**EFFECTS OF PERCEPTUAL CONTEXT ON
EVENT-RELATED BRAIN POTENTIALS DURING
AUDITORY SPATIAL ATTENTION**

by

Stephen Richard Arnott

**A thesis submitted in conformity with the requirements
for the degree of Master's of Arts
Graduate Department of Psychology
University of Toronto**

© Copyright by Stephen Richard Arnott (2000)

**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-50465-4

Arnott, Stephen Richard

Department of Psychology

University of Toronto

Effects of perceptual context on event-related brain potentials during auditory spatial attention.

Master's Degree, 2000.

Performance on a selective attention task was measured in different conditions promoting the grouping of auditory locations. Short and long noise bursts were randomly presented at three azimuth locations. Participants were required to attend to one location and respond to infrequent, longer duration sounds occurring there. The spatial separation between task-relevant and task-irrelevant stimuli was varied between blocks of trials. In the Evenly-Spaced (ES) block, middle and far distractors were 60° and 120° respectively from the attended location. In the distractor clustering blocks, either the “farthest” distractor was moved 30° closer to the middle distractor, or vice versa. In the attended clustering blocks, the attended location was moved 30° closer to the middle tone, or vice versa. Relative to ES, event-related potential indexes of attention were larger when the distractors were spatially clustered together. Results are interpreted in terms of an object-based hypothesis of auditory attention.

Acknowledgements

I would like to thank my thesis advisor, Dr. Claude Alain, for his help throughout the preparation of this thesis as well as Dr. Bruce Schneider for his suggestions in the design of the experiment and preparation of the written work. Constantine Chernenko created the program used to analyze the behavioural data. This research was supported by a University of Toronto Open Fellowship and a Catherall Fund grant from the Baycrest Centre for Geriatric Care in Toronto.

Table of Contents

	Page
List of Tables	v
List of Figures	vi
List of Appendices	viii
Introduction	
• Early Behavioural Auditory Selective Attention Research	1
• Object-Based Attention	3
• Event-Related Potentials and Selective Attention	6
• Stimulus-Set Account.....	9
• Attentional-Trace Account	10
• ERPs and Auditory Selective Attention to Spatial Locations	12
• Mismatch Negativity	13
Method	
• Participants.....	15
• Stimuli and Task	15
• EEG Recording	18
• Data Analysis	19
Results	20
Discussion	29
References	42

List of Tables

Table 1. Mean performance for each condition collapsed over attended location

Table 2. Mean RT for targets in each condition according to the preceding sound location

Table 3. Mean FC1 and FC2 voltage (125-165 ms) for each condition collapsed over attended location

Table 4. Mean Pz voltage (350-550 ms) for each condition collapsed over attended location

List of Figures

Figure 1. Schemata for location clustering paradigm. Participants completed five conditions; one Evenly Spaced (ES) condition where distracting sounds could occur 60° or 120° from the attended location, and four clustering conditions. In the far-to-middle (FM) condition, the far distractors were moved 30° closer to the middle distractor; in the middle-to-far (MF) condition, the middle distractors were moved 30° closer to the farther distractors; in the attended-to-middle (AM) condition, the attended location was moved 30° closer to the middle distractors; and in the middle-to-attended (MA) condition, the middle distractors were moved 30° closer to the attended location. All five conditions were run once when the attended stream was on the extreme left (denoted by L) and again when the attended stream was on the extreme right (R). The middle sounds were always task-irrelevant.

Figure 2. Forty-seven channel electrode montage.

Figure 3. Grand average ERPs for standard sounds (dashed line) and deviant/target sounds (solid line) at the attended location collapsed over condition and side of presentation. The P1, N1, P2, and P3 deflections of the ERPs are indicated. A box above the scale indicates onset and duration of stimuli (standard and target). In this and subsequent figures, ERP polarity is plotted with negativity up.

Figure 4. ERPs recorded over the left (F3) and right (F4) frontal region as a function of the standard sounds (collapsed over attention).

Figure 5. a) P1, b) N1, and c) P2 voltage distributions to standard sounds (collapsed over attention) according to location of sound presentation. Each contour line reflects a change of 0.1 μ V.

Figure 6. Comparison of Nd waves for standard sounds. Two rows of electrode sites are displayed, FC1 and FC2. Five conditions are presented: NdES (black), NdAM (green), NdMA (blue), NdFM (orange), and NdMF (red). Scalp Nd_e voltage distributions are displayed across the bottom for each condition (ES, AM, MA, FM, and MF from left to right). Each contour line reflects a change of 0.1 μ V.

Figure 7. ERPs elicited by attended (solid line) and unattended (dashed line) standard sounds. Two electrode sites are displayed, FC1 and FC2. ERPs for the different conditions are differentiated by colour: ES (black), AM (green), MA (blue), FM (orange), and MF (red). Solid lines represent ERPs to standard sounds at the attended stream.

Figure 8. ERPs to standard sounds as a function of condition. ERPs for each condition are differentiated by colour: ES (black), AM (green), MA (blue), FM (orange), and MF (red). Sound location is represented by line thickness: attended (thick), middle (medium), and far (thin).

Figure 9. ERPs to deviant sounds presented at the attended (thick line), middle irrelevant (medium line), and far irrelevant (thin line) locations at the midline parietal electrode (Pz) for each of the five conditions.

Figure 10. MMN responses elicited by deviant sounds presented at the attended location (thick line), middle irrelevant (medium line), and far irrelevant (thin line). Electrode sites Fz, FC1, FC2, and Cz are shown for each of the five conditions.

Figure 11. Mean MMN voltage between 150 and 250 ms (average of Fz, FC1, FC2, and Cz electrodes) as a function of distance from the attended source (attended location, 30°, 60°, 90°, or 120°).

List of Appendices

Appendix A. Instructions to participants

Appendix B. Calculation of the Nd waves

**Effects of perceptual context on event-related brain potentials
during auditory spatial attention**

Auditory selective attention can be considered as the influence of behavioural goals on the processing priority of different aspects of auditory input. Traditionally, selective attention has been studied by requiring participants to attend to a subset of stimuli among other task irrelevant stimuli. To successfully accomplish this, a listener may use information provided by the physical features of the target stimuli (e.g., frequency and/or location of the sound source). One might hypothesize however, that attentional performance could be partially facilitated by organizing or grouping those attended and unattended features in some pre-attentive fashion. The aim of the present study was to test this hypothesis by manipulating parameters that promoted perceptual grouping. Although the notion of 'grouping' elements into objects is not novel, an understanding of the role of perceptual grouping during auditory selective attention remains elusive.

Early Behavioural Auditory Selective Attention Research

In early and now classical studies of auditory selective attention, individuals were asked to shadow one of two messages presented separately to each ear over headphones (Cherry, 1953). Although participants found this easy to do, they were often unable to report the contents of the other ear following the termination of the shadowing. Moreover, when told that the message in the unattended ear had switched from English to German, or to backward-playing English, participants reported being unaware that the change had occurred. Subsequent experiments by Moray (1959) demonstrated that words from word lists repeated up to 35 times in the unattended ear had a likelihood of being

recognized that was no greater than chance. Although individuals had difficulty remembering stimuli presented at the so-called 'unattended' ear, they were usually aware if the unattended ear's message switched gender, or pitch, or was replaced with a 400 Hz tone (Cherry, 1953). As well, listeners were able to respond to the occurrence of their own name (Moray, 1959), or the meaning of a significant item (Treisman, 1960) presented at the unattended ear.

These results lead to the formation of various models of selective attention. Broadbent, for example, proposed a filter model (Broadbent 1958, see also 1970, 1982). Within this model, all stimuli are stored in a short-term buffer and analyzed (preattentively) in parallel for relevant features. A subset of those stimuli are then transferred through a 'channel' of limited capacity allowing only stimulus with relevant features (e.g., location in space, or pitch) to pass on to the second stage of higher level processing. Treisman (1960, 1964) re-interpreted Broadbent's filter-model to include three stages: 1) an automatic discrimination of 'channels' based on general physical characteristics; 2) selective discrimination or attenuation of irrelevant channels by a limited-capacity filter; and 3) the identification of words and semantics carried out only for the 'selected' signals.

Other attempts to explain performance during auditory selective attention tasks suggested that the filtering occurred at a later stage of processing. Deutsch and Deutsch (1963) posited that all information was fully analyzed and that the limits of attention applied only to conscious awareness, memory and response. Stimuli are automatically processed at the semantic level, regardless of whether attention was being paid to stimuli or not (e.g., even if there were strong physical differences between stimuli at both

attended and unattended 'channels'). Because the human system was thought to be limited in the ability to organize responses to all sensory input, selective attention was utilized only to control the access of stimuli to consciousness, memory, and response using a system of assigned weightings of importance.

Object-Based Attention

Early versus late selection aside¹, it may be asked what effect distracting information has on performance during selective attention. Imagine a person selectively attending to a particular conversation in the midst of others at a party. In order to follow this conversation, the listener may use information provided by the physical features of the stimuli such as the frequency and/or location of the speaker's voice. The ability to monitor the relevant message would likely be made worse by increasing the physical similarity between the distracting and the relevant messages. Distractors physically similar to a target produce interference because they also pass through the attentional 'filter', and are therefore selected for further processing and a potential response.

However, information provided by the perceptual context can also be used during selective listening. Perceptual context can be derived from top-down and bottom-up processes. For example, semantic characteristics (e.g., priming effects or word expectancy) can play a role in forming perceptual context. Individuals are more likely to correctly identify a sentence-final word embedded in noise if the target word is predictable from the context of the sentence (Pichora-Fuller, Schneider, and Daneman, 1995). The temporal relationship of sounds also contributes to the perceptual context. Bregman (1990) provides the example of an alternating pair of tones, one of high (H) and

¹ Indeed, Lavie (1995) argues that differences in 'perceptual load' can resolve the early vs. late selection issue.

one of low (L) frequency. At slow presentation speeds, one hears a succession of tones that alternate regularly in pitch, at intermediate speeds one is given the impression of melodic jumps between the H and L, and at high speeds the sequence of Low and High tones split into two separate streams or objects.

