

The gradient of spatial auditory attention in free field: An event-related potential study

WOLFGANG A. TEDER-SÄLEJÄRVI and STEVEN A. HILLYARD
University of California, San Diego, La Jolla, California

Young adult subjects attended selectively to brief noise bursts delivered in free field via a horizontal array of seven loudspeakers spaced apart by 9° of angle. Frequent "standard" stimuli (90%) and infrequent "target/deviant" stimuli (10%) of increased bandwidth were delivered at a fast rate in a random sequence equiprobably from each speaker. In separate runs, the subjects' task was to selectively attend to the leftmost, center, or rightmost speaker and to press a button to the infrequent "target" stimuli occurring at the designated spatial location. Behavioral detection rates and concurrently recorded event-related potentials (ERPs) indicated that auditory attention was deployed as a finely tuned gradient around the attended sound source, thus providing support for gradient models of auditory spatial attention. Furthermore, the ERP data suggested that the spatial focusing of attention was achieved in two distinct stages, with an early more broadly tuned filtering of inputs occurring over the first 80–200 msec after stimulus onset, followed by a more narrowly focused selection of attended-location deviants that began at around 250 msec and closely resembled the behavioral gradient of target detections.

In an acoustically complex environment, human listeners are able to attend selectively to one sound source in the presence of other competing sound sources—a phenomenon commonly known as the *cocktail party effect*. Extensive studies of selective listening to spoken messages have shown that relevant voices can be segregated and attended to on the basis of various cues, including spatial position, pitch, and voice dynamics (Cherry, 1953; Wood & Cowan 1995; Yost, 1992). When multiple (more than two) sources are concurrently active, however, spatial cues become increasingly important for solving the cocktail party problem (Yost, Dye, & Sheft, 1996). Despite the obvious role of spatial cues in distinguishing attended and unattended sound sources, little is known about the spatial allocation of auditory attention among multiple competing inputs, primarily because most studies of the cocktail party effect have presented only two sound sources (Yost, 1997). In particular, only a few studies have examined how narrowly auditory attention may be focused or tuned spatially to a relevant source when adjacent sources may also be active.

Mondor and Zatorre (1995) investigated this question in a free-field tone discrimination task in which the target stimuli were randomly presented at one of six spatial positions located 15°, 45°, and 75° left and right of cen-

ter (their Experiment 4). Targets were preceded by a cue that indicated the most probable location of target occurrence. It was found that discriminative reaction times (RTs) were fastest to targets presented at the cued location and increased progressively as a function of target distance from that location. The authors concluded that, in their trial-by-trial cuing paradigm, "auditory attention appeared to be distributed as a gradient, with the density of resources declining gradually with distance from an attentional focal point." This conclusion is in line with numerous studies of visual spatial attention in which gradient models of attentional allocation provided a good account of the results (Cheal, Lyon, & Gottlob, 1994; Downing, 1988; Downing & Pinker, 1985; LaBerge & Brown, 1989; Shulman, Wilson, & Sheehy, 1985; Yund, Efron, & Nichols, 1990).

Behavioral studies of attentional gradients have not provided definitive evidence about the level(s) of processing at which the spatially graded allocation of resources takes place. An early sensory locus for visuospatial attention has been proposed on the basis of observed gradients of signal detectability (Downing, 1988), but this interpretation has been contested and debated (Hawkins et al., 1990; Luck et al., 1994; Müller & Humphreys, 1991). It is particularly difficult to make inferences about processing levels of attentional selection when RT is used as a dependent measure, since response speed is multiply determined by decision-level factors, such as expectancy, response bias, and surprise, and by early sensory effects (Sperling, 1984).

The question of whether attentional gradients are determined at an early sensory level can also be investigated by recording event-related brain potentials (ERPs) in con-

This work was supported by Grants MH-25594 from NIMH, Grant N00014-93-1-0942 from ONR, and Grant NS17778 from NIH. The authors thank Carlos Nava and Matt Marlow for technical assistance. Correspondence should be addressed to W. A. Teder-Sälejärvi, Department of Neurosciences 0608, School of Medicine, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0608 (e-mail: wat@sdepl.ucsd.edu).

junction with behavioral data. ERP recordings provide precisely timed measures of sensory processing at successive levels of the afferent pathways and cortical areas, and specific ERP components have been found to be increased in amplitude when attention is focused on the eliciting stimuli (for a review, see Näätänen, 1992). In the visual modality, Mangun and Hillyard (1988) found that the amplitudes of early cortical ERP components (latency 80–200 msec) were enhanced to attended-location stimuli and declined progressively to stimuli at more distant locations, in parallel with a decline in target detectability. They concluded that gradients of visual attention were determined at least in part at early sensory levels of processing. Longer latency components in the range of 250–500 msec, however, were found to be elicited only by the attended-location stimuli. Mulder, Wijers, Brookhuis, Smid, and Mulder (1994) extended these findings and proposed that selection for location involves two distinct stages: The first produces a relatively broad and shallow spatial gradient, whereas the second results in a steeper, more narrowly focused gradient.

In the auditory modality, the chronometry of stimulus processing may be monitored from brainstem to cortex by means of ERP recordings. In experiments on selective listening to dichotic tone sequences, the earliest ERP component found to be affected by attention was a positive wave elicited in the auditory cortex at a latency of 20–50 msec (Woldorff et al., 1993; Woldorff & Hillyard, 1991). A much more robust ERP attention effect, however, is an enhanced negative wave that is elicited in the auditory cortex by attended stimuli starting at 60–70 msec after stimulus onset (for reviews, see Hillyard, Mangun, Woldorff, & Luck, 1995; Näätänen, 1992). This attention-related negativity overlaps in time with the evoked N1 wave (peaking at 100–130 msec) and may extend for a few hundred milliseconds thereafter. This attention effect may be measured either as the amplitude of the N1 component itself ("N1 effect") or as the negative difference (Nd) wave obtained by subtracting the ERP elicited by a given stimulus when unattended from the ERP to the same stimulus when attended.

The enhanced N1/Nd negativity (also termed *processing negativity*) is typically elicited with shorter latencies and greater amplitudes when the attended and unattended sounds are distinguished by rapidly discernible physical cues, such as location or pitch (Näätänen, 1982, 1992). Accordingly, this negative ERP has been considered a signature of a rapid, early selection between "channels" of competing auditory inputs defined by such cues as location and pitch (Hansen & Hillyard, 1988). The N1/Nd amplitude reportedly increases as a function of attentional allocation to a particular input channel, in parallel with improved behavioral accuracy at detecting target events in that channel (Hink, Voorhis, Hillyard, & Smith, 1977). The functional significance of the N1/Nd has been variously interpreted in terms of the amount of sensory information passing through an early channel selection mechanism (Hillyard, Hink, Schwent, & Picton, 1973),

the further processing of attended-channel information (Näätänen, 1982; Okita, 1981), and the goodness of match between the eliciting stimulus and the cue characteristics of the attended input channel (Näätänen, 1992).

In ERP studies of auditory attention, subjects are typically required to monitor a specified input channel (ignoring the other channel or channels) and to detect infrequent occurrences of target events (usually difficult to discriminate) within the attended channel. The identification of such targets is associated with a longer latency positive ERP component (the "P3" or "P300") wave having a peak latency of 300–400 msec. A number of studies have shown that the P3 is specifically elicited by targets in the attended channel and not by comparable deviant stimuli in an unattended channel (e.g., Hillyard et al., 1973; Hink, Fenton, Pfefferbaum, Tinklenberg, & Kopell, 1978). These findings have suggested that the N1/Nd and P3 waves are indices of two hierarchically organized levels of selection, with the N1/Nd reflecting an initial between-channel selection based on easily discriminable cues and the P3 indexing a subsequent within-channel target selection following a more elaborated and detailed processing of the relevant stimulus properties. The actual postselection processing events reflected in the P3 are still being debated, however. One widely discussed proposal is that the P3 is associated with the updating of working memory upon detection of a task-relevant event (for recent reviews, see Picton, 1992, and Verleger, 1997).

Several studies have used ERPs to examine the spatial allocation of auditory attention among multiple input channels. Schwent and Hillyard (1975) presented through headphones randomized sequences of tones belonging to four channels that were readily distinguishable both in pitch and in perceived location. Subjects attended to one of the four channels at a time. The results showed that the N1/Nd was enlarged to tones only in the attended channel, with no amplitude enhancement spreading to spatially adjacent channels. A narrow focus of attention indexed by N1/Nd was also reported by Schwent, Snyder, and Hillyard (1976) in a three-channel design. However, in a similar type of study that presented tones over five channels distinguished by location and pitch, Hink et al. (1978) found that N1/Nd amplitudes were enhanced not only by the attended-channel tones but also by the tones in adjacent channels. Possible factors that might account for these differing patterns of results among studies include differences in stimulus presentation rate, spatial separation between auditory channels, frequency separation between channels, and type of attention-directing task. Thus, the steepness of the gradient or spatial tuning curve of auditory attention around a relevant sound source may be influenced by a number of stimulus and task variables.

In the present experiment we made a systematic attempt to combine behavioral and ERP measures of the spatial allocation of auditory attention in free field using an array of seven closely spaced loudspeakers. One of the goals was to obtain more precise information on the

early sensory basis for attentional gradients than was provided by previous ERP studies by relating ERP waveform measures to gradients of target detection rates and RTs. Specifically, we wanted to find out to what extent spatial gradients of auditory attention are determined at early levels of sensory input filtering (N1/Nd latency range) and at later levels of target selection and recognition (indexed by P3). By examining the time course of these ERP attention effects, it was possible to obtain information about the temporal properties of the spatial focusing of attention in free field.

The present study also included the first behavioral investigation of attentional gradients in a sustained attention paradigm in which auditory stimuli were presented in rapid, continuous sequences from multiple sources. Mondor and Zatorre's (1995) study of auditory gradients used a trial-by-trial cuing task in which target stimuli were presented individually and attention was redirected by a cue on each trial. These two types of tasks may well engage different types of attentional processes (Schröger & Eimer, 1997; Spence & Driver, 1996). Indeed, ERP studies have demonstrated that the attention-related Nd waves are elicited with longer latency and a more posterior scalp distribution in trial-by-trial cuing tasks than in sustained attention tasks (Schröger & Eimer, 1993, 1997).

METHOD

Subjects

Twelve young right-handed adult subjects were paid for participation (5 women, 7 men; age range = 18–39 years; $M = 23$ years). According to self-report, all subjects had normal hearing.

Stimuli and Apparatus

The stimuli were 82-msec bursts of pink noise (including a total of 10 msec of rise/fall time) with a bandwidth of either 500–5000 Hz (frequent "standard" stimuli) or 500–15000 Hz (infrequent "deviant/target" stimuli). The stimuli were band-pass-filtered with IV-order Bessel filters (-24 dB/octave) from computer-generated broadband noise fed into a pink noise converter (constant energy per octave) and were presented through an array of loudspeakers.¹ The speakers in the array were spaced 9° of angle apart (altogether 54° of angle) on a horizontally oriented hoop, with the middle speaker centered with respect to the head of the subject at a distance of 1.0 m (Figure 1). Noise bursts were presented in random order from the speakers in the array at an intensity of 76 dB SPL as measured at the subject's head. Stimulus onset asynchronies (SOAs) varied randomly between 90 and 270 msec (180 msec on the average with a rectangular distribution). For each speaker, the probability of the occurrence of the standards was $p = .87$ and that of the deviants was $p = .13$.

The subject's head position was monitored with a custom-made device that reflected a narrow beam of infrared (IR) light from a small mirror mounted on top of the subject's head. This device registered vertical displacements of the head-mounted mirror with a sensitivity of $\pm 4^\circ$ of angle and registered horizontal shifts with a sensitivity of $\pm 2^\circ$ of angle. The output of this device, in addition to serving as a monitor for the experimenter outside the chamber, provided feedback to the subject by illuminating the fixation point (a light-emitting diode [LED] mounted under the center speaker), only when his/her head was correctly positioned (see Figure 1).

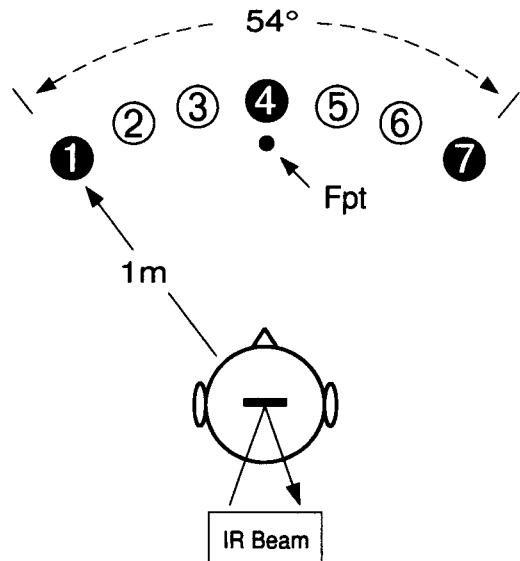


Figure 1. Experimental arrangement of loudspeaker array, head movement monitor with infrared (IR) beam and active fixation point (Fpt). The spacing between the speakers was 9° of angle, yielding a total of 54° of angle. The labels S1 through S7 for the speaker positions will be used throughout the text.

EEG Recording

The EEG was recorded from 44 electrodes mounted in an elastic cap (see Clark, Fan, & Hillyard, 1995). The scalp sites were as follows: frontal/central, Fp1, Fp2, FC1, FC2, FC5, FC6, F7, F3, Fz, F4, F8; temporal/central, T3, T5, CT5, C1, C3, C5, Cz, C2, C4, C6, CT6, T6, T7; central/parietal, CP1, CP2, P3, Pz, P4; temporal/parietal/occipital, TO1, PO1, TO2, PO2, O1, O2, IPz, INz, IN3, IN4, IN5, IN6; left mastoid (A1); and right mastoid (A2). The horizontal EOG (HEOG) was a bipolar montage between the outer canthi of the eyes and the vertical EOG (VEOG) electrode was placed under the left eye. All electrodes were initially referenced to the nose and subsequently rereferenced off line to the averaged mastoids. EEG and EOG signals were amplified with a bandpass of 0.1–100 Hz (-6 -dB points) and digitized at 250 Hz.

Prior to averaging, a battery of artifact rejection algorithms was applied to eliminate artifacts due to blinking, muscle activity, and amplifier blocking. In a first pass, all electrode sites were tested for blocking, and amplitude thresholds were set to reject voltage deviations exceeding three times the typical amplitude level of ongoing EEG. In addition, selected electrodes were tested with more specific algorithms for the detection of high-frequency activity (e.g., muscle activity) at frontal, temporal, and mastoid sites, alpha rhythm (at occipital sites), and eye blinks or slow shifts of gaze evident in the vertical or horizontal EOG. After visually assessing the rejected trials, the rejection thresholds were iteratively adjusted until artifact contamination was no longer evident in the averaged waveforms. Since an effect of eye position on ERPs during auditory selective attention has been reported (Okita & Wei, 1993), special efforts were made to exclude trials with detectable ocular deflections (roughly 20 – 30 μ V, corresponding to 2° – 3°) in the horizontal EOG. The typical number of trials included in an average was 900–1,200 trials for standard stimuli and 120–160 trials for deviant stimuli. Usually, subjects with more than 25% rejections are discarded. However, none of the subjects met this criterion. After averaging, the data were digitally low-pass-filtered with a gaussian finite impulse function, yielding a -6 -dB point (50% down) at about 40 Hz.

Given the short SOAs between successive stimuli, there was undoubtedly some overlap from ERPs to preceding stimuli included in the averaged waveforms that were analyzed. However, we did not consider it necessary to apply methods for reducing this overlap, such as ADJAR (Woldorff, 1993), because the overlap would similarly affect ERPs to stimuli from all speaker positions in the array and thus would not distort the assessment of spatial gradients.

Procedure

The experiment was carried out in an electrically and acoustically shielded room. There were three experimental conditions, with 10 runs of 1,092 stimuli (duration of each run = 3.27 min) in each condition. In counterbalanced order, the subjects' task was to attend to either the leftmost speaker (*attend left*), the center speaker (*attend center*), or the rightmost speaker (*attend right*). In all three conditions, the subjects were asked to press a button to the infrequent higher pitched noise bursts (targets) appearing at the designated location and to ignore stimuli coming from adjacent locations. Several practice runs preceded the experiment to ensure that the subjects could perform the task. Additionally, the subjects were asked to avoid excessive blinking and extraneous head movements during the runs and were made familiar with the feedback provided by the active fixation point.

Behavioral Data

The subjects' performance was evaluated by measuring the percentage of correctly detected targets in the attended channel and the percentage of deviant stimuli incorrectly responded to in adjacent channels. Due to the very fast delivery rate of about 6 stimuli per second, not all of the recorded buttonpresses could be unambiguously assigned to a particular target/deviant. We defined a correct target detection response as a buttonpress occurring after a target stimulus within a window of 200–800 msec. We employed the same algorithm for the classification of detection responses following the nontarget deviant stimuli presented at adjacent speaker positions. The distributions of correct detections and incorrect responses to neighboring locations and the RTs of these responses were subjected to one-way analysis of variance (ANOVAs; with a factor location), including Tukey–Kramer posttests for comparisons between adjacent locations. Selected comparisons of detection response rates between the three experimental conditions were made with *t* tests.

ERP Data

To examine attentional gradients of ERPs, amplitude measures of N1/Nd to the standard tones, N1 to the target/deviant tones, and P3 to the target/deviant tones were taken for all seven speakers under each experimental condition. The amplitudes of the waveforms were calculated with respect to the mean voltage over a 200-msec baseline preceding stimulus onset. The N1/Nd attention effects were quantified both in terms of N1 amplitudes elicited by standard tones at each speaker and as the Nd difference waves. For the attend-left condition, the Nd for each speaker was calculated as the ERP to stimuli from that speaker (S1–S4) when S1 was attended minus the ERP to the same speaker when S7 was attended. The Nds for the attend-right condition were calculated in a similar fashion, as the ERPs to speakers S4–S7 when S7 was attended minus the ERPs to the same stimuli when speaker S1 on the opposite side of the array was attended. For the attend-center condition, the Nds for speakers S2–S6 were calculated as the ERP to each speaker when the central speaker S4 was attended minus the ERP to the same speakers averaged over the conditions when S1 and S7 were attended.

Because the N1 and Nd measures yielded equivalent attentional gradients (see Results section) and because the N1 measure could be more uniquely associated with the subjects' behavioral responses, the N1 measure was preferred in this analysis.²

For the N1 and for the P3, respectively, the mean amplitude during 120–200 msec (site Cz) and that during 350–550 msec (site Pz) were used to assess attentional gradients. These measurement intervals encompassed the major attention-related changes in each component. Attention effects on ERP components were evaluated either by one-way ANOVAs, including Tukey–Kramer posttests, or by repeated measures ANOVAs. Selected comparisons of ERP amplitudes at different electrode sites were made with paired *t* tests.

RESULTS

Behavioral Data

In all three experimental conditions, the distribution of detection responses to deviant stimuli at the designated (attended) and adjacent spatial locations indicated a relatively steep falloff. In Figure 2, the percentages of correct detections during a window of 200–800 msec following the deviants at the designated location and the percentages of responses made incorrectly to deviants in neighboring speakers are shown for each condition. The percentage of targets/deviants responded to at each location is termed the *detection response rate*. The labeling of the speaker positions in Figure 2 and throughout the text corresponds to the labels shown in Figure 1.

In the attend-left condition, the falloff of detection responses with increasing distance from the designated location was highly significant [$F(3,33) = 111.5, p < .0001$] and was mainly due to significant differences between speaker positions S1 and S2 ($q = 17.96, p < .001$) and between S2 and S3 ($q = 4.03, p < .05$); the difference between positions S3 and S4 was not significant. In the attend-right condition, the falloff was also highly significant [$F(3,33) = 123.0, p < .0001$], mainly due to differences between speaker positions S6 and S7 ($q = 18.41, p < .001$), and S5 and S6 ($q = 4.84, p < .01$), whereas responses given to positions S4 and S5 did not differ significantly from one another. The falloff curves of detection responses across the four adjacent sound sources were very similar for the attend-left and attend-right conditions; these two curves were correlated with $r = .99 (p < .0002)$.

For the attend-center condition, a significant falloff of responses was confirmed for both the left half-field speakers S2, S3, and S4 [$F(2,33) = 132.40, p < .0001$] and the right half-field speakers S4, S5, and S6 [$F(2,33) = 130.20, p < .0001$]. The difference between locations S3 and S4 ($q = 19.45, p < .001$) for the left half-field and the falloff between locations S4 and S5 ($q = 19.64, p < .001$) for the right half-field were mainly responsible for the highly significant main effects. The number of detection responses given to positions S2 versus S3 and to S5 versus S6 did not differ significantly.

In order to compare the steepness of lateral and central falloffs in detection response rate, two-tailed *t* tests were performed on the normalized proportional falloffs between the target locations and their nearest neighbors. Since the falloffs between the attend-left and attend-right conditions were similar, the data were collapsed across the two conditions. Furthermore, since the proportional falloffs

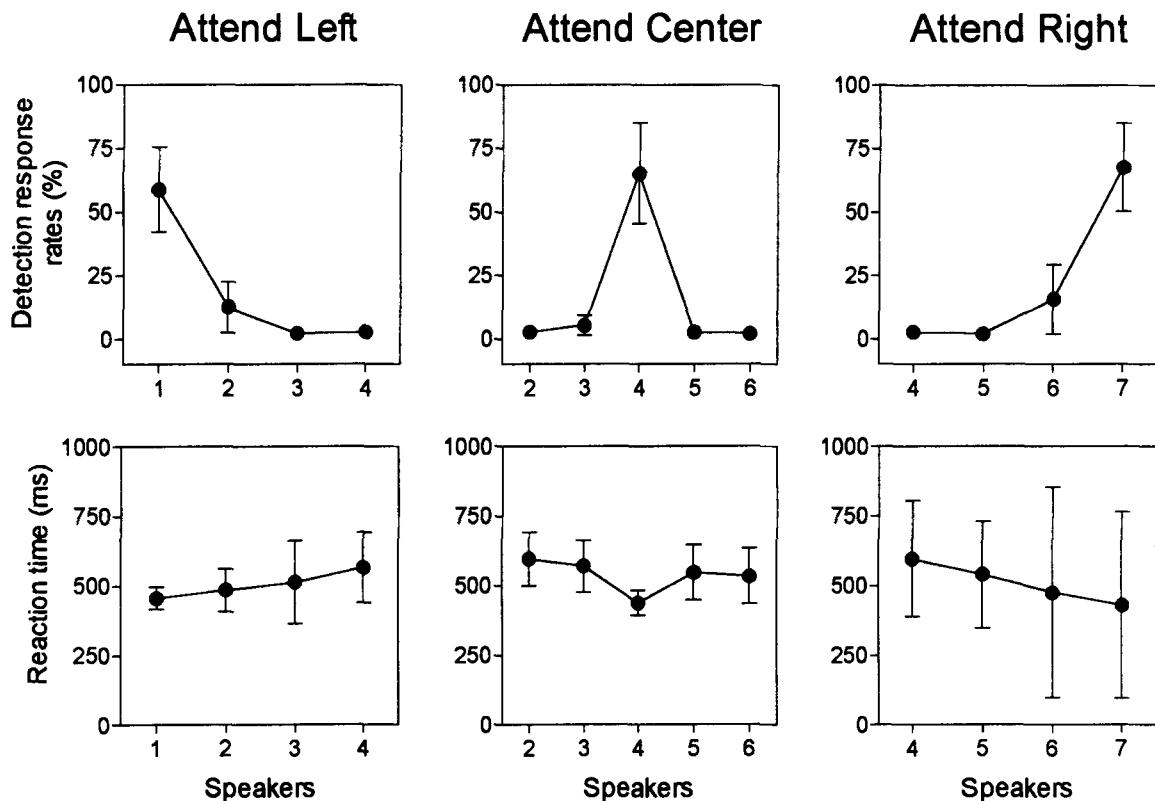


Figure 2. The distribution of behavioral detection responses (hits to attended locations, false alarms to neighboring sources) made during a window of 200–800 msec after target/deviant stimuli at each location under the three different experimental conditions (error bars show standard errors of the mean, SEMs). Top row: Distribution of detection response rates. Bottom row: Distribution of reaction times (RTs).

for the left and right half-fields in the attend-center condition were similar, the response rates of locations S3 and S5 in relation to S4 were averaged together to quantify the central falloff. The central falloff to the nearest neighbors (on the average, 93.8%) was significantly steeper than the averaged lateral falloff (79.0%) [$t(11) = 3.85, p < .003$].

Figure 2 also shows the RTs for buttonpresses to the designated locations and neighboring locations. In all three conditions, the shortest RTs were for the correct detections at the designated target locations (i.e., to S1, S4, and S7), with a gradual increase in RTs with distance. These effects failed to reach significance for lateral target conditions; however in the attend-center condition, the RTs to stimuli at positions S2 and S3 were significantly longer relative to the RTs to stimuli at position S4 [$F(2,33) = 12.99, p < .0001$], as were the RTs to positions S5 and S6 [$F(2,33) = 5.91, p < .006$].

ERP Data

In all three experimental conditions, the ERPs to the attended locations had larger negative amplitudes in the N1 latency range (100–200 msec poststimulus onset) than the ERPs to the same locations when unattended. Figure 3 shows the distribution of attended and unattended

ERPs to the frequent standard stimuli coming from speaker locations S1, S4, and S7 in the three different attention conditions. The attended N1 amplitudes to the standard stimuli (N1) showed a broad distribution across the scalp.

Figure 4 shows ERPs and the associated negative difference waves (Nd)s elicited by frequent standard stimuli at each speaker location in the three experimental conditions (site Cz). The Nd showed a biphasic morphology, with the early Nd peak (range 160–200 msec) having a similar amplitude and latency to that of the N1 to standards.³

Both the N1 and the Nd deflections were largest in amplitude to stimuli occurring at the attended locations, with a marked falloff across the neighboring sound sources.⁴ The late Nd, peaking at 350–400 msec, however, was actually largest to stimuli coming from the nearest neighbor locations, as discussed below.⁵

Figure 5 shows the ERPs to infrequent deviant stimuli (noise bursts with an increased bandwidth) at site Cz. The top row shows the N1 waves elