et al. (2000) were also concerned that the RT measure might not be appropriate, and used a detectability (d') outcome measure instead. Their reasoning was that a detectability measure is ideally suited to detecting changes in thresholds and sensitivity, and that it is "not clear that the primary determinants of RT are low level and sensory" (p. 717). It should be noted though that even with a different type of outcome measure, some of Sach et al.'s spatial orienting experiments also produced variable results across subjects.

Influence of Stimulus Presentation Method

Auditory spatial-orienting effects have been found with a range of different stimulus-presentation techniques. Stimuli presented in freefield offer the greatest wealth of cues with which to localise sounds (ITD, ILD, and spectral cues). Stimuli presented over headphones are typically presented either monaurally, or are lateralised using ITD or ILD cues. With monaural headphone presentation it is possible that listeners are able to orient attention to an ear, rather than to a genuine spatial location. This is also possible if stimuli are presented from a speaker directed towards one ear (at +/-90° azimuth). The final experiment in this chapter directly compared spatial orienting effects with stimuli presented in freefield and over headphones. Freefield presentation was from speakers located directly opposite each ear. In the headphone presentation conditions, stimuli were lateralised to the left and right using monaural presentation, and using binaural presentation with an ITD of +/-600 μs. The results show spatial orienting benefits in the freefield and monaural conditions, but not in the condition where stimuli were lateralised using ITDs. This pattern suggests that listeners were attending to an ear, rather than to a spatial location, but there are two alternative explanations. First, that listeners did not gain a strong percept of lateralisation in the ITD condition, and were therefore less motivated to attend to the cued location. Second, that listeners were unable to orient attention to an internal sound source. These possibilities cannot be separated in this experiment, but could be examined by positioning stimuli at a smaller azimuthal angle (e.g. at +/-45° from the midline). Stimuli presented from speakers located at 45° azimuth would still favour a better-ear strategy, as the signal would still be louder at the ear closest to the sound source. However, the strategy might be less effective than when stimuli are located directly opposite one ear. Stimuli lateralised to 45° azimuth using ILDs would also favour a better-ear strategy. As with ITDs, sounds lateralised using ILDs appear to originate from sources located inside the head. By comparing performance with sounds lateralised to 45° azimuth using ILDs and ITDs, it would therefore be possible to determine whether the null effect with ITDs in Experiment 7 related to a better-ear strategy or to difficulties orienting to an internal sound source.

Conclusions

Exogenous auditory spatial orienting appears to be a robust phenomenon, insensitive to differences in task design. Exogenous orienting may be mediated by the superior colliculus, which is known to contain spatiotopic maps of auditory and visual space, and is involved in reflexive head and eye movements. In contrast, endogenous auditory spatial orienting appears to be a weak effect with large inter-subject variability. The lack of robust voluntary orienting effects might reflect the way in which auditory information is coded in the cortex. A direct comparison of auditory spatial orienting effects using different presentation methods provided evidence that subjects may be using a 'better-ear' strategy, in which one ear is attended rather than a genuine spatial location. However, alternative explanations related to the spatial precision with which stimuli were localised using ITD cues, and difficulty orienting to an internal sound source, might also account for this effect. Experiments designed to address this question are currently being conducted at IHR.

Chapter 6: Summary of Findings and Discussion

The research reported in this thesis investigated several key aspects of auditory attention using behavioural and neuroimaging methods. In this chapter, the main findings will be summarised, and then discussed in light of the research aims. Some further research directions will also be considered. To begin with, the aims of the thesis are recapped.

Research Aims

The thesis aimed to investigate auditory attention. Specifically, the attentional skills of alerting, orienting, and attentional control were assessed using behavioural and neuroimaging methods. There were both theoretical and applied motivations for this research. Theoretically, it is not yet clear whether attention is a supramodal facility, or whether there are modality-specific attentional resources. In addition, auditory attention is under-researched relative to visual attention, and merits further research in its own right. In the applied field, there is emerging evidence that certain clinical groups may experience difficulty with situations involving auditory attention (e.g. Gatehouse & Noble, 2004). As yet, there are no clinical tests designed to evaluate auditory attention skills. An improved understanding of auditory attention would enable an informed decision to be made as to whether clinical tests of visual attention, such as the ANT, are appropriate for evaluating auditory attention deficits. The thesis therefore had two key aims: first, to contribute to an understanding of auditory attention, and how it relates to visual attention; and second, to develop a test of auditory attention that could be used with clinical groups.

Summary of Findings

Chapter 2: The Attention Network Test

The attention network test (ANT) (Fan et al., 2002) is a single test of visual attention which separately evaluates the attentional skills of alerting, spatial orienting, and executive control. A cueing task (Posner, 1978) is used to obtain measures of subjects' ability to increase their alertness in response to a warning cue (alerting), and to orient their attention to a cued location (orienting). Executive control is assessed using a flanker task (Eriksen & Eriksen, 1974), which gives a measure of subjects' ability to resolve conflict. An auditory analogue of the test was developed in which alerting and orienting were assessed using cues, as in the visual task, and executive control was assessed using an auditory Stroop task. The same group of subjects participated in both the auditory and visual ANTs. The results showed more variability in the auditory measures of all three types of attention, compared with the visual measures. Despite this, subjects gained a significant benefit from an alerting cue in both auditory and visual tasks, and the effects were of a similar magnitude. The executive control (conflict resolution) measures were also of a similar magnitude in both auditory and visual tests, and were significantly correlated across subjects. In contrast, while subjects gained a significant benefit from spatial orienting cues in the visual modality, there were no corresponding benefits in the auditory modality. The results from this study formed the basis for much of the rest of the thesis, and the experiments that followed were designed to investigate the auditory attentional skills of alerting, orienting, and executive control in more detail.

Chapter 3: Neural Correlates of Auditory and Visual Conflict Resolution

An fMRI study was conducted to investigate whether the similar (and correlated) conflict-resolution measures obtained in the auditory and visual ANTs were the result of common cortical mechanisms. Initially, a meta-analysis of conflict-resolution studies was conducted to identify regions commonly associated with conflict monitoring and resolution. The meta-analysis identified a number of regions, including anterior cingulate cortex (ACC), bilateral dorsolateral prefrontal cortex (DLPFC), bilateral insula, and bilateral parietal lobe. There was also some suggestion that activation in DLPFC was task-dependent, while ACC activation appeared to be independent of specific task demands. An fMRI study was then conducted, in which subjects took part in a visual colour-word Stroop task and an auditory pitch-word Stroop task. Activation on incongruent trials (in which subjects had to resolve conflict) was contrasted with activation on neutral trials (in which there was no conflict). Overlapping activation from the visual and auditory tasks was found in ACC and bilateral DLPFC, consistent with a supramodal anterior network for conflict monitoring and resolution. Activation in the parietal lobe appeared segregated, which may have reflected either differential flow of information from sensory to higher-order areas, task-specific processing, or modality-specific selective attention processes.

Chapter 4: Orienting to Spatial and Non-Spatial Stimulus Features

The Vowels study used a cueing paradigm to investigate alerting and orienting using concurrently presented synthesised vowels as target stimuli. The task enabled four types of attention to be addressed: alerting, orienting to a spatial location, orienting to a pitch, and orienting to both a location and a pitch. In addition, these effects were investigated in the presence of a distractor stimulus. The data showed significant alerting effects with SOAs of 450 and 1050 ms, and significant location, pitch, and combined location-and-pitch orienting effects at the 1050-ms SOA. There were no additive benefits from having a cue to both location and pitch, suggesting that attention may be directed towards an auditory object encompassing both location and pitch information. The lack of orienting effects at the 450-ms SOA might reflect an interaction between alerting and orienting mechanisms. Orienting cues additionally provided an alerting cue, and so pure orienting measures were obtained by subtracting benefits obtained from an alerting cue from benefits obtained from alerting and orienting cues. This method assumes independence between measures of alerting and orienting, but at the 450-ms SOA there were significant negative correlations between the alerting and orienting measures, indicating a lack of independence. The experiment also revealed a spatial conflict effect, in which subjects were slower to respond when the target and response were located on opposite sides than when they were located on the same side, and a contrast effect, whereby novel stimuli were easier to detect when presented amid a sequence of repeated stimuli.

Chapter 5: Spatial Orienting of Auditory Attention: Effect of Different Cueing Strategies

A series of seven cueing experiments investigated possible explanations for the variable results found in tests of auditory spatial orienting. The spatial relevance hypothesis states that spatial orienting benefits will be found when space is relevant to task performance. Informative spatial cues should have been sufficient to meet this criterion, but the auditory ANT found no evidence for auditory spatial orienting, while the Vowels task only found spatial orienting benefits at a relatively long SOA. A review of the literature revealed several other inconsistencies with the spatial relevance hypothesis, and some suggestion that there might be high variability across subjects. Experiment 1 replicated the auditory ANT but with some methodological changes. Experiments 2 to 6 used steady-state target stimuli, and a range of cueing protocols. These studies revealed two key effects. First, that exogenous auditory spatial orienting is a robust phenomenon which occurs in response to peripheral spatial cues at short SOAs. This effect may reflect the involvement of the superior colliculus, which contains spatiotopic maps of auditory, visual, and tactile space, and is responsible for reflexive head and eye movements. Second, endogenous auditory spatial orienting is a weak effect, with large inter-subject variability. Voluntarily orienting attention to the source of auditory stimuli might be difficult given the non-spatiotopic representation of auditory information in the cortex. The final experiment in Chapter 5 investigated the influence of stimulus presentation methods. Auditory spatial orienting effects were found with freefield and monaural headphone presentation, but not when stimuli were presented binaurally over headphones and lateralised using ITDs. This suggests that subjects were attending to an ear, rather than to a spatial location, but alternative explanations are that subjects did not gain such a strong percept of laterality in the ITD condition, or that subjects were unable to orient to an intracranial location.

Discussion

Theoretical Implications: Attention as a supramodal resource

Alerting

Auditory alerting was evaluated by the auditory and visual ANTs (Chapter 2), the Vowels experiment (Chapter 4), and cueing Experiments 1 and 2 (Chapter 5). Subjects gained a similar amount of benefit from alerting cues in the auditory and visual ANTs, although the effect was more variable in the auditory task than in the visual task. Alerting benefits found in the Vowels task and in cueing Experiment 2 increased with increased time between cue and target onsets, suggesting that alerting benefits in these studies were slow to build. In contrast, the alerting benefit found in cueing Experiment 1 was of a similar magnitude at all SOAs, although this study did use linguistic stimuli which may have provided some additional alerting benefit. The literature reviewed in the introductory chapter suggested that alerting may be a supramodal facility, since similar neural activation is found when subjects perform alerting tasks with stimuli presented in different modalities (Kinomura et al., 1996; Sturm & Willmes, 2001). The results from the studies reported in this thesis are consistent with this conclusion. While there was some suggestion that alerting benefits might be more variable in the auditory modality than in the visual modality, auditory alerting benefits were detected in all studies in which they were investigated and therefore appear to be robust.

Attentional Control

Behavioural measures of attentional control were obtained in the visual ANT (flanker task); auditory ANT (pitch-word Stroop task), colour-word Stroop task, and Vowels task (spatial conflict). Neural correlates of auditory and visual Stroop conflict were investigated using fMRI (Chapter 3). The results suggest both supramodal and intramodal elements of overcoming conflict. Behaviourally, measures of auditory attentional control appear to be robust, with significant effects being found in all tasks, although, as was found with alerting, there was more variability in measures of auditory attentional control than visual attentional control. The fMRI study revealed areas of overlapping activation associated with auditory and visual conflict monitoring and resolution, and areas of modality-specific activation. The slight discrepancies in the behavioural data, combined with the pattern of results from the fMRI study, suggest both supramodal and intramodal elements of attentional control. One way to interpret this finding is to consider the operations being conducted during conflict resolution. Some elements, such as monitoring for conflict and inhibiting incorrect responses, operate at a high level of processing, or on the response itself. These operations might be expected to be supramodal. In contrast, any attempts to process the stimuli differentially in order to overcome conflict at the perceptual level might be expected to be modality-dependent. Since both types of operation might be involved during conflict tasks, it is reasonable to expect both supra- and intra-modal components to attentional control.

Orienting

While the visual ANT elicited reliable spatial orienting benefits, there were no comparable benefits associated with spatial orienting in the auditory ANT (Chapter 2). Auditory spatial orienting was then investigated further in the Vowels experiment (Chapter 4) and a series of seven cueing experiments (Chapter 5). The main finding from these studies was that while exogenous auditory spatial orienting is a reliable effect which is relatively insensitive to task differences, endogenous auditory spatial orienting is highly variable across subjects, and therefore not very robust. There has been some debate over whether attention is a supramodal facility (Farah et al., 1989) or whether there are separate attentional facilities for each perceptual modality (Wickens, 1980). Crossmodal studies have provided evidence that contradicts both of these extreme hypotheses. For example, it has been shown that subjects can only direct attention to one location in one modality and to a different location in a different modality, under certain circumstances (Spence & Driver, 1996). Alternative hypotheses are that there are either separate-but-linked attentional facilities (Spence & Driver, 1996), or that there are separate attentional facilities which subserve a higher-level supramodal facility (Posner, 1990). Typically, this debate is informed by crossmodal studies of spatial orienting. However, unimodal investigations of auditory spatial orienting might also contribute. The experiments conducted for this thesis produced reliable exogenous orienting effects, but weak endogenous orienting effects. One interpretation of this finding is that the weak endogenous spatial orienting effect does not arise from different attentional mechanisms, but from an interaction between a supramodal attentional mechanism (or separate attentional mechanisms which operate in a similar manner) and non-spatiotopic representations of the auditory stimuli. It seems important that any efforts to investigate whether spatial attention is supra- or intra-modal should remain aware that attention must operate upon very different neural representations of task stimuli when they are presented in different modalities.

Theoretical Implications: Auditory spatial orienting

The above discussion is based on the premise that auditory spatial orienting is the auditory equivalent of visual spatial orienting. However, more appropriate auditory analogues to visual space might be frequency or time. While visual information is coded spatiotopically, both on the sensory epithelia and in the cortex, auditory information is processed tonotopically. Physiologically, therefore, orienting auditory attention to a pitch might operate in a manner analogous to orienting visual attention to a spatial location. Kubovy and Van Valkenburg (Kubovy & Van Valkenburg, 2001; Van Valkenburg & Kubovy, 2003) propose a 'theory of indispensable attributes', which they use to specify what constitutes an object in the auditory and visual modalities. The theory states than an attribute is indispensable if it is necessary for the perception of more than one object. For example, if two differently coloured spotlights are directed to a surface they will be perceived as one light if they are directed to the same location, and two different lights if they are directed to different locations. Space is therefore an indispensable attribute of visual perception, according to the theory, while colour is not. Similarly, if sounds of the same frequency are presented from two speakers, a single sound will be heard, while if two different frequencies are presented from the same speaker, two sounds could be heard. Time is considered to be an indispensable attribute of both the visual and auditory modalities. Temporal information has a critical role in auditory processing, in many ways analogous to the role that spatial information has in vision. In interpreting a visual scene, much of the information is static, and can be processed by orienting to spatial locations of interest. In contrast, auditory scenes are dynamic, constantly changing over time. A practical example is that of language perception. In order to understand speech it is necessary to process both spectral and temporal features of the signal, whereas written information is presented in a static display which can be processed without regard to time.

There is some evidence to suggest that the same neural regions are involved in attending to both spatial and temporal features of a stimulus. Orienting to a temporal interval has been shown to activate a number of regions also active during orienting to a spatial location (Coull & Nobre, 1998). In addition, evidence from patients with right-hemisphere lesions following stroke damage provides evidence for a common basis for spatial and temporal attention (Husain & Rorden, 2003). Following right-hemisphere stroke, patients can experience two related spatial attentional deficits: unilateral neglect, in which they fail to attend to the contralesional side of space, and extinction, in which they fail to detect the stimulus presented further into the contralesional side of space when two stimuli are presented concurrently. Both types of spatial deficit have been found with both auditory and visual stimuli (Pavani, Ládavas, & Driver, 2003). In addition, non-spatial deficits have been detected in patients exhibiting neglect. Performance on an auditory sustained attention task was found to be significantly correlated with the severity of neglect symptoms in patients with right-hemisphere lesions (Robertson et al., 1997). Further, patients with neglect have been shown to perform worse on a frequency discrimination task than patients with right-hemisphere lesions but no neglect, even when stimuli are presented one at a time, and close to the midline (Cusack, Carlyon, & Robertson, 2000). In the visual modality, Husain and Rorden (2003) found that patients with damage to the right superior temporal gyrus (STG) and inferior parietal lobe (IPL) show an enhanced attentional blink, relative to healthy individuals. The attentional blink refers to subjects' inability to detect the second of two rapidly presented targets in a stream of non-targets. In healthy subjects, the attentional blink lasts for around 400 ms, while in patients with right-hemisphere STG and IPL damage the attentional blink was shown to last for more than 1200 ms. This indicates that the time taken for the first target to be processed, before attention is free to attend to a subsequent target, is considerably longer in these patients than in healthy subjects (Husain & Rorden, 2003). It therefore appears that there are links between the attentional systems involved in spatial neglect, and attentional systems involved in temporal and sustained attention, and that deficits in both types of attention are found in both auditory and visual modalities.

Applied Implications: Progress towards a test of auditory attention

The visual attention network test (ANT) was used as a start point from which to evaluate the auditory attention skills of alerting, orienting, and attentional control. Part of the motivation for this was the aim to create or identify a test that could be used with clinical groups to evaluate deficits in auditory attention skills. Elderly hearing-impaired adults who responded to the SSQ (Gatehouse & Noble, 2004) reported difficulties with situations involving auditory attention, such as following one person speaking and using the telephone at the same time, and following multi-talker conversations without missing the start of each new talker. These self-reported attentional difficulties correlated with the extent of their self-perceived handicap, even when controlling for their degree of hearing loss. Control of alertness, ability to selectively attend to one talker, and top-down attentional control over multiple inputs appear to be components of the difficulties Gatehouse and Noble's respondents report. A preliminary step towards evaluating these self-reported problems was to design an auditory analogue of the visual ANT and to test it alongside the visual ANT in a group of 40 healthy subjects.

Had the auditory and visual ANTs elicited the same results, with correlations between auditory and visual measures of each attentional skill, this would have been consistent with supramodal attentional resources. Under these circumstances, visual tests of attention, like the ANT, might have been considered appropriate for evaluating self-reported difficulties with auditory attention. There may even be advantages to using tests of visual attention, since they are more established than auditory tests, and would not be subject to interactions with degree of hearing loss. The latter reason is an important consideration given that Gatehouse and Noble's sample had all presented with a hearing impairment. Alternatively, had the auditory and visual ANTs produced different measures of the three attention skills being evaluated, which did not correlate across subjects, then this would have indicated differences in the way in which attention operates across modalities. These differences would not necessarily suggest intramodal attentional resources. They could instead result from methodological differences, or from differences in the interaction between perceptual processing and attentional control. Under these circumstances a test of visual attention would not necessarily be unreliable for investigating auditory attention skills, but the nature of the difference would need to be understood in order to interpret the results of a visual test of attention, and relate them to subjective reports of difficulties with auditory attention.

In fact, the visual and auditory ANTs produced a mixed pattern of results. Similar alerting measures were produced by the two tests, as were similar measures of attentional control. The implication is that a test of visual attention might be appropriate for evaluating auditory alerting and attentional control. However, the spatial orienting effect was very different in the auditory and visual ANTs, and further experiments have demonstrated that reaction-time measures of endogenous auditory spatial orienting are not robust, and therefore do not resemble measures of visual spatial orienting. Different spatial orienting effects in the auditory and visual modalities might indicate intramodal attentional facilities, which must be evaluated by modality-specific tests. Alternatively, different spatial orienting effects might indicate an interaction between a supramodal attentional facility and modality-specific neural representations of task stimuli. Unfortunately, the research conducted for this thesis did not identify a reliable test of endogenous auditory spatial orienting. Evaluation of auditory spatial orienting must therefore be conducted through reliable, established tests of visual spatial orienting, or through unreliable tests of auditory spatial orienting. Neither alternative is ideal, and one interpretation of being presented with these two inadequate options is that auditory spatial orienting is not in fact the attentional skill of interest.

Gatehouse and Noble's (2004) respondents did not specifically report difficulties in orienting attention to a spatial location, but rather difficulties in dividing attention across two talkers, and difficulty re-orienting attention to a new talker. While orienting to a spatial location is one mechanism which might be used in these situations, it is clearly not the only option for performing these tasks. Since target location is rarely the only cue available for auditory orienting in the real world, tests of spatial orienting in isolation may not be as beneficial as tests which investigate orienting to ecologically-valid stimuli. Subjects had no difficulty attending to just one stream of information in the dichotic listening task (e.g. Cherry, 1953), in which a number of cues were available for subjects to orient towards. Using more controlled stimuli, Darwin and Hukin (1999; 2000) have demonstrated the usefulness of dichotic listening tasks in determining which stimulus features can be used to attend to one of two competing sentences. Tests such as these might be more beneficial in testing listeners' ability to orient attention to one talker than tests which present artificial stimuli in which only one type of cue is available at a time. There is evidence that auditory attention can operate upon an auditory object comprising location and frequency information (Mondor, Zatorre et al., 1998; Zatorre et al., 1999), and that cues to pitch, prosody, vocal tract size, and location can all be used to direct attention to a talker (Darwin & Hukin, 2000). In an applied setting, a test which incorporates a variety of cues might be more informative about listeners' ability to orient attention than a pure test of spatial attention.

An alternative approach which is also more ecologically valid is to present target sounds in the presence of distractor sounds. If attention operates to suppress unattended stimuli, instead of (or in addition to) enhancing attended stimuli, then spatial orienting effects would be more readily found when targets are presented in combination with competing stimuli. The Vowels task presented concurrent distractor sounds and found spatial orienting benefits at the 1050-ms SOA. Another methodology in which subjects are required to separate target and distractor stimuli is the rhythmic masking release paradigm (Sach & Bailey, 2004), in which a target rhythm is combined with masking tones. Sach and Bailey successfully used this technique to demonstrate that masking is maximal when target and distractor tones are presented to the same perceived location, and therefore demonstrated that attention could be used to minimise the effect of masking tones when they were presented at different perceived locations.

Directions for Further Research

Since auditory attention in general remains poorly understood relative to visual attention, numerous investigations could be proposed here. I will restrict myself to three proposals which stem from the research reported in the thesis.

Neuroimaging techniques are a valuable tool for determining whether attention is a supramodal or modality-specific resource. In matched tests of auditory and visual spatial orienting, fMRI would identify both sources and sites active during the tasks. Overlapping activation in areas associated with top-down selective attention (such as around the intraparietal sulcus (Corbetta et al., 2000)), but segregated activation in primary sensory cortices, would suggest a supramodal network involved in spatial orienting, even if reaction-time measures of spatial orienting are not reliably found in the auditory task. Alternatively, segregated activation in areas identified as sources of selective attention would suggest modality-specific attentional facilities. If separate top-down sources of attention are found, the behavioural literature strongly suggests that there are crossmodal links between them, biasing attention to be oriented to the same location in all modalities.

Whether auditory spatial attention is directed to an ear or to a genuine spatial location was investigated in Experiment 7 (Chapter 5). However, this experiment was unable to discriminate between three possible explanations for the null spatial-orienting effect found when stimuli were lateralised using ITDs alone. The first possibility is that subjects were unable to direct their attention to a spatial location, and instead attended to the ear at which the stimulus was louder (an ear-selection strategy). The second possibility is that subjects were unable to orient attention to the intracranial sound source produced when stimuli are lateralised using ITD cues. Finally, subjects may have obtained a weak percept of lateralisation from the ITD cues alone. An experiment currently being conducted at IHR investigates these possibilities by presenting stimuli in freefield, and over headphones lateralised using three different methods: ITDs, ILDs, and a generic head-related transfer function (HRTF), which simulates ITD, ILD, and spectral cues. If subjects are using an ear-selection strategy, spatial orienting effects would be expected in the freefield, ILD, and HRTF conditions, in which the stimulus will be louder at one ear than the other. If however subjects simply experience difficulty orienting to an intracranial sound source or to a diffuse sound source, no spatial-orienting effects would be expected in the ILD condition, since stimuli in this condition will also be perceived internally and with a diffuse sound source. Preliminary results from this study show spatial-orienting benefits in all conditions except the ITD condition, consistent with the suggestion that subjects are using an ear-selection strategy. Further experiments could determine whether this result is task-specific. An ear-selection strategy may be particularly effective in this task, in which the target sound is presented in isolation and the task is non-spatial. It is possible that subjects attend to a genuine spatial location in other tasks, especially those with a spatial task or competing stimuli. The spatial-orienting benefits found in the Vowels study support this hypothesis.

The research reported in this thesis was conducted on healthy subjects, none of whom were expected to have attentional deficits. It would be beneficial to test the ANTs on a sample of elderly, hearing-impaired adults who do report difficulties with attentionally-demanding situations, such as those who responded on the SSQ (Gatehouse & Noble, 2004). By comparing reaction time and accuracy measures on the auditory and visual ANTs with responses on the SSQ, it would be possible to evaluate the effectiveness of the auditory and visual tests in providing an objective measure of respondents' self-reported difficulties. Using the ANTs in combination with the SSQ would allow the relative involvement of alerting, spatial orienting, and executive control to be evaluated, and in addition it would be possible to investigate any interaction between age, hearing impairment, and the attentional skills measured by the ANT.

Conclusions

The attentional skills of alerting, orienting, and attentional control were evaluated using behavioural and neuroimaging methods. The studies showed alerting to be comparable across auditory and visual tasks. This is consistent with previously reported research suggesting that alerting is a supramodal mechanism. Attentional control, evaluated through auditory and visual conflict tasks, showed both supramodal and modality-specific components. Spatial orienting of attention differed substantially between a visual task, which elicited reliable spatial orienting effects, and auditory tasks, which produced weak effects which varied across subjects. The spatial-orienting experiments reported here contribute to a small and variable literature, and provide evidence that the spatial relevance hypothesis is insufficient to account for all results found in auditory spatial orienting studies. It is hypothesised that the relative unimportance of spatial location in auditory processing might account for the unreliable spatial orienting effects found with non-spatial tasks. Differences between auditory and visual spatial orienting effects may result from an interaction between top-down attentional influences, and the non-spatiotopic representation of auditory information in the cortex.

References

Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., et al. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage, 16*, 61-75.

Alho, K., Medvedev, S. V., Pakhomov, S. V., Roudas, M. S., Tervaniemi, M., Reinikainen, K., et al. (1999). Selective tuning of the left and right auditory cortices during spatially directed attention. *Cognitive Brain Research, 7*, 335-341.

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In A. F. Sanders & H. Heuer (Eds.), *Perspectives on perception and action* (pp. 395-419). Hillsdale, NJ: Erlbaum.

Arbogast, T. L., & Kidd, G. (2000). Evidence for spatial tuning in informational masking using the probe-signal method. *Journal of the Acoustical Society of America, 108*(4), 1803-1810.

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

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron, 41*, 473-487.

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience, 12*(6), 988-1000.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cerebral Cortex, 11*, 837-848.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12*(2), 298-309.

Bédard, M. A., El Massioui, F., Pillon, B., & Nandrino, J. L. (1993). Time for reorienting of attention: a premotor hypothesis of the underlying mechanism. *Neuropsychologia, 31*(3), 241-249.

Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., et al. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia, 31*(9), 907-922.

Blauert, J., & Lindeman, W. (1986). Spatial mapping of intracranial auditory events for various degrees of interaural coherence. *Journal of the Acoustical Society of America, 79*(3), 806-813.

Boehnke, S. E., & Phillips, D. P. (1999). Azimuthal tuning of human perceptual channels for sound location. *Journal of the Acoustical Society of America, 106*(4), 1948-1955.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*, 179-181.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*(12), 539-546.

Bowtell, R., & Peters, A. (1999). Analytic approach to the design of transverse gradient coils with co-axial return paths. *Magnetic Resonance in Medicine, 41*, 600-608.

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *NeuroImage, 14*, 1416-1423.

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

Broadbent, D. E. (1958). *Perception and communication.* Oxford: Oxford University Press.

Brown, G. G., Kindermann, S. S., Siegle, G. J., Granholm, E., Wong, E. C., & Buxton, R. B. (1999). Brain activation and pupil response during covert performance of the Stroop color word task. *Journal of the International Neuropsychological Society, 5*, 308-319.

Büchel, C., Josephs, O., Rees, G., Turner, R., Frith, C. D., & Friston, K. J. (1998). The functional anatomy of attention to visual motion. *Brain, 121*, 1281-1294.

Buchtel, H. A., & Butter, C. M. (1988). Spatial attentional shifts: implications for the role of polysensory mechanisms. *Neuropsychologia, 26*(4), 499-509.

Buchtel, H. A., Butter, C. M., & Ayvasik, B. (1996). Effects of stimulus source and intensity on covert orientation to auditory stimuli. *Neuropsychologia, 34*(10), 979-985.

Bunge, S., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron, 33*, 301-311.

Bush, G., Frazier, J. A., Rauch, S. L., Seidman, L. J., Whalen, P. J., Jenike, M. A., et al. (1999). Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the counting Stroop. *Biological Psychiatry, 45*, 1542-1552.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialised for functional neuroimaging - validation study with functional MRI. *Human Brain Mapping, 6*, 270-282.

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M.-J., Tian, B., Rauschecker, J. P., et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localisation in humans. *Nature Neuroscience, 2*(8), 759-766.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience, 12*(1), 1-47.

Callejas, A., Lupiáñez, J., & Tudela, P. (2004). The three attentional networks: On their independence and interactions. *Brain and Cognition, 54*, 225-227.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*, 747-749.

Carter, C. S., MacDonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D. C., et al. (2000). Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *PNAS, 97*(4), 1944-1948.

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An $H_{20}{}^{15}O$ PET study of Stroop task performance. *NeuroImage, 2*, 264-272.

Casey, B. J., Thomas, K. M., Welsh, T. F., Badgaiyan, R. D., Eccard, C. H., Jennings, J. R., et al. (2000). Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. *PNAS, 97*(15), 8728-8733.

Chambers, C. D., Stokes, M. G., & Mattingley, J. B. (2004). Modality-specific control of strategic spatial attention in parietal cortex. *Neuron, 44*, 925-930.

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.

Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews: Neuroscience, 3*, 553-562.

Cohen, Y. E., & Knudsen, E. I. (1999). Maps versus clusters: different representations of auditory space in the midbrain and forebrain. *Trends in Neurosciences, 22*(3), 128-135.

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., et al. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional Stroop tasks. *Cognitive, Affective, & Behavioral Neuroscience, 3*(2), 81-96.

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*, 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. *The Journal of Neuroscience, 11*(8), 2383-2402.

Coull, J. T. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology, 55*, 343-361.

Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia, 34*(11), 1085-1095.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience, 18*(18), 7426-7435.

Coull, J. T., Nobre, A. C., & Frith, C. D. (2001). The noradrenergic alpha2 agonist clonidine modulates behavioural and neuroanatomical correlates of human attentional orienting and alerting. *Cerebral Cortex, 11*, 73-84.

Craft, J. L., & Simon, J. R. (1970). Processing symbolic information from a visual display - interference from an irrelevant directional cue. *Journal of Experimental Psychology, 83*(3), 415-420.

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*(2), 785.

Culling, J. F., Summerfield, Q., & Marshall, D. H. (1994). Effects of simulated reverberation on the use of binaural cues and fundamental-frequency differences for separating concurrent vowels. *Speech Communication, 14*, 71-95.

Cusack, R., Carlyon, R. P., & Robertson, I. H. (2000). Neglect between but not within auditory objects. *Journal of Cognitive Neuroscience, 12*(6), 1056-1065.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378*, 279-281.

Darwin, C. J., & Hukin, R. W. (1999). Auditory objects of attention: The role of interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 617-629.

Darwin, C. J., & Hukin, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *Journal of the Acoustical Society of America, 107*(2), 970-977.

de Zubicaray, G. I., Wilson, S. J., McMahon, K. L., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: An event-related fMRI study employing overt responses. *Human Brain Mapping, 14*, 218-227.

Dehaene, S., Le Clec'H, G., Poline, J.-B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *NeuroReport, 13*, 321-325.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*(1), 51-61.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*, 53-78.

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

Duncan, J., Humphreys, G. W., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology, 7*, 255-261.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*, 475-483.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*(2), 161-177.

Eimer, M. (1999). Can attention be directed to opposite locations in different modalities? An ERP study. *Clinical Neurophysiology, 110*, 1252-1259.

Eimer, M., & van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: Evidence from event-related potentials. *Psychophysiology, 39*, 437-449.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143-149.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). *3D statistical neuroanatomical models from 305 MRI volumes.* Paper presented at the Proceedings of IEEE-Nuclear Science Symposium and Medical Imaging Conference, Piscataway, NJ.

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and brain consequences of conflict. *NeuroImage, 18*, 42-57.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26*, 471-479.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*(3), 340-347.

Farah, M. J., Wong, A. B., Monheit, M. A., & Morrow, L. A. (1989). Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal? *Neuropsychologia, 27*(4), 461-470.

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *PNAS, 95*(3), 914-921.

Fossella, J. A., Sommer, T., Fan, J., Wu, Y., Swanson, J. M., Pfaff, D. W., et al. (2002). Assessing the molecular genetics of attention networks. *BMC Neuroscience, 3*(14), 1-11.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 3*, 165-189.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., et al. (1995). Analysis of fMRI time-series revisited. *NeuroImage, 2*, 45-53.

Frith, C. D., & Friston, K. J. (1996). The role of the thalamus in 'top down' modulation of attention to sound. *NeuroImage, 4*, 210-215.

Gatehouse, S., & Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology, 43*, 85-99.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science, 4*(6), 385-390.

George, M. S., Ketter, T. A., Parekh, P. I., Rosinsky, N., Ring, H., Casey, B. J., et al. (1994). Regional brain activity when selecting a response despite interference: An H215O PET study of the Stroop and an emotional Stroop. *Human Brain Mapping, 1*, 194-209.

George, M. S., Ketter, T. A., Parekh, P. I., Rosinsky, N., Ring, H. A., Pazzaglia, P. J., et al. (1997). Blunted left cingulate activation in mood disorder subjects during a response interference task (the Stroop). *The Journal of Neuropsychiatry and Clinical Neurosciences, 9*, 55-63.

Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., Kim, Y.-H., Meyer, J. R., et al. (1999). A large-scale distributed network for covert spatial attention. *Brain, 122*, 1093-1106.

Golob, E. J., Pratt, H., & Starr, A. (2002). Preparatory slow potentials and event-related potentials in an auditory cued attention task. *Clinical Neurophysiology, 113*, 1544-1557.

Green, E. J., & Barber, P. J. (1983). Interference effects in an auditory Stroop task: Congruence and correspondence. *Acta Psychologica, 53*, 183-194.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience, 15*(8), 1176-1194.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 7*(5), 555-562.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends in Cognitive Sciences, 7*(3), 125-133.

Hazeltine, E., Poldrack, R., & Gabrieli, J. D. E. (2000). Neural activation during response competition. *Journal of Cognitive Neuroscience, 12*(Supplement 2), 118-129.

Hill, N. I., & Darwin, C. J. (1996). Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning. *Journal of the Acoustical Society of America, 100*(4), 2352-2364.

Hillis, A. E., Newhart, M., Heidler, J., Barker, P. B., Herskovits, E. H., & Degaonkar, M. (2005). Anatomy of spatial attention: Insights from perfusion imaging and hemispatial neglect in acute stroke. *The Journal of Neuroscience, 25*(12), 3161-3167.

Hugdahl, K., & Nordby, H. (1994). Electrophysiological correlates to cued attentional shifts in the visual and auditory modalities. *Behavioral and Neural Biology, 62*, 21-32.

Husain, M., & Rorden, C. (2003). Non-spatially lateralized mechanisms in hemispatial neglect. *Nature Reviews: Neuroscience, 4*, 26-36.

James, W. (1890). *Principles of Psychology.* New York: Holt.

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

Jerger, J., & Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing disorders in school-aged children. *Journal of the American Academy of Audiologists, 11*, 467-474.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews: Neuroscience, 1*, 91-100.

Karnath, H.-O., Ferber, S., & Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature, 411*, 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. *Neuron, 22*, 751-761.

Keith, R. W. (2000). Diagnosing central auditory processing disorders in children. In R. J. Roeser, H. Hosford-Dunn & M. Valente (Eds.), *Audiology: Diagnosis* (pp. 337-353). New York: Thieme Medical Publishers.

Kerns, J. G., Cohen, J. D., MacDonald, A. M., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*, 1023-1026.

Kinomura, S., Larsson, J., Gulyas, B., & Roland, P. E. (1996). Activation by attention of the human reticular formation and thalamic intralaminar nuclei. *Science, 271*(5248), 512-515.

Klatt, D. H. (1980). Software for a cascade / parallel formant synthesizer. *Journal of the Acoustical Society of America, 67*, 838-844.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*(4), 138-147.

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

Laarni, J. (1999). Allocating attention in the visual field: The effects of cue type and target-distractor confusability. *Acta Psychologica, 103*, 281-294.

LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance, 9*(3), 371-379.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451-468.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56*(2), 183-197.

Liu, X., Banich, M. T., Jacobson, B. L., & Tanabe, J. L. (2004). Common and distinct neural substrates of attentional control in an integrated Simon and spatial Stroop task as assessed by event-related fMRI. *NeuroImage, 22*, 1097-1106.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4*(11), 432-440.

Macaluso, E., Eimer, M., Frith, C. D., & Driver, J. (2003). Preparatory states in crossmodal spatial attention: Spatial specificity and possible control mechanisms. *Experimental Brain Research, 149*, 62-74.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4*(10), 383-391.

Maclin, E. L., Gratton, G., & Fabiani, M. (2001). Visual spatial localization conflict: an fMRI study. *NeuroReport, 12*(16), 3633-3636.

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., & Robertson, I. H. (2001). The differential assessment of children's attention: the Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *Journal of Child Psychology and Psychiatry, 42*(8), 1065-1081.

Marrocco, R. T., & Davidson, M. C. (1998). Neurochemistry of attention. In R. Parasuraman (Ed.), *The Attentive Brain*. Cambridge, MA: MIT Press.

Martínez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., et al. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience, 2*(4), 364-369.

Matthews, S. C., Paulus, M. P., Simmons, A. N., Nelesen, R. A., & Dimsdale, J. E. (2004). Functional subdivisions within anterior cingulate cortex and their relationship to automatic nervous system function. *NeuroImage, 22*, 1151-1156.

McAlpine, D., Jiang, D., & Palmer, A. R. (2001). A neural code for low-frequency sound localization in mammals. *Nature Neuroscience, 4*(4), 396-401.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*(7), 293-299.

McClain, L. (1983). Stimulus-response compatibility affects auditory Stroop interference. *Perception & Psychophysics, 33*(3), 266-270.

McDonald, J. J., & Ward, L. M. (1999). Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. *Journal of Experimental Psychology: Human Perception and Performance, 25*(5), 1234-1252.

Mead, L. A., Mayer, A. R., Bobholz, J. A., Woodley, S. J., Cunningham, J. M., Hammeke, T. A., et al. (2002). Neural basis of the Stroop interference task: Response competition or selective attention? *Journal of the International Neuropsychological Society, 8*, 735-742.

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology, 56*(3), 640-662.

Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development, 75*(5), 1373-1386.

Middlebrooks, J. C. (2000). Cortical representations of auditory space. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, Massachusetts: The MIT Press.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research, 12*, 467-473.

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., et al. (2002). Attentional control in the aging brain: insights from an fMRI study of the Stroop task. *Brain and Cognition, 49*, 277-296.

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

Mondor, T. A., & Amirault, K. J. (1998). Effect of same- and different-modality spatial cues on auditory and visual target identification. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 745-755.

Mondor, T. A., & Breau, L. M. (1999). Facilitative and inhibitory effects of location and frequency cues: Evidence of a modulation in perceptual sensitivity. *Perception & Psychophysics, 61*(3), 438-444.

Mondor, T. A., Breau, L. M., & Milliken, B. (1998). Inhibitory processes in auditory selective attention: Evidence of location-based and frequency-based inhibition of return. *Perception & Psychophysics, 60*(2), 296-302.

Mondor, T. A., & Bregman, A. S. (1994). Allocating attention to frequency regions. *Perception & Psychophysics, 56*(3), 268-276.

Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(2), 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 and Performance, 24*(1), 66-79.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140.

Motter, B. C. (2000). Neurophysiology of visual attention. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, Massachusetts: The MIT Press.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence' in the auditory cortex. *Trends in Neurosciences, 24*(5), 283-288.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86*(3), 214-255.

Nee, D. E., Jonides, J., & Wager, T. D. (2004). *A meta-analysis of interference-resolution tasks.* Paper presented at the Meeting of the Cognitive Neuroscience Society.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In A. F. Sanders & H. Heuer (Eds.), *Perspectives on perception and action* (pp. 361-394). Hillsdale, NJ: Erlbaum.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*, 653-660.

Nishitani, N., Nagamine, T., & Shibasaki, H. (1998). Modality-specific subregions in human inferior parietal lobule: A magnetoencephalographic study during cognitive tasks. *Neuroscience Letters, 252*, 79-82.

Nobre, A. C. (2001). The attentive homunculus: Now you see it, now you don't. *Neuroscience and Behavioural Reviews, 25*, 477-496.

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting attention to locations in perceptual versus mental representations. *Journal of Cognitive Neuroscience, 16*(3), 363-373.

Nobre, A. C., Gitelman, D. R., Dias, E. C., & Mesulam, M. M. (2000). Covert visual spatial orienting and saccades: Overlapping neural systems. *NeuroImage, 11*, 210-216.

Norris, D. G., Zysset, S., Mildner, T., & Wiggins, C. J. (2002). An investigation of the value of spin-echo-based fMRI using a Stroop color-word matching task and EPI at 3T. *NeuroImage, 15*, 719-726.

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.

Pardo, J. V., Fox, P. T., & Raichle, M. E. (1991). Localization of a human system for sustained attention by positron emission tomography. *Nature, 349*, 61-64.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *PNAS, 87*, 256-259.

Paus, T., Zatorre, R. J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., et al. (1997). Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. *Journal of Cognitive Neuroscience, 9*(3), 392-408.

Pavani, F., Làdavas, E., & Driver, J. (2003). Auditory and multisensory aspects of visuospatial neglect. *Trends in Cognitive Sciences, 7*(9), 407-414.

Peelen, M. V., Heslenfeld, D. J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *NeuroImage, 22*, 822-830.

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of Stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. *Biological Psychiatry, 45*, 1237-1258.

Posner, M. I. (1978). *Chronometrics explorations of mind.* Hillsdale, NJ: Erlbaum.

Posner, M. I. (1990). Hierarchical distributed networks in the neuropsychology of selective attention. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics: Advances in models of cognitive function and impairment* (pp. 187-210). Hillsdale, NJ: Erlbaum.

Posner, M. I., & DiGirolamo, G. J. (1998). Executive attention: Conflict, target detection, and cognitive control. In R. Parasuraman (Ed.), *The Attentive Brain.* Cambridge, MA: MIT Press.

Posner, M. I., & Fan, J. (in press). Attention as an organ system. In J. Pomerantz (Ed.), *Neurobiology of Perception and Communication: From Synapse to Society. The IVth De Lange Conference.* Cambridge, UK: Cambridge University Press.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Posner, M. I., Rothbart, M. K., Vizueta, N., Levy, K. N., Evans, D. E., Thomas, K. M., et al. (2002). Attentional mechanisms of borderline personality disorder. *Proceedings of the National Academy of Science, USA, 99*(25), 16366-16370.

Potenza, M. N., Leung, H.-C., Blumberg, H. P., Peterson, B. S., Fulbright, R. K., Lacadie, C. M., et al. (2003). An fMRI Stroop task study of ventromedial prefrontal cortical function in pathological gamblers. *American Journal of Psychiatry, 160*, 1990-1994.

Prime, D. J., & Ward, L. M. (2002). Auditory frequency-based inhibition differs from spatial IOR. *Perception & Psychophysics, 64*(5), 771-784.

Quinlan, P. T., & Bailey, P. J. (1995). An examination of attentional control in the auditory modality: Further evidence for auditory orienting. *Perception & Psychophysics, 57*(5), 614-628.

Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science, 286*, 2504-2507.

Rhodes, G. (1987). Auditory attention and the representation of spatial information. *Perception & Psychophysics, 42*(1), 1-14.

Robertson, I. H., & Manly, T. (1999). Sustained attention deficits in time and space. In G. W. Humphreys, J. Duncan & A. Treisman (Eds.), *Attention, Space and Action* (pp. 298-310). Oxford: Oxford University Press.

Robertson, I. H., Manly, T., Beschin, N., Daini, R., Haeske-Dewick, H., Hömberg, V., et al. (1997). Auditory sustained attention is a marker of unilateral spatial neglect. *Neuropsychologia, 35*(12), 1527-1532.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). Structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society, 2*(6), 525-534.

Robinson, D. L., & Kertzman, C. (1995). Covert orienting of attention in macaques. III. Contributions of the superior colliculus. *Journal of Neurophysiology, 74*(2), 713-721.

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience, 2*(12), 1131-1136.

Rorden, C., & Driver, J. (1999). Does auditory attention shift in the direction of an upcoming saccade? *Neuropsychologia, 37*, 357-377.

Rosen, A. C., Rao, S. M., Caffarra, P., Scaglioni, A., Bobholz, J. A., Woodley, S. J., et al. (1999). Neural basis of endogenous and exogenous spatial orienting: a functional MRI study. *Journal of Cognitive Neuroscience, 11*(2), 135-152.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*, 1029-1040.

Ruff, C. C., Woodward, T. S., Laurens, K. R., & Liddle, P. F. (2001). The role of the anterior cingulate cortex in conflict processing: evidence from reverse Stroop interference. *NeuroImage, 14*, 1150-1158.

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

Sach, A. J., Hill, N. I., & Bailey, P. J. (2000). Auditory spatial attention using interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 717-729.

Scharf, B., Quigley, S., Aoki, C., Peachey, N., & Reeves, A. (1987). Focused auditory attention and frequency selectivity. *Perception & Psychophysics, 42*(3), 215-223.

Schmitt, M., Postma, A., & De Haan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *Quarterly Journal of Experimental Psychology, 53A*(1), 105-130.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition, 80*, 1-46.

Schröger, E., & Eimer, M. (1996). Effects of lateralized cues on the processing of lateralized auditory stimuli. *Biological Psychology, 43*, 203-226.

Schröger, E., & Eimer, M. (1997). Endogenous covert spatial orienting in audition: "Cost-benefit" analyses of reaction times and event-related potentials. *The Quarterly Journal of Experimental Psychology, 50A*(2), 457-474.

Simon, J. R., & Berbaum, K. (1990). Effect of conflicting cues on information processing: the 'Stroop effect' vs. the 'Simon effect'. *Acta Psychologica, 73*, 159-170.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*, 1657-1661.

Sparks, D. L. (1999). Conceptual issues related to the role of the superior colliculus in the control of gaze. *Current Opinion in Neurobiology, 9*, 698-707.

Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nature Reviews: Neuroscience, 3*(12), 952-964.

Spence, C., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20*(3), 555-574.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 1005-1030.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics, 59*(1), 1-22.

Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*(4), 1298-1319.

Steel, C., Haworth, E. J., Peters, E., Hemsley, D. R., Sharma, T., Gray, J. A., et al. (2001). Neuroimaging correlates of negative priming. *NeuroReport, 12*, 3619-3624.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Sturm, W., de Simone, A., Krause, B. J., Specht, K., Hesselmann, V., Radermacher, I., et al. (1999). Functional anatomy of intrinsic alertness:

Evidence for a fronto-parietal-thalamic-brainstem network in the right hemisphere. *Neuropsychologia, 37*, 797-805.

Sturm, W., Longoni, F., Fimm, B., Dietrich, T., Weis, S., Kemna, S., et al. (2004). Network for auditory intrinsic alertness: A PET study. *Neuropsychologia, 42*, 563-568.

Sturm, W., & Willmes, K. (2001). On the functional neuroanatomy of intrinsic and phasic alertness. *NeuroImage, 14*, S76-S84.

Summerfield, A. Q., & Akeroyd, M. A. (1998). Computational approaches to modelling auditory selective attention: monaural and binaural processes. In S. Greenberg & M. Slaney (Eds.), *Course Reader for the NATO Advanced Study Institute on Computational Hearing*. Berkeley: International Computer Science Institute.

Swainson, R., Cunnington, R., Jackson, G. M., Rorden, C., Peters, A. M., Morris, P. G., et al. (2003). Cognitive control mechanisms revealed by ERP and fMRI: Evidence from repeated task-switching. *Journal of Cognitive Neuroscience, 15*(6), 785-799.

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

Tamm, L., Menon, V., Johnston, C. K., Hessl, D. R., & Reiss, A. L. (2002). fMRI study of cognitive interference processing in females with fragile X syndrome. *Journal of Cognitive Neuroscience, 14*(2), 160-171.

Taylor, S. F., Kornblum, S., Lauber, E. J., Minoshima, S., & Koeppe, R. A. (1997). Isolation of specific interference processing in the Stroop task: PET activation studies. *NeuroImage, 6*, 81-92.

Taylor, S. F., Kornblum, S., Minoshima, S., Oliver, L. M., & Koeppe, R. A. (1994). Changes in medial cortical blood flow with a stimulus-response compatibility task. *Neuropsychologia, 32*(2), 249-255.

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

Tipper, S. P., Lortie, C., & Baylis, G. C. (1992). Selective reaching: Evidence for action-centered attention. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 891-905.

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*, 1409-1422.

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

Treisman, A. (1969). Strategies and models of selective attention. *Psychological Review, 76*, 282-299.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*(1), 97-136.

Tsal, Y., & Lamy, D. (2000). Attending to an object's color entails attending to its location: Support for location-special views of visual attention. *Perception & Psychophysics, 62*(5), 960-968.

Tsal, Y., & Lavie, N. (1993). Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 131-139.

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage, 14*, 1387-1401.

Vallar, G. (1998). Spatial hemineglect in humans. *Trends in Cognitive Sciences, 2*(3), 87-97.

Van Valkenburg, D., & Kubovy, M. (2003). In defense of the theory of indispensable attributes. *Cognition, 87*, 225-233.

van Veen, V., Cohen, J. D., Botvinick, M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage, 14*, 1302-1308.

Videbech, P., Ravnkilde, B., Gammelgaard, L., Engander, A., Clemmensen, K., Rasmussen, N. A., et al. (2004). The Danish PET/depression project: Performance on Stroop's test linked to white lesions in the brain. *Psychiatry Research: Neuroimaging, 130*, 117-130.

Wang, K., Fan, J., Dong, Y., Wang, C.-q., Lee, T. M. C., & Posner, M. I. (2005). Selective impairment of attentional networks of orienting and executive control in schizophrenia. *Schizophrenia Research, 78*(2-3), 235-241.

Ward, L. M. (1994). Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. *Canadian Journal of Experimental Psychology, 48*(2), 242-259.

Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception & Psychophysics, 62*(6), 1258-1264.

Weis, S., Fimm, B., Longoni, F., Dietrich, T., Zahn, R., Herzog, H., et al. (2000). The functional anatomy of intrinsic and phasic alertness - a PET study with auditory stimulation. *NeuroImage, 5*(Supplement 1), S10.

Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and Performance VIII* (pp. 239-257). Hillsdale, NJ: Erlbaum.

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 Science, USA, 90*, 8722-8726.

Woodruff, P. W. R., 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, D. L., & Alain, C. (1993). Feature processing during high-rate auditory selective attention. *Perception & Psychophysics, 53*(4), 391-402.

Woods, D. L., Alho, K., & Algazi, A. (1994). Stages of auditory feature conjunction: An event-related brain potential study. *Journal of Experimental Psychology: Human Perception and Performance, 20*(1), 81-94.

Wright, B. C., & Wanley, A. (2003). Adults' versus children's performance on the Stroop task: Interference and facilitation. *British Journal of Psychology, 94*, 475-485.

Wright, R. D., Richard, C. M., & McDonald, J. J. (1995). Neutral location cues and cost/benefit analysis of visual attention shifts. *Canadian Journal of Experimental Psychology, 49*(4), 540-548.

Zatorre, R. J., Mondor, T. A., & Evans, A. C. (1999). Auditory attention to space and frequency activates similar cerebral systems. *NeuroImage, 10*, 544-554.

Zysset, S., Müller, K., Lohrmann, G., & von Cramon, D. Y. (2001). Color-word matching Stroop task: Separating interference and response conflict. *NeuroImage, 13*, 29-36.