An auditory object refers to the percept of a group of sounds as a coherent whole, seeming to emanate from a single source (e.g., location or area). Sounds may be grouped into objects according to various principles. Gestalt psychology offers several ideas some of which include: similarity (elements that are similar in pitch, intensity, and spatial location tend to be grouped together); proximity (elements occurring close together in time tend to be grouped together); common fate (elements that move together tend to be grouped together); and symmetry and closure (elements that form symmetrical and enclosed objects tend to be grouped together) (Handel, 1987; Bregman, 1990).

Few studies have examined the effect of perceptual context during selective listening. In an early study, Treisman (1964) provided some of the first evidence for perceptual grouping when she noted that participants' shadowing of monaural prose slightly improved as the locations of distracting prose were moved closer together (by varying intensity differences between the ears). This improvement in performance may be attributable, in part, to the perceptual grouping of distractors. However, the interpretation is confounded by the fact that as the distractors were moved closer together, they were also being moved farther away from the attended location. This illustrates a significant problem interpreting results from most studies in terms of an object-based or a feature-based account; that is, an auditory object may also be defined by its features. This confound is most readily dealt with in studies that contrast the effects of grouping with

the effects of physical similarity.

Bregman and Rudnicki (1975) had participants judge the order of two target tones (T_1 and T_2 at 2200 Hz and 2400 Hz, respectively). The judgment was made more difficult by flanking the tones with two distractor tones (D) (at 1460 Hz each). Bregman and Rudnicki then added captor tones (C) that flanked the distractor and target sequence (e.g., CCCDT₁T₂DCCC). When the captors were of a lower frequency than the distractors, target order judgment remained difficult. However, raising the frequency of the captor tones to values more similar to the distractor frequency, improved target judgment performance. These results are consistent with an object-based account predicting that increasing the similarity between the distractor and captor frequency allowed for those elements to be grouped together as an object and thus more readily filtered out (Alain and Arnott, 2000).

Alain and Woods (1993, 1994) also showed that the information provided by frequency clustering could override physical similarity as the major influence of selective processing mechanisms. In their 1994 study, participants were required to listen to binaural tones of low, middle, or high frequency randomly played over headphones. In separate blocks of trials, participants were asked to attend to low or high frequency sounds and respond whenever they heard a tone of longer duration (i.e., 90 ms instead of 50 ms) at that attended frequency. Alain and Woods used three conditions that they termed: evenly-spaced (ES), clustered-easy (CE), and clustered-hard (CH). In the ES condition, the three frequencies (two distractors and one attended) were equidistant on the musical scale. In the other conditions, the attended frequency was either distinct from the clustered distractors (CE) or clustered itself with one of the distractors (CH).

Relative to the ES condition, they found that participants performed better (i.e., faster RTs and fewer errors) in the CE condition but worse in the CH condition. Like their 1993 study, these results are in line with the hypothesis that selective attention can operate on objects (i.e., sound features grouped into a perceptual stream) rather than on simple features alone. Importantly, their paradigm differed from their previous study in that brain activity was recorded concurrently with behavioural performance. Event-related brain recordings have the distinct advantage over behavioural studies of being able to examine so called 'unattended' information. As well, they allow one to dissociate the effect of stimulus selection and response processes (see Woods, 1990).

Event-Related Potentials and Selective Attention

Scalp recorded event-related potential (ERP) waveforms reflect synchronized postsynaptic activity in relatively large populations of neurons time-locked to a sensory or cognitive event (Picton, 1990). ERPs are extracted from background noise by averaging several brain responses elicited by a particular stimulus. Averaging is based on three assumptions (Glaser and Ruchkin, 1976): 1) that the recorded waveform is the linear sum of the signal and noise; 2) that the signal waveform is the same for each event repetition; and 3) that the noise waveforms are sufficiently irregular from event-to-event that they can be considered to be statistically independent samples of a random process. In most cases, averaging will reduce noise to a residual that is inversely proportional to the square-root of the number of repetitions (Picton, Lins, and Scherg, 1995).

Averaged ERPs to auditory stimuli follow a typical pattern of positive and negative voltage peaks following sound onset. Because it is not unreasonable to assume that different information processing mechanisms occur in parallel, any wave deflection

may therefore be reflecting more than one underlying component (Coles and Rugg, 1995). Components may be formally defined as the contribution to the recorded waveform of a particular generator process (Näätänen and Picton, 1987). The identification of these components is a current concern in the field and has been termed “the Inverse Problem”, as there exists no unique solution to the possible underlying component(s). Attempts to estimate these locations include source analysis which is based on properties of electrical activity and brain physiology (c.f. Scherg 1990; Picton et al., 1999).

Auditory evoked potentials can be subdivided into three groups on the basis of their latency: early auditory brainstem responses (ABRs) occurring from stimulus onset to 10 ms post-stimulus; middle latency responses (MLRs) occurring between 10 and 50 ms post-stimulus; and longer latency evoked potentials arising 50 ms after the stimulus onset (Hillyard and Kutas, 1983). ABRs reflect cochlear and brainstem responses. MLRs reflect the earliest responses from the auditory cortex (Vaughan and Arezzo, 1988). Longer latency deflections are larger than ABRs or MLRs (approximately 5-10 times the size of ABRs) and reflect activity from a distributed neural network. The longer latency deflections are labeled according to polarity and time sequence (e.g., P1, N1, P2, and N2). For example, N1 refers to the first large negative wave following sound onset, while P2 refers to the second occurring positive deflection. The ABR has been found to be insensitive to attentional manipulation (Picton and Hillyard, 1974, Picton Stapells, and Campbell, 1981). Although there is some evidence that MLRs are sensitive to attentional manipulation, especially when the task requires focussed attention (Linden, Picton, Hamel & Campbell, 1987; Woldorff and Hillyard, 1991), reliable selective

attention effects on auditory evoked potentials do not occur until 50 ms post-stimulus onset (see Woods, 1990).

The first study to unequivocally demonstrate the effect of selective attention on the (late) auditory evoked potential was carried out by Hillyard, Hink, Schwent, and Picton (1973). In their paradigm, participants were required to attend to tones of one frequency in one ear and count the number of frequency deviants occurring there, while ignoring tones and deviants of another frequency in the other ear. Importantly, the stimuli were delivered at a rate (2-4 per second) fast enough to require selective attention to one ear, and presented randomly to the left and right ear so that no anticipatory response preparations could be made. Because the stimulus set was always the same (only the direction of attention was changed), the experimenters could infer effects of selective attention by comparing ERP responses from a sound when it was attended to the same sound when it was being 'ignored'. Hillyard and colleagues (1973) found enhanced N1 amplitude over the frontocentral scalp sites for sounds when they were attended versus unattended. As well, deviant sounds in the attended stream were found to elicit a P3 wave maximal over the parietal region that was not present for deviant sounds in the unattended ear.

Since this seminal study, subsequent studies have often followed a similar design, recording ERPs to both attended and unattended stimuli in the same block. To reveal the physiological effects of attention, it has become standard practice to examine 'difference waves'. These are obtained by subtracting 'ignored stimulus' ERPs from ERPs to the same stimulus when it is being attended to. The resulting slow-negative difference wave (Nd) is characterized by two negative, occasionally overlapping, phases: an early (Nd₁)

and a late (Nd_l) component (see Hansen and Hillyard, 1980).

The Nd_e can begin as early as 50 ms after tone onset, peaking around 200 ms. It is present over various signal intensities as well as during attention to sounds in the presence of multiple distractors. The Nd_e peaks earlier and is larger in amplitude when the stimuli are presented at rapid rates (Parasuraman, 1980; Schwent, Hillyard, and Galambos, 1976a). Similarly, greater amplitude is found when stimuli are faint and difficult to detect (Schwent, Hillyard, and Galambos, 1976b). The Nd_e seems to be more closely related to discriminability of attended and unattended sequences rather than to the discriminability of standard and target stimuli within the attended sequence (Parasuraman, 1980). The Nd_e has been shown to increase in amplitude when the frequency (Alho, Sams, Paavilainen, and Näätänen, 1986) or location (Alho, Donauer, Paavilainen, Reinkainen, Sams, and Näätänen, 1987) differences between the task-relevant and task-irrelevant stimuli increase. The Nd_l is not as well understood as the Nd_e , however it has been associated with a 'further' processing of stimuli that have already been selected (Schröger and Eimer, 1996).

Stimulus-Set Account

Two models have been proposed to explain the attention effect on ERPs. Hillyard's Stimulus Set account posits that the Nd_e reflects the enhanced sensory processing of attended stimuli (Hillyard et al., 1973). Thus, enlargement of the N1 for both the standard and deviant tones at the attended locations relative to the ERPs to the same sounds when they were unattended was taken to indicate facilitated sensory transmission. Furthermore, the lack of a $P3b^2$ component to the deviant sounds occurring

² Note: $P3b$, as opposed to $P3a$, refers to the classic $P3/P300$ component maximal over the parietal region. It is distinct from the more frontal $P3a$ component (Knight, Scabini, Woods, and Clayworth, 1989).

at unattended locations was interpreted as inhibition of non-relevant stimuli. Within the Stimulus Set account the $N1/Nd_e$ indexes a 'gating' function on perceived stimuli. Stimuli are accepted or rejected based on their dominant relevant physical attributes and only those 'accepted' stimuli receive further perceptual analysis. According to Hillyard, the Nd_e reflected a late stage of selection of targets from non-targets within the attended channel based on less discriminable cues.

Hansen and Hillyard (1983) challenged their Stimulus Set account using four auditory 'channels' defined by pitch (600 or 1500 Hz) and location (left or right ear): i) high pitch/left ear, ii) high pitch/right ear, iii) low pitch/left ear, and iv) low pitch/right ear. In some blocks, the difficulty of the pitch discriminations or the location discriminations was increased. Participants were required to attend to a certain pitch/location combination and respond to deviant longer duration tones. The Nd_e 's were observed for stimuli that shared one of the relevant features and their summed amplitudes roughly approximated the size of the Nd_e to the feature combination. This finding suggests that the Nd_e indexes a self-terminating process in which stimulus features are processed in parallel. Irrelevant stimuli are rejected rapidly after an analysis of their salient features.

Attentional-Trace Account

Näätänen (1992) proposed a slightly different and more specific perspective of the processes underlying the Nd_e wave. According to Näätänen, the degree of $N1/Nd_e$ negativity is related not to a true enhancement of the sensory response per se, but rather to an endogenous processing negativity (PN) that overlaps with the $N1$ response. The PN indexes a comparison process between the incoming stimulus and a voluntarily

maintained representation (template) of the physical features of the task-relevant stimuli. The more similar a stimulus is to this template, the longer the duration of the PN. Thus, for an irrelevant stimulus, the matching process terminates earlier for more conspicuous differences from the relevant stimuli. Within this 'Attentional-Trace' model, the Nd_t reflects either a late selection of the target within the attended channel or a rehearsal of the attentional trace.

Of relevance to these and other potential models of selective attention however, is whether or not the criteria that is being used to perform the selection is purely feature-based or whether this can accommodate a more abstract filter or template; i.e., an object-based one. Alain and Woods (1994) examined the effect of selective attention on ERPs in situations that promoted the auditory stream segregation. As previously mentioned, participants were randomly presented with a rapid series of tones varying among three frequencies. Participants were required to attend to deviant tones occurring at a specified frequency. In the ES block, frequencies were evenly-spaced apart; in the CH blocks, the task-relevant frequency was changed to a pitch more similar to the middle distractor frequency (clustered hard); and in the CE (clustered easy), the extreme distractor was changed to a frequency more similar to the middle distractor frequency.

Relative to the ES condition, the Nd_t amplitude was enhanced when distractors were clustered (i.e., CE), and reduced when the task-relevant stream was clustered with the middle distractors (i.e., CH). The enhanced Nd amplitude reflected i) an increased negativity to attended tones, and ii) an enhanced positivity to unattended tones. These results suggest that the processing of stimuli belonging to the attended stream were promoted, while those not belonging to the attended stream were inhibited. In other

words, clustering distractors together whether nearer or further to the attended frequency, resulted in larger Nd_t amplitudes. The effect of frequency clustering on the Nd_t are not easily accounted for by the Attentional Trace model because the attention-related negativity to the middle tones differed between the ES and CE conditions despite constant frequency separation. The trace model would have predicted similar ERP amplitudes to the middle distractors. Also, the PN associated with the trace model assumes that the negativity is maximal and constant for the attended tone whereas it varies at the unattended stimuli according to the similarity between the attended and unattended stimuli. However, in the CE condition, the ERP to attended stimuli showed a negative enhancement compared with the ERP to attended stimuli in the ES condition. The results appear more compatible with an object-based account wherein attended stimuli are preattentively grouped together into one perceptual object separated from the background distractors.

ERPs and Auditory Selective Attention to Spatial Locations

To this point, the object-based hypothesis has been discussed solely in the frequency realm. If this grouping principle is to be generalized, one might expect it to apply to features other than frequency. Over the last decade, there has been a resurgence of interest in studying the mechanisms of auditory spatial attention (Mondor and Zatorre, 1995; Teder, Kujala and Näätänen, 1993; Teder and Näätänen, 1994; Teder-Sälejärvi, Hillyard, Röder, and Neville, 1999a; Yost, 1997). This is likely the result of an increased awareness of the role that spatial cues play in selective attention, especially when more than two sound sources are simultaneously active (Yost, Dye, and Sheft, 1996).

Schwent and Hillyard (1975) were among the first to examine the effect of

selective attention on ERPs using multiple sound sources. The N1 effects were found to be greater than those found in Hillyard et al.'s (1973) two-channel paradigm prompting the suggestion that the task-relevant location is processed more selectively under conditions of high information load. In a different experiment, they also noted an N1 effect for distant distractors presented at an adjacent position to the attended location.

To examine the spatial gradient of auditory attention, Teder-Sälejärvi and Hillyard (1998) had participants seated in front of an array of seven speakers spaced at 9° intervals on a horizontal axis. Standard (500-5000 Hz bandwidth) and deviant (500-1500 Hz bandwidth; 0.13 probability at each speaker) noise bursts were played randomly from each speaker with an ISI varying from 8-188 ms. Participants were required to attend to a speaker (e.g., the speaker to the extreme left) and press a button in response to deviant sounds (targets) at that location, ignoring stimuli from the other six locations. The investigators found a narrow gradient of auditory attention surrounding the attended location. The attention-related negativity decreased progressively as one examined ERPs to sounds presented at locations farther away from the attended location. This finding is in accordance with several other behavioural and electrophysiological studies (Mondor and Zatorre, 1995; Spence, Nichols, Gillespie, and Driver, 1997; Teder-Sälejärvi, et al., 1999a).

Mismatch negativity

A secondary goal of the present study was to evaluate the effect of selective attention on the processing of task-irrelevant material. When a 'deviant' stimulus occurs during a repetitive sequence of homogenous stimuli, a broad mismatch negative (MMN) wave is elicited in the ERP, superimposing the N1 and P2 deflections. The MMN is

maximal over the frontocentral scalp region and inverts in polarity over the mastoid sites (see Näätänen, 1992 for a review). The MMN is elicited by changes in pitch, intensity, duration, location, timing, and phonetic content. Because it is obtained when attention is directed away from the auditory stimuli, (e.g., having the participant read a book), it is thought to reflect the automatic analysis of stimulus features. Recent studies however, have shown that auditory attention can modulate the amplitude of the MMN (Woldorff, Hackley, and Hillyard, 1991; Woldorff, Hillyard, Gallen, Hampson, and Bloom, 1998). In the grouping study of Alain and Woods (1994) the MMN was shown to be modulated by attention (great negativity for the attended frequency, moderate negativity for the middle distractor frequency, and least negativity for the extreme distracting frequency), but not by the clustering manipulation itself.

The present study

The current study was a selective listening task where participants were required to attend to one sound location and ignore distracting sounds at a moderate and a far distance from the attended location. In different blocks, the spatial separation between the attended and distracting streams was manipulated. Based on previous object-based accounts (i.e., Alain and Arnott, 2000; Alain and Woods, 1993; 1994), it was expected that clustering of spatial location would influence performance (RT and accuracy) and ERP attention effects. Specifically, the clustering of the middle and far distractor sounds was predicted to improve performance (reduce RT, increase accuracy etc.) and generate a larger Nd amplitude relative to the ES condition. Alternatively, according to a feature-based account (e.g., Hillyard et al., 1973; Näätänen, 1992), moving extreme distractor sounds closer to middle sounds should impair performance because the distractor would

now be made much more similar to the task-irrelevant stimuli.

Method

Participants

Twelve right-handed young adults (5 males, aged 18-39 years) were recruited for this study. One participant was excluded due to excessive muscle artifacts. Volunteers were remunerated for their participation and all provided their informed consent according to University of Toronto and the Baycrest Centre for Geriatric Care guidelines. All participants reported normal hearing.

Stimuli and Task

The 3D auditory environment was created using Tucker-Davis Technology hardware. The auditory stimuli, were comprised of two types of broadband gaussian noise bursts (RMS intensity of 75 dB SPL and a peak intensity of 90 dB SPL; 5 ms rise/fall times). Binaural stimuli were presented over Sennheiser HD 265 Linear headphones at a variable interstimulus interval (ISI) of 100-300 ms. Deviant probability was 22% overall with the constraint that at least one standard sound precede the deviant stimuli. On any given trial, stimuli randomly emanated from one of three possible locations.

Realistic emulation of sound localization makes use of interaural time differences (ITD; the lag in time for the same sound source to reach the left versus the right ear) in addition to interaural intensity differences (IID; the amount of sound energy arriving at each ear from the same source varies with the source's location). Free-field sound presentation over loudspeakers achieves this realism, however there are instances when this type of presentation is either not feasible (e.g., when anechoic chambers are not available) or, as in the case of brain-imaging studies, not practical.

In the current study, sounds were presented at virtual locations using head-related transfer functions (HRTFs). HRTFs are mathematical formulae (three space coordinate variables and one frequency variable) that replicate pinnae and interaural time and frequency differences based on actual recordings made within the artificial (McKinley and Ericson, 1988) or the real (Middlebrooks, Makous, and Green, 1989) human inner ear. Studies requiring listeners to discriminate between ‘real’ speaker sounds versus ‘virtual’ HRTF headphone sounds demonstrate chance performance, even when the magnitude spectrums of the HRTF files are smoothed (Kulkarni and Colburn, 1998). However, the HRTFs in these studies are based on actual head and ear measurements from the participants. Ideally, HRTFs should be tailored to each listener because of population variances in interaural distance and pinnae size and shape. However, this is more of a problem when front/back or elevation discriminations have to be made as these cues rely more heavily on the pinnae (Han, 1994; Handel, 1987). With these guidelines in mind, HRTF files can be used to create a realistic yet virtual, externalized auditory environment over headphones.

In the present study, virtual locations varied along the azimuth only (see Figure 1). In the baseline (ES) condition, sound locations were evenly spaced along the azimuth, originating from -60° (to the left of midline), 0° (straight-ahead), and 60° (to the right of midline) locations. Participants were asked to attend to one of the extreme locations and to ignore the stimuli presented at the other locations. Sounds at the middle location were always task-irrelevant. For example, in the ES attend left (ES_L) participants were required to attend (underlined font) to the -60° location, ignoring the sounds occurring at the 0° (Middle) and 60° (Far) locations. The other eight conditions (four types by two

sides) involved changing the separation between the source comprising the auditory scene. In the middle-to-far (MF) conditions, the middle distractor was moved closer to the attended location; in the far-to-middle (FM) conditions, the far stream was moved closer to the the middle stream; in the middle-to-attended (MA) conditions the middle distractor was moved closer to the attended location; and in the attended-to-middle (AM) conditions, the attended location was moved closer to the middle location.

Apparent spatial locations were: -60°, 30° and 60° for the MF_L condition; -60°, -30° and 60° for the MF_R condition; -60°, 0° and 30° for the FM_L condition; -30°, 0° and 60° for the FM_R; -60°, -30° and 60° in the MA_L condition; -60°, 30° and 60° in the MA_R condition; -30°, 0°, and 60° in the AM_L condition; and -60°, 0°, and 30° in the in the AM_R condition. Side of attended stream location (left or right) was counterbalanced within subjects.

Insert Figure 1 about here

The HRTF file used for each participant was chosen among a set of HRTF files. Participants listened to a series of one second noise bursts that ‘rotated’ clockwise at 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315° azimuth locations (0° elevation). This rotation was repeated using six HRTF files (five previously obtained from ear measurements of five non-participating individuals, and one from the measurements of an average head model). The HRTF file that provided the listener with the most distinct, localizable sound sources around their head was used for the experiment.

Participants first read an instruction sheet (see Appendix A) requiring them to respond as quickly and as accurately as possible to the occurrence of a deviant tone at a specified location (i.e., a target tone). Standard and deviant tones occurring at the other

two locations were therefore task-irrelevant and participant's were told that they should ignore them. Participants then completed a practice block of 200 trials that included real-time onscreen feedback of response times, false alarms and correct hits. The experiment proper consisted of ten blocks of 1100 trials (no performance feedback provided). The first 100 trials of each block were considered practice and were not included in the data analysis. Responses were made by depressing a button with the right index finger. Similar to Alain and Woods (1994) and Teder-Sälejärvi and Hillyard (1998), correct responses were defined as those trials with RTs occurring between 200 and 800 ms after target presentation. Responses at other times were classified as false alarms (FA). Misses were defined as no responses made to targets within 1200 ms after presentation.

EEG Recording

Electrophysiological signals were recorded continuously with a sampling rate of 250 Hz from an Electro-Cap with 35 tin electrodes adhering to the ten-twenty system (American Electrophysiological Society, 1991) (see Figure 2). Vertical and horizontal eye movements were monitored with electrodes placed at the outer canthi and superior and inferior orbits. Six additional off-cap electrodes were placed at the nasion, neck, left and right mastoids, and over the left and right sphenoids. Electrophysiological signals were amplified (gain of 2500) and filtered (bandpass 0.5-30 Hz) via Syn Amps. During acquisition, all electrodes were referenced to Cz. After acquisition, all data were re-referenced a common reference (i.e., the reference voltage at any point of an electrode's recorded epoch was made equal to that of the mean voltage of all scalp electrode voltages at that same moment in time) and the Cz electrode was then reinstated.

Averaging occurred offline. The ERPs were averaged separately for each

stimulus type, and for each condition according to each electrode. The epoch extended over -100 ms (pre-stimulus) to 800 ms (post-stimulus). ERPs to targets included only the individual trials that were followed by a correct response (within 200-800 ms following target presentation). Trials contaminated by excessive peak-to-peak deflection (i.e., greater than 100 μ V or less than -100 μ V) at non-eye electrode sites were excluded from the ERP average. The proportion of rejected trials was less than 0.05.

Insert Figure 2 about here

In each individual average, the ocular (electro-oculogram (EOG)) artefacts (e.g., blinks and lateral eye movements) not removed by the artefact rejection criteria were corrected using ocular source components (Berg and Scherg, 1994; Lins, Picton, Berg, and Scherg, 1993a, 1993b). Using the same montage, a separate EOG calibration recording was made. During this calibration, participants made saccades to a fixation point that alternated between the centre and the edge of the screen, separately in the up, down, right and left directions. As well, intermittent blinking was recorded. This allowed for the creation of an ocular data set by concatenating average recordings of each of the saccades and the blinks. A principal component analysis of this data set provided a set of components that represented the variance related to the eye movements. Those components that explained more than 1% of the variance and were specifically related to the EOG waveforms were used as source components to subtract EOG contamination from the average ERPs (see Lins et. al., 1993a; 1993b for a more in-depth discussion of EOG correction).

Data Analysis

The experiment was a five-factor within subjects design with repeated measures

of Stimulus Type (standard or deviant), Condition type (ES, MF, FM, MA, and AM), Hemispace (attend to the left auditory field or attend to the right auditory field), Location (attended, middle, and far locations), and Laterality (hemisphere contralateral or ipsilateral to the attended sound location). Reaction time and accuracy measures were analyzed by Stimulus Type, Condition, and Hemispace as were ERP component (peak amplitude and latency) measures with electrode site as an additional factor.

Following ERP averaging, Nd waves were calculated and measured. It is most common to examine Nd attention effects on frequent standard sounds because, unlike deviants, standard sounds usually possess 1) a greater number of trials (thus a better signal-to-noise ratio), as well as 2) less contamination by response-related potentials such as motor components. Peak amplitudes and latencies of Nd_x waves were quantified. Greenhouse-Geisser corrections for multiple comparisons and violations of sphericity were employed where appropriate using Epsilon (ϵ). Significant values were based on p values of 0.05 or less.

Results

Behavioural Data

Group mean behavioural data are summarized in Table 1. The task was relatively easy. The participants responded quickly (~475 ms) and accurately, correctly responding to approximately 90% of targets. Misses and RTs were not affected by the clustering manipulations, although participants tended to be slower when the middle distractors were clustered with the attended location. There was a main effect of Condition (ES, MA, AM, FM, and MF) on the FA rate, [$F(4,40)=11.34$, $p<0.001$, $\epsilon=0.556$]. Pairwise comparisons revealed that participants made significantly more FAs when the middle

distractors were clustered with the attended location ($p < 0.001$, in all cases)³.

Insert Table 1 about here

The RT has been shown to be influenced by the features of distracting tones preceding the target (Woods, 1990). Moreover, this 'priming effect' is affected by frequency clustering manipulations (Alain and Woods, 1993). To assess the influence of spatial clustering on the priming effect, the effect of the location of the preceding sound type on RT was compared across conditions. Mean RTs for the three types of preceding sound locations (Attended, Middle, or Far) are displayed for each condition in Table 2. A main effect of Preceding Location [$F(2,20)=12.44$, $p < 0.01$, $\eta^2=0.693$] prompted pairwise comparisons between the three preceding locations. The RT for the location identical to the target location (463 ms) was faster than was the RT to the targets preceded by a middle (475 ms) or the far (480 ms) location ($p < 0.02$ and $p < 0.01$ respectively). Moreover, mean RTs for targets preceded by middle location stimuli were faster than those preceded by far location stimuli ($p < 0.05$). There was no interaction between Previous Location type and Condition type [$F(8,80)=1.86$] suggesting that the priming effect was not influenced by the clustering manipulations.

Insert Table 2 about here

Within each condition, the priming effect was examined. The main effect of location was primarily the result of the FM condition. In the ES, AM, and MF conditions, RTs did not differ from each other according to where the previous stimulus emanated from. In the FM condition, the RTs were fastest when the previous stimulus

³ Note: The finding of increased errors for the MA but not the AM condition is in agreement with recent findings demonstrating that the gradient of auditory spatial attention is less narrowly focussed for stimuli in the peripheral auditory field than for stimuli in the central auditory field (Teder-Sälejärvi and Hillyard, 1998).

originated from the task-relevant location, moderately fast when the previous stimulus was at the middle location ($p < 0.05$), and slowest when the previous stimulus was at the far location ($p < 0.005$). For the MA condition, RT was actually slowest when preceded by a middle sound ($p < 0.01$), but moderately slow when preceded by a far sound ($p < 0.005$). Middle and Far locations did not differ.

Electrophysiological Data

The ERP response to short frequent and long infrequent duration noise bursts were characterized by a positive-negative-positive deflection over the frontal scalp area that reversed in polarity at mastoid sites. At frontal sites, the P1 for standard sounds peaked at approximately 75 ms, the N1 at 145 ms, the P2 at 250 ms. The target stimuli elicited an additional positive deflection (P3b also referred to as P300) between 350 and 650 ms, peaking at 510 ms over the parietal region (see Figure 3). An ANOVA with Laterality (using FC1, F3, AF3 vs. FC2, F4, AF4) indicated that the P1 and N1 waves were larger over the hemisphere contralateral to the sound location [$F(4,172)=4.30$ and 27.61 respectively, $p < .01$, $\eta^2=0.815$] (see Figure 4). No significant Laterality effect was evident for the P2 deflection [$F(4,172)=0.81$]. Furthermore, the peak of the voltage distribution across the scalp shifted closer to midline as sound location was moved towards 0° (see Figure 5).

Insert Figure 3 about here

Insert Figure 4 about here

Insert Figure 5 about here

The interaction between Hemisphere (left auditory field vs. right auditory field) and Attention (attend vs. ignore) was not significant, nor was the interaction between

Hemisphere, Attention, and Condition. To reduce the number of displays, all electrodes during an attend right condition except for the midline sites of Fz, Cz, Pz, Oz, Iz, and Neck, were flipped (i.e., FC1 became FC2 and FC2 became FC1, T7 became T8 and T8 became T7, etc.). Unless otherwise stated, all subsequent figures collapse Hemisphere and can be considered to reflect brain responses to sounds when attention was directed to left auditory space.

Attention Effects

The selective attention effects on ERPs are most easily illustrated by subtracting ERPs elicited by task-irrelevant stimuli from the ERPs elicited by the same stimuli when they are task-relevant. Specific subtractions used to calculate the Nd waves in the present study can be found in Appendix B. The Nd waves were characterized by early (Nd_e) and late (Nd_l) components (see Figure 6).

Insert Figure 6 about here

Nd_e

The Nd_e peaked at approximately 150 ms and was largest over frontocentral scalp sites. Furthermore, it was larger over the hemisphere contralateral to the attended location (Attention by Laterality [$F(1,10)=64.44$, $p<0.001$]). The effect of attention was quantified by measuring the mean amplitude of ERPs to standard sounds from 120 to 180 ms at two frontocentral electrodes, FC1 and FC2.

A four-way ANOVA with Attention, Condition, Hemisphere, and Laterality (using FC1 and FC2) as factors yielded, a main effect of Attention [$F(1,10)=56.63$, $p<0.001$]. ERPs to sounds at the attended location were more negative than those elicited by the

same sounds when unattended (task-irrelevant)⁴. There was a main effect of Condition [$F(4,40)=4.47$, $p<0.02$, $\eta^2=0.668$] and a significant interaction between Attention and Condition [$F(4,40)=2.99$, $p<0.05$, $\eta^2=0.864$]. Planned comparisons revealed that the Nd_e was larger in the MF than the ES, AM and MA conditions ($p<0.02$) but not the FM condition. This lack of difference between the FM and MF condition could not be attributed to an increase in variance in one or the other condition as a pairwise comparison between the standard deviations of the two means was not significant ($p=0.2$). This is important with regard to the initial predictions. It implies that moving a far distractor closer to the attended location (FM) moderately increases the Nd_e amplitude.

To determine whether the enhanced Nd_e amplitude in the MF condition reflected increased processing negativity to attended stimuli or decreased processing negativity to task-irrelevant stimuli, the effect of clustering on ERP amplitude was examined separately for the attended and unattended stimuli (see Figure 7). An ANOVA on the ERP amplitude to attended stimuli revealed a smaller ERP amplitude (less negative) to attended stimuli when the middle and far distractors were clustered together, $NdMA$ ($p<0.025$) and $NdAM$ ($p<0.02$) relative to each of the other three conditions. An analysis of the task-irrelevant stimuli comparison of the unattended component comprising the $NdMF$ wave revealed that it was less negative than the unattended $NdES$ ($p<0.02$) and $NdFM$ ($p<0.02$) components. Thus, like Alain and Woods' (1994) results, it was determined that the Nd_e modulations were affected both at the attended location and at the unattended location. Specifically, a decreased negativity to attended $NDAM$ and $NDMA$

⁴ 'Unattended' in this context is loosely defined as stimuli or stimulus locations that are irrelevant to the task. It does not necessarily preclude the possibility that attention was being allocated to these stimuli.

sounds as well as an increased positivity to unattended NDMF sounds relative to the NDES condition

Insert Figure 7 about here

Nd_t

The Nd_t peaked at approximately 410 ms post-stimulus. A four-way ANOVA on the mean amplitude between 350 and 500 ms with Attention, Condition, Hemispace, and Laterality (using FC1 and FC2) yielded a main effect of Attention [$F(1,10)=30.19$, $p<0.001$]. ERPs to sounds occurring at the attended location were more negative than ERPs to task-irrelevant sounds presented at the same location. There was a marginal trend for a Condition effect [$F(4,40)=3.10$, $p=0.051$, $\eta^2=0.646$]. Unlike the Nd_e , there was no Attention by Condition [$F(4,40)=1.29$] interaction nor was there an Attention by Laterality interaction [$F(1,10)=1.77$].

$N1$ effects

The effect of spatial clustering was examined on the $N1$ amplitude (125-165 ms at FC1 and FC2 for standard sounds; see Table 3). Comparisons of attended stimuli located 60° from midline (i.e., ES, FM, MF, and MA conditions)⁵ yielded a marginal main effect of Condition [$F(3,30)=2.91$, $p=0.07$, $\eta^2=0.748$]. Pairwise comparisons revealed that sounds in the ES condition were more negative than those in the MA condition ($p<0.01$). For the distractor sounds at 0° midline (i.e., ES, AM, and FM conditions) all $N1$ s were similar in amplitude, with no significant effect of Condition [$F(2,20)=0.06$]. This did not change when all middle sounds (including MF and MA conditions) were compared. The

⁵ Because different sound locations evoke different scalp voltage topographies, inferences regarding the effects of conditional manipulation can only be made by comparing responses to sounds occurring in the same location under different conditions. Thus, only the attended locations of ES, FM, MF, and MA (60°) but not AM (30°), were compared.

N1 amplitude to far distractor sounds presented at -60° however yielded a main effect of Condition [$F(3,30)=4.06$ $p<.03$, $\eta^2=0.822$]. Pairwise comparisons revealed that the N1 mean amplitude was less negative in the far MA location than the N1 to the same sound in the ES or the AM condition ($p<0.01$ in both cases).

Insert Table 3 about here

The gradient of attention was examined according to the three locations in each condition. An ANOVA revealed a main effect of Location (attended, middle, and far) [$F(2,20)= 60.58$ $p<0.001$]. No significant main effect of Condition was found [$F(4,40)= 2.61$ $p=0.07$, $\eta^2=0.808$]. There was an interaction between Location and Condition [$F(8,80)= 3.87$ $p<0.01$, $\eta^2=0.565$]. The ERPs are illustrated in Figure 8 (refer also to Table 3). In the ES condition, mean N1 amplitudes for the three locations were all different from each other (all $p<0.005$). The N1 for the middle location stimuli was more negative than the far stimuli but more positive than the attended stimuli. Similarly for the AM condition, N1 amplitudes for the different locations all differed from each other (all $p<0.05$) with N1s for middle stimuli being more negative than in the far stimuli but more positive than for the attended. For the MA condition, all location comparisons differed (all $p<0.001$) except for the middle and far stimuli ($p=0.063$) which had a trend for the middle N1 to be more negative than the far. The attended N1 was more negative than either the middle or the far. Within each of the FM and MF conditions, all locations differed from one another (all $p<0.001$) except for the far and middle stimuli. The N1 for the attended stimuli were more negative than middle and far. Similarly, in the MF condition, all locations differed from one another (all $p<0.001$) except for the middle and far.

Insert Figure 8

P3b

The target stimuli generated a P3b response that was maximum over the midline parietal site. The P3b peak amplitude defined as the maximum positivity between 350-550 ms at the midline parietal electrode (Pz), did not vary as a function of the clustering condition [$F(4,40)=0.80$] nor did the P3b peak latency [$F(4,40)=0.69$]. A broader P300 window of 350-650 ms did not change this.

To examine whether attention was drawn toward deviants at unattended locations, mean P3bs were compared at the different locations for each condition. Mean voltages (350-550 ms) at Pz for each Location (attended, middle, and far) according to each Condition are displayed in Table 4. Multiple t-tests demonstrated that all amplitudes at attended locations were significantly greater than 0 μV ($p<0.01$ for ES, FM, MF, MA, and AM). No other amplitudes in the middle or far distractor conditions were significantly greater than 0. Three conditions in the middle distractor had mean amplitudes that were significantly less than 0 (ES $p<0.05$, AM $p<0.02$, and FM $p<0.01$). Of these, only the FM condition had a middle sound that was different from its far sound ($p<0.01$). Examination of the ERPs at the Pz site over the P3b interval then, revealed that P3b deflections were only elicited for the attended locations (see Figure 9).

Insert Table 4 about here

Insert Figure 9 about here

MMN

Deviant sounds presented at the task-irrelevant location generated a mismatch negativity (MMN) deflection that superimposed the N1 and P2. The MMN peaked at

approximately 200 ms and was maximum over frontocentral regions. The MMN was quantified as the mean amplitude at Fz, Cz, FC1, and FC2 (150-250 ms) (see Figure 10). As there was no main effect of Hemispace [$F(1,10)=0.95$], attend left conditions were combined with attend right conditions. There was no main effect of Condition [$F(4,40)=2.14$] on the MMN amplitude. However, there was a significant effect of Location (attended, middle, or far streams) [$F(2,20)=9.16$, $p=0.01$, $\eta^2=0.856$], with attended streams demonstrating the greatest MMN amplitude, followed by the next nearest stream, and finally the stream farthest away from the attended stream generating the least MMN amplitudes. Pairwise comparisons substantiated this with the MMN being significantly larger at attended location than at the middle ($p<0.05$), and the far ($p<0.01$) unattended streams. Moreover, the MMN for the middle location was found to be significantly more negative than the far stream ($p<0.05$). This provides evidence that the MMN is, at least for spatial locations, modulated by attention.

Insert Figure 10 about here

The MMN amplitude was also examined as a function of the distance from the attended location regardless of the clustering conditions. Distractors were either 30°, 60°, 90° or 120° away from the attended location. The MMN was found to vary in amplitude as a function of spatial distance from the attended source [$F(4,84)=9.36$ $p<.001$, $\eta^2=0.754$] (see Figure 11). Pairwise comparisons demonstrated that the MMNs located 60°, 90°, and 120° away were less negative than the MMNs at the attended location (all $p<0.005$), or the MMNs located 30° away from the attended location (all $p<0.05$).

Insert Figure 11 about here

Discussion

The aim of the present study was to examine the effect of perceptual context during auditory spatial attention. Results provided some evidence that the perceptual context, as defined by the relative spatial separation between the distractors and target sounds in the auditory scene, influenced selective processing. In particular, moving a distractor closer to an attended location increased the number of errors, while clustering distractors together increased N_d amplitudes.

The enhanced N_d amplitude with the clustering of distractors is consistent with the object-based of selective attention (Alain and Arnott, 2000). The hypothesis predicted that if the relevant feature(s) of an auditory scene (e.g., pitch, location) are presented in such a way that promotes the perceptual grouping of distracting elements, attentional performance should improve. This improvement was predicted to occur independently of whether or not the distracting feature was made more or less similar to the attended feature. The alternative hypothesis (referred to as ‘feature-based’) predicted that by making distracting features more similar to the attended stimulus, performance would decline. The critical condition then was the FM, where the far distracting location was clustered with the other middle distractor, and at the same time made more similar to the attended stimulus. Relative to the ES condition, the feature-based model predicted a decline in performance and N_d amplitude, while the object-based predicted an increase. Results were not as clear-cut. The FM behavioural performance measures did not differ from any other condition. Moreover, the FM manipulation did not demonstrate any significant electrophysiological attention differences (i.e., N_d). However, the object-based model did receive indirect support. Specifically, the MF distractor clustering

condition was the only condition that demonstrated a significant Nd_e difference. This condition differed (i.e., a larger Nd_e amplitude) from all conditions except FM. Because the FM similarity to the MF was not found to be due to an increase in FM variance, one may infer that the Nd_e received attentional enhancement relative to the ES, MA, and AM conditions. A purely feature-based hypothesis would have predicted that that attentional component would have decreased under those circumstances because the distractor was made more similar (spatially) to the relevant channel.

There are at least three possible reasons why the spatial attention effect was smaller than expected: 1) participants were not ignoring the deviant sounds at distracting locations; 2) the virtual locations were not all perceived at the expected azimuth location; 3) the clustering together of auditory locations was not optimal.

Attention being allocated to unattended locations

Despite being instructed not to attend to distractor locations, selective attention effects could potentially be reduced if participant's were actually dividing their attention to the distractor locations (Schwent, et. al, 1976a). Given that the mean ISI rate of 200 ms was as fast or faster than most selective attention studies (including the clustering experiment of Alain and Woods, 1994) it seems unlikely that participants would be spreading their attention. However, the analysis of RTs did demonstrate that participants were fastest to respond to targets preceded by a stimulus occurring at the same location; moderately fast when the preceding target occurred at a moderate distance, and slowest when the preceding stimulus occurred at a distant location. This was predominately the result of the FM condition, however. Nevertheless, this priming effect seems consistent with the idea that subjects were attending to the distracting locations however, it does not

specify what type of attention was being allocated or at what stage this occurred. For example, does the priming effect reflect an early, automatic, sensory-type orienting response, or a later, decision making stage? The difficulty involved in making such inferences is that RT measures are an end product of expectancy, response bias, surprise and early sensory effects (Teder-Sälejärvi and Hillyard, 1998). For this reason, electrophysiological data with its high temporal resolution is more suited for examining such processes.

The ERP data provides two reasons why it does not seem likely that focussed attention was allocated to task-irrelevant locations. First, P3bs were only elicited for deviant sounds at the attended location (and not to any of the deviant sounds at distractor locations). The role or function of the P3b/P300 is still not completely understood, however it has been associated with the transference of information to consciousness, and/or context-updating in working memory (for a review, see Picton, 1992). Regardless, the P3b reflects a late-stage analysis of stimulus features. If participants were extensively processing stimuli presented at these unattended locations, one would have expected P3bs to be present at these locations too.

As well, the analysis of the MMN components suggests that attention was not being divided. The MMN component as previously discussed, has been taken to reflect a neural mismatch between the incoming stimulus and a representation of the previous stimuli. An analysis of the MMN at the three different locations in the present study (attended, middle, or far) revealed that the middle location MMN was less negative than at the attended location but more negative than at the far location. In addition to other evidence (Woldorff et al., 1991; Woldorff et al., 1998), this gradient clearly challenges

the claim that the MMN is an automatic process, independent of attention (see Näätänen, 1992). A re-evaluation of the MMN suggests that the processes reflected in the MMN are ‘partially automatic’ in that their amplitudes can be suppressed or gated if attention is strongly focussed elsewhere (Woldorff et al., 1998). Because the MMN data in the present study were gated according to distance from the task relevant location, one may surmise that attention was strongly focussed on the designated location.

For these reasons, it does not seem likely that divided attention played a significant role in the effects of the clustering manipulation. Also, the priming effect observed in the RT data most likely reflects an early, perhaps automatic sensory response to stimuli at the task-irrelevant location. This explanation is consistent with studies contending that lateralized auditory stimuli affect the processing of subsequent stimuli (e.g., Schröger and Eimer, 1996; Spence and Driver, 1994). For example, Schröger and Eimer (1996) have shown that ERPs to targets are modulated by the spatial position of preceding stimuli, regardless of whether those preceding stimuli are task-relevant or task-irrelevant. They conclude that positional information of task-irrelevant stimuli may trigger exogenous attentional capture. The priming effect may be accounted for in a similar manner.

Virtual auditory locations

There are some indications that the middle location may have been perceived at a 0° location that was not on the same medial plane as the other locations. For example, post-experiment anecdotal feedback from a few participants indicated that the middle tone was not clearly localized in front of the body at 0°, but was heard as being localized “ontop” of the head, or at the centre, “inside” of the head. The cause of inside-the-head

locatedness (IHL) has been attributed to many factors (for a review see Blauret, 1997) however the most compelling factor may have had to do with the fact that there were no interaural time or intensity differences between the two ears for 0° sounds. In the real world, even small head movements will introduce ITDs and IIDs between the ears. However, with headphones, no time or intensity differences are apparent even when the head is moved. To resolve this apparent discrepancy, the auditory system may interpret the sound as coming from a position somewhere along the midline where head movements do not produce any interaural differences. That place is anywhere along the vertical axis of the centre of the head (i.e., 'inside' or 'ontop' of the head) (R. Duda, personal communication, December 1999).

The uniqueness of the 0° sounds was also potentially indicated during the analysis of the P3b interval. Of the ten deviant distractors (middle and far locations in each of the five conditions) only the three 0° locations (i.e., middle ES, middle FM, and middle AM) had mean amplitudes that were significantly less than 0 μ V. Exactly what this reflects is not apparent. It does not seem to be the result of a negativity bias in the baseline (e.g., a processing negativity occurring during the 100 ms baseline prior to stimulus onset). The ISI ranged from 100 to 300 ms, thus a baseline of 100 ms should not have included any processing as no sound was occurring during this interval. The difference does not seem to be related simply to variance either at least in the FM condition, because the 0° FM sound was significantly different from the far (30°) FM sound.

For these reasons, it seems that the perception of the middle tone may have not have been along the same medial plane as the rest of the sounds. If this is true, the implication may be that the clustering manipulations (especially AM and FM) may not

have been as effective as initially predicted.

Auditory location clustering manipulation

It is also possible that a 30° separation between sounds was too far apart to promote effective grouping. Aside from an increase in FAs for the MA condition, the behavioural data provided little indication that grouping occurred. This result must be interpreted with caution however because auditory spatial processing differences observed in some ERPs are not always reflected in RT measures (Schröger and Eimer, 1996; Simon, Acosta, and Mewaldt, 1975; Spence and Driver, 1994; Teder-Sälejärvi and Hillyard, 1998). This may be due to differences in the relative sensitivity between the two types of measurements and/or to what each measure was quantifying (e.g., behavioural measures may have predominately reflected within block variability rather than variability between conditions, whereas ERP measures were reflecting the perceptual processes).

It is likely that clustering is dependent on the task difficulty. The harder the task, the more likely it would seem that ‘helpful’ strategies such as grouping would be adopted (consciously or unconsciously). The high Hit and low Error rates of the behavioural performance suggests that the task was relatively easy. One might expect that a spatial separation of less than 30° may have increased the difficulty and thus taxed the attentional system more. A comparison of studies from the Teder-Sälejärvi group revealed that the attentional gradient became steeper and narrower when distracting speakers were located at intervals of 6° from the attended location (Teder-Sälejärvi et al., 1999a) rather than 9° (Teder-Sälejärvi and Hillyard, 1998). Moreover, both of these studies used much smaller attentional windows of 18° in comparison to the 90° or 120° used in the present study.

Thus, the most likely reason why smaller-than-expected attentional effects were observed in the present study was due to the clustering manipulation itself. The spatial separations adopted may have been too large. Moreover, it seems that the perceptual location of the 0° sounds, as presented over headphones using HRTFs, may have also contributed to a reduced perceptual grouping effect. Despite these difficulties, the present results provide support for an object-based model of selective attention.

Spatial gradient of auditory attention

A somewhat unexpected finding was that the spatial gradient of attention did not simply decrease with increasing distance from the attended location (refer to Figure 8). In the current study, far distractor sounds produced an N1 amplitude that was intermediate between attended and middle distractors. The most striking example was in the ES condition where N1 amplitudes were most negative for the attended location, least negative for the middle distractor, and moderately negative for the far distractor. Although not illustrated, this was true for both the ES_L and ES_R conditions. This is surprising on several accounts.

First, given the finding that central sounds are more easily localized than peripheral sounds (Perrott, Costantino, and Cisneros, 1993; Röder, Teder-Sälejärvi, Sterr, Rösler, Hillyard, and Neville, 1999; Teder-Sälejärvi and Hillyard, 1998), electrophysiological responses to a central distractor location were expected to have been greater than responses to a peripheral distractor sound. Second, two recent auditory localization ERP studies (Teder-Sälejärvi and Hillyard, 1998; Teder-Sälejärvi et al., 1999a) reported that attention to a sound location was distributed as a gradient that tapers off as a function of the distance from the attended point in space. In these studies much

smaller sound separations of 9° and 6° were used. While it is tempting to propose that the spatial separations adopted in the current paradigm were too large to observe a gradient, results by Mondor and Zatorre (1995) suggest otherwise. Using behavioural measures, they report evidence of spatial gradients using comparable separation distances of 30° and 45° between speakers. Finally, Alain and Woods' (1994) clustering study demonstrated an 'attention gradient' by manipulating the frequency separation. They demonstrated greatest N1 negativity for the attended high pitch, moderate negativity to the medium distractor pitch, and minimal negativity to the lowest distractor pitch. Although these amplitude distributions varied with each clustering condition, the amplitude gradient was always maintained. All of these findings indicated that a gradient should have been observed in the current study.

One possible explanation that can easily be dismissed is a problem of baseline bias. If several targets in a row directly preceded the unattended sound, consecutive target stimuli could build up a long-duration negative shift (Michie, Solowij, Crawford, and Glue, 1993). ERPs to unattended stimuli would therefore appear to be positively shifted relative to the attended stimuli. This however, cannot be the case in the present study, as a minimum of one standard sound had to precede any deviant sound.

Another possibility is that of active rejection of distracting locations (Alho, Woods, Algazi, and Näätänen, 1994; Donald, 1987; Michie, Bearpark, Crawford, and Glue, 1990; Michie et al., 1993; Woods, 1990). Since Broadbent's filter model, attention has been characterized by a broad array of metaphors including a 'spotlight' that moves (Tsal, 1983), and a 'zoom' lens (Eriksen and St. James, 1986). Central to these models is that of enhanced processing of task-relevant information. However, there is some

evidence for the active rejection of unwanted stimuli. Results from an auditory ERP study of Michie et al. (1990) demonstrated changes in both attended (greater negativity) and unattended ERPs (less negativity) relative to that of a control condition. Woods (1990) thus proposed a 'parasol' metaphor whereby an inhibitory surround is centred on the rejected sounds. This would function to reduce the processing of frequently repeated, distracting sounds that were similar to the target.

This idea is a compelling one because it is consistent with the N1 gradient observed in the present study. The ERPs for the middle task-irrelevant location may be less negative than those for the farthest task-irrelevant location because they are being actively inhibited. The reason why a positivity for the far distracting location is not observed would presumably be a result of it being far enough away from the task-relevant location that it is no longer a source of confusion. Also, active rejection of distracting locations would explain why the two unattended components of the NdMF and NdFM wave differed in amplitude despite their both being separated by the same 30° distance. Perhaps when a distractor is moved out to the periphery, 30° separation becomes harder to distinguish (see Note 1). Thus, the middle distractor is not heard to be that different from the far distractor and as such receives the enhanced positivity that is applied to 'middle' tones (or at least tones nearest to the attended tone). However, if this is true, one must ask why a similar positivity for ERPs in neighbouring distracting locations was not observed in other spatial selection studies like Teder-Sälejärvi and Hillyard's (1998).

Probably the simplest explanation as to why the unexpected attentional gradient was obtained relates to a previous observation; i.e., there was something 'different' about the 0° HRTF noise bursts relative to peripheral sounds. Pairwise comparisons revealed

that only certain conditions demonstrated the unexpected gradient. Specifically, only the middle sounds in the ES and AM conditions were more positive than either the attended or far locations. This AM gradient pattern was all the more surprising given that the attended sounds were spatially clustered with the middle sounds (i.e., if the attended sound was more similar to the middle sound, the middle sound should have received more attention and therefore more negativity than the far distractor sound). Importantly, the middle sounds in the ES and AM conditions were both located at 0°. Although the other 0° position (FM condition) did not differ from the far location, this is most likely the result of a confound with the clustering manipulation (i.e., the far sound was made to sound more spatially similar to the middle sound).

In those conditions where the middle stimuli were located 30° off midline, the expected gradient was obtained. In the MF condition, one would have expected the N1 to be more negative than at the other two locations, but the middle and far (clustered) N1s to be similar. This was true. In the MA condition, one would have expected the attended location N1 to be similar to the middle location (clustered) but to be more negative than the far location. This too, was true. These observations indicate that the 0° stimuli were processed differently in some way. What could have caused this difference?

Neural refractoriness and habituation (N1 response decrement due to repeated, rapid stimulation of the same neural populations) provide a possible answer (Budd, Barry, Gordon, Rennie, and Michie, 1998). It has been shown for instance, that N1 components decrease (in amplitude) as the frequency similarity between the tones comprising the sequence increases (Butler, 1968). Presumably this is due to the fact that as the stimuli become physically more alike, they begin to stimulate similar populations

of neurons. The resulting neural response therefore becomes weaker as the stimulation rate begins to overlap with the recovery cycle of the neural generators.

Along a similar line, neuronal populations responding to 'middle' sounds may also be reduced because they are always flanked by peripheral stimuli. This relates to an observation made by Schwent, Snyder, and Hillyard (1976c). They noted that the N1 waves of their middle channel (binaural presentation over headphones) were not as sensitive to attention as were the N1s of the other two flanking channels. Drawing comparisons from early work by Treisman (1964b), they suggested that the middle channel was perhaps more susceptible to 'perceptual interference' from the two peripheral channels. Regardless of what process(es) underlie perceptual interference (e.g., neural refractoriness?), if perceptual interference was affecting the middle sounds in the current paradigm, one might expect there to have been some differences among the N1 amplitudes of the middle sounds as a function of the clustering manipulations. For example, in the ES condition the middle stimuli were located 60° from the attended location, and 60° from the task-irrelevant location. However, in the MA condition, the middle stimuli were located 30° from the attended location and 90° from the task-irrelevant location (and vice versa in the MF condition). Despite these manipulations, N1 amplitudes were similar across all conditions for middle stimuli. While this finding argues against a strong role of perceptual interference for the middle tones (there was after all a 30° separation between sounds), it does not rule out its potential contribution as the non-significance may simply reflect a problem of power.

The evidence reviewed here as well as that discussed previously regarding the P3b seems however, to suggest that the 0° middle sound was particularly problematic. If

these sounds were perceived as being located from inside the head (IHL), then it is possible that the evoked response components were reduced in amplitude relative to the more salient, externalized locations. It has been shown that the N1 evoked by binaural stimuli are substantially smaller than what would be expected from the addition of two monaural stimuli (see Näätänen and Picton, 1987). This may result from mutual inhibition of left and right auditory cortices (Pantev, Lütkenhoner, Hoke, and Lehnertz, 1986). Recall that the 0° sound was created by playing exactly the same sound to each ear at the same time (i.e., no ITD or IID differences). Mutual inhibition would likely be optimal for the 0° stimulus. It therefore seems that the middle stimulus, specifically the 0° sound was responsible for the unexpected attentional gradient that was observed in some clustering manipulations.

Nd laterality

As a final note, there is some controversy in the literature as to whether or not the Nd wave is lateralized over the contralateral hemisphere. For example Michie et al. (1990) did not demonstrate Nd_c laterality. In that study, participants were instructed to attend to brief tones of a particular location (left/right ear) and pitch (900/960 Hz) combination and respond to those targets of longer duration. Although the N1 component demonstrated a laterality effect (hemisphere contralateral to the attended ear producing a larger N1), the Nd_c did not. The results reported here are however, consistent with more recent results from Alho et al. (1994) and Teder-Sälejärvi, Münte, Sperlich, and Hillyard (1999b). These groups found a larger Nd_c over the hemisphere contralateral to the side of sound presentation, but a symmetrically distributed Nd_i. Such findings are consistent with the notion that the early and late components of the Nd wave reflect

different neural generators and index distinct attentional processes (c.f. Näätänen, 1992; Teder-Sälejärvi and Hillyard, 1998). The analysis of the Nd₁ data in the present study, was not affected by the condition manipulations.

Summary

The results presented here are consistent with studies demonstrating that the manipulation of perceptual context can affect performance on attentional tasks (Bregman and Rudnick, 1975; Alain and Woods, 1993; 1994). These studies support the object-based hypothesis of selective attention and therefore place a requirement on any general selective attention theory to be able to accommodate the existence of higher-order interactions such as grouping. In addition, the present study provides evidence that the MMN may be modified by attention, and that early and late components of the Nd wave index distinct neural processes.

References

- Alain, C., & Arnott, S.R. (2000). Selectively attending to auditory objects. Frontiers in Bioscience, 5, 202-212.
- Alain, C., & Woods, D.L. (1993). Distractor clustering enhances detection speed and accuracy during selective listening. Perception and Psychophysics, 54, 509-514.
- Alain, C., & Woods, D.L. (1994). Signal clustering modulates auditory cortical activity in humans. Perception and Psychophysics, 56, 501-516.
- Alho, K., Donauer, N., Paavilainen, P., Reinikainen, K., Sams, M., & Näätänen, R. (1987). Stimulus selection during auditory spatial attention as expressed by event-related potentials. Biological Psychology, 24, 153-162.
- Alho, K., Sams, M., Paavilainen, P., & Näätänen, R. (1986). Small pitch separation and the selective-attention effect on the ERP. Psychophysiology, 23, 189-197.
- Alho, K., Woods, D.L., & Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. Psychophysiology, 31, 469-479.
- American Electroencephalographic Society: Guidelines for standard electrode position nomenclature. (1991). Journal of Clinical Neurophysiology, 8, 200-202.
- Berg, P., & Scherg, M.A. (1994). A multiple source approach to the correction of eye artifacts. Electroencephalography and Clinical Neurophysiology, 90, 229-241.
- Blauret, J. (1997). Spatial Hearing: The Psychophysics of Human Sound Localization. Cambridge, MA: The MIT Press.
- Bregman, A.S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: The MIT Press.

- Bregman, A.S., & Rudnick, (1975). Auditory segregation: Stream or streams? Journal of Experimental Psychology: Human Perception and Performance, 1, 263-267.
- Broadbent, D.E. (1958). Perception and Communication. New York: Pergamon.
- Broadbent, D.E. (1970). Stimulus set and response set: Two kinds of selective attention. In D.I. Mostofsky (Ed.), Attention: Contemporary theory and analysis (pp.51-60).
- Broadbent, D.E. (1982). Task combination and selective intake of information. Acta Psychologica, 50, 253-290.
- Budd, T.W., Barry, R.J., Gordon, E., Rennie, C., & Michie, P.T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: Habituation vs. refractoriness. International Journal of Psychophysiology, 31, 51-68.
- Butler, R.A. (1968). Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. Journal of the Acoustical Society of America, 44, 945-950.
- Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two experiments. Journal of the Acoustical Society of America, 25, 975-979.
- Coles, M.G.H., & Rugg, M.D. (1995). Event-related brain potentials: An introduction. In M.D. Rugg and M.G.H. Coles (Eds.), Electrophysiology of Mind: Event-Related Brain Potentials and Cognition (pp. 1-23). Oxford: Oxford University Press.
- Deutsch, J.A., & Deutsch, D. (1963). Attention: Some theoretical considerations. Psychological Review, 70, 80-90.
- Donald, M.W. (1987). The timing and polarity of different attention-related ERP changes inside and outside of the attentional focus. In R. Johnson, Jr., J.W. Rohrbaugh, &

R. Parasuraman (Eds.), Current Trends in Event-Related Potential Research, (pp.81-86).

New York: Elsevier.

Duda, R.O. (Personal communication Dec. 17, 1999).

Eriksen, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. Perception and Psychophysics, **25**, 249-263.

Glaser, E.M., & Ruchkin, D.S. (1991). Principles of Neurobiological Signal Analysis. New York: Academic Press.

Han, H.L. (1994). Measuring a dummy head in search of pinna cues. Journal of the Audio Engineering Society, **42**, 15-37.

Handel, S. (1987). Listening: An Introduction to the Perception of Auditory Events. Cambridge: MIT Press.

Hansen, J.C., & Hillyard, S.A. (1980). Endogenous brain potentials associated with selective auditory attention. Electroencephalography and Clinical Neurophysiology, **49**, 277-290.

Hansen, J.C., & Hillyard, S.A. (1983). Selective attention to multidimensional auditory stimuli. Journal of Experimental Psychology: Human Perception and Performance, **9**, 1-19.

Hillyard, S.A., Hink, R.F., Schwent, V.L., & Picton, T.W. (1973). Electrical signs of selective attention in the human brain. Science, **182**, 177-180.

Hillyard, S.A. & Kutas, M. (1983). Electrophysiology of cognitive processing. Annual Review of Psychology, **34**, 33-61.

Hink, R.F., Van Voorhis, S.T., & Hillyard, S.A. (1977). The division of attention and the human auditory evoked potential. Neuropsychologia, **15**, 597-605.

Knight, R.T., Scabini, D., Woods, D.L., & Clayworth, C.C. (1989). Contributions of temporal-parietal junction to the human auditory P3. Brain Research, 502, 109-116.

Kulkarni, A., & Colburn, H.S. (1998). Role of spectral detail in sound-source localization. Nature, 396, 747-749.

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

Linden, R.D., Picton, T.W., Hamel, G., & Campbell, K.B. (1987). Human auditory steady-state evoked potentials during selective attention. Electroencephology and clinical Neurophysiology, 66, 145-159.

Lins, O.G., Picton, T.W., Berg, P., & Scherg, M. (1993a). Ocular artifacts in EEG and event-related potentials. I Scalp topography. Brain Topography, 6, 51-63.

Lins, O.G., Picton, T.W., Berg, P., & Scherg, M. (1993b). Ocular artifacts in recording EEGs and event-related potentials. II Source dipoles and source components. Brain Topography, 6, 65-78.

McKinley, R.L., & Ericson, M.A. (1988). Digital synthesis of binaural auditory localization azimuth cues using headphones. Journal of the Acoustical Society of America, Suppl. 1, 83, S18.

Michie, P.T., Bearpark, H.M., Crawford, J.M., & Glue, L.C.T. (1990). The nature of selective attention effects on auditory event-related potentials. Biological Psychology, 30, 219-250.

Michie, P.T., Solowij, N., Crawford, J.M., & Glue, L.C. (1993). The effects of between-source discriminability on attended and unattended auditory ERPs. Psychophysiology, 30, 205-220.

Middlebrooks, J.C., Makous, J.C., & Green, M.D. (1989). Directional sensitivity of sound-pressure levels in the human ear canal. Journal of the Acoustical Society of America, 86, 89-108.

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

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. Quarterly Journal of Experimental Psychology, 11, 56-60.

Näätänen, R. (1992). Attention and Brain Function. Hillsdale, New Jersey: Lawrence Erlbaum.

Näätänen, R., & Picton, T.W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology, 24, 375-425.

Pantev, C., Lütkenhoner, B., Hoke, M., & Lehnertz, K. (1986). Comparison between simultaneously recorded auditory-evoked magnetic fields and potentials elicited by ipsilateral, contralateral, and binaural tone burst stimulation. Audiology, 25, 54-61.

Parasuraman, R. (1980). Effects of information demands on slow negative shift latencies and N100 amplitude in selective and divided attention. Biological Psychology, 11, 217-233.

Perrott, D.R., Costantino, B., & Cisneros, J. (1993). Auditory and visual localization performance in a sequential discrimination task. Journal of the Acoustical Society of America, 93, 2134-2138.

Pichora-Fuller M. K., Schneider B. A., & Daneman M. (1995). How young and

old adults listen to and remember speech in noise. Journal of the Acoustical Society of America, 97, 593-608.

Picton, T.W. (1990). Auditory evoked potentials. In D.D. Daly and T.A. Pedley (Eds.), Current Practice of Clinical Electroencephalography, (2nd ed.). New York: Raven Press.

Picton, T.W. (1992). The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology, 9, 456-479.

Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Baynard, J., & Trujillo, N.J. (1999). Intracerebral sources of human auditory-evoked potentials. Audiology and Neuro-Otology, 4, 64-79.

Picton, T.W., & Hillyard, S.A. (1974). Human auditory evoked potentials. II: Effects of attention. Electroencephalography and Clinical Neurophysiology, 36, 191-199.

Picton, T.W., Lins, O.G., & Scherg, M. (1995). The recording and analysis of event-related potentials. In Boller and Grafman (Eds.), Handbook of Neuropsychology Vol. 10, (pp. 3-73). Amsterdam: Elsevier Science.

Picton, T.W., Stapells, D.R., & Campbell, K.B. (1981). Auditory evoked potentials from the human cochlea and brainstem. The Journal of Otolaryngology, 10S, 1-41.

Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S., & Neville, H.J. (1999). Improved auditory spatial tuning in blind humans. Nature, 400, 162.

Scherg, M. (1990). Fundamentals of dipole source analysis. In F. Grandor, M. Hoke, and G.L. Romani, (Eds.), Auditory Evoked Magnetic Fields and Electric

Potentials. Advances in Audiology, Vol. 5, (pp. 40-69). Basel: Karger.

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

Schwent, V.L., & Hillyard, S.A. (1975). Evoked potential correlates of selective attention with multichannel auditory inputs. Electroencephalography and Clinical Neurophysiology, 38, 131-138.

Schwent, V.L., Hillyard, S.A., & Galambos, R. (1976a). Selective attention and the auditory vertex potential: I. Effects of stimulus delivery rate. Electroencephalography and Clinical Neurophysiology, 40, 604-614.

Schwent, V.L., Hillyard, S.A., & Galambos, R. (1976b). Selective attention and the auditory vertex potential: II. Effects of signal intensity and masking noise. Electroencephalography and Clinical Neurophysiology, 40, 615-622.

Schwent, V.L., Snyder, E., & Hillyard, S.A. (1976c). Auditory evoked potentials during multichannel selective listening: Role of pitch and localization cues. Journal of Experimental Psychology: Human Perception and Performance, 2, 313-325.

Simon, J.R., Acosta, E., & Mewaldt, S.P. (1975). Effect of locus warning tone on auditory choice reaction time. Memory and Cognition, 3, 167-170.

Spence, C., Nichols, M., Gillespie, N., & Driver, J. (1997). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. Perception and Psychophysics, 60, 544-557.

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

Teder, W., Kujala, T., & Näätänen, R. (1993). Selection of speech messages in free-field listening. Neuroreport, 5, 307-309.

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

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

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

Teder-Sälejärvi, W.A., Münte, T.F., Sperlich, F-J., & Hillyard, S. (1999b). Intra-modal and cross-modal spatial attention to auditory and visual stimuli. An event-related brain potential study. Cognitive Brain Research, 8, 327-343.

Treisman, A.M. (1960). Contextual cues in selective filtering. Quarterly Journal of Experimental Psychology: Human Perception and Performance, 12, 242-248.

Treisman, A.M. (1964a). The effect of irrelevant material on the efficiency of selective listening. The American Journal of Psychology, 77, 533-546.

Treisman, A.M. (1964b). Selective attention in man. British Medical Bulletin, 20, 12-16.

Tsal, Y. (1983). Movement of attention across the visual field. Journal of Experimental Psychology: Human Perception and Performance, 9, 523-530.

Vaughan, H.G. Jr., & Arezzo, J.C. (1988). The neural basis of event-related potentials. In T.W. Picton (Ed.). Handbook of Electroencephalography and Clinical

Neurophysiology, Revised Series, Vol. 3, (pp. 45-96). Amsterdam: Elsevier.

Woldorff, M.G., Hackley, S.A., & Hillyard, S.A. (1991). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones.

Psychophysiology, 28, 30-42.

Woldorff, M.G., & Hillyard, S.A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalography and Clinical Neurophysiology, 79, 170-191.

Woldorff, M.G., Hillyard, S.A., Gallen, C.C., Hampson, S.R., & Bloom, F.E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. Psychophysiology, 35, 283-292.

Woods, D.L. (1990). The physiological basis of selective attention: Implications of event-related potential studies. In J.W. Rohrbaugh, R. Parasuraman, and R. Johnson, Jr. (Eds.). Event-Related Brain Potentials, (pp. 178-209). New York: Oxford University Press.

Yost, W.A. (1997). The cocktail party problem: Forty years later. In R.H. Gilkey and T.R. Anderson (Eds.) Binaural and Spatial Hearing in Real and Virtual Environments. (pp. 329-347). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Yost, W.A., & Dye, R.H. Jr., & Sheft, S. (1996). A simulated "cocktail party" with up to three sound sources. Perception and Psychophysics, 58, 1026-1036.

Table 1

Mean Performance for Each Condition Collapsed Over Attended Location

Condition	Misses %	False Alarms %	Reaction Time (SD) msec
ES	7.5	0.19	472 (13)
FM	7.4	0.23	469 (10)
MF	7.9	0.17	469 (11)
MA	13.2	0.49*	480 (14)
AM	11.5	0.31	482 (14)

* $p < 0.05$

Table 2

Mean RT for Targets in each Condition According to the Preceding Sound Location

Condition	Preceding Sound Location		
	Identical	Middle	Far
ES	465 (12)	467 (14)	477 (15)
FM	454 (12)	466 (9)	479 (11)
MF	459 (10)	469 (13)	470 (13)
MA	461 (15)	492 (12)	490 (17)
AM	477 (19)	481 (18)	487 (16)

Table 3

Mean FC1 and FC2 Voltage (125-165 ms) for Each Condition Collapsed Over Attended Location

Condition	Location		
	Attended	Middle	Far
ES	-0.95±0.15	0.05±0.08	-0.34±0.07
FM	-0.95±0.09	0.02±0.08	-0.03±0.08
MF	-0.94±0.10	-0.04±0.09	-0.19±0.09
MA	-0.70±0.14	-0.11±0.06	-0.01±0.08
AM	-0.64±0.12	0.02±0.07	-0.24±0.09

Table 4**Mean Pz Voltage (350-550 ms) for Each Condition Collapsed Over Attended Location**

Condition	Location		
	Attended	Middle	Far
ES	1.78±0.58	-0.38±0.16	-0.15±0.19
FM	1.42±0.59	-0.29±0.14	-0.06±0.13
MF	1.62±0.53	-0.47±0.17	0.17±0.21
MA	1.56±0.49	-0.20±0.20	-0.18±0.09
AM	1.32±0.56	0.01±0.16	-0.06±0.16

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

Figure 9

Figure 10

Figure 11

Appendix A

Instructions to Participants

In this experiment you will hear very brief sounds played to you over headphones. They will be played to you at a fast rate (about 3-4 sounds per second). Each sound will be located in one of three possible locations (for example: to the left of your head, directly in front of your head, or to the right of your head). Occasionally, you will hear sounds of longer duration. You will be required to attend to a designated location and press the response button only to the longer duration sounds that occur at that location. You should try to respond as fast as you can while at the same time trying to make as few errors as possible.

Appendix B

Calculation of the Nd Waves

Nd condtion	Calculation of Nd*
NdES	ES _L left stream – ES _R left stream ES _R right stream – ES _L right stream
NdFM	FM _L left stream – AM _R left stream FM _R right stream – AM _L right stream
NdMF	MF _L left stream – MA _R left stream MF _R right stream – MA _L right stream
NdMA	MA _L left stream – MF _R left stream MA _R right stream – MF _L right stream
NdAM	AM _L left stream – FM _R left stream AM _R right stream – FM _L right stream

* subscripts indicate the attended location: left-most stream (L) or right-most stream (R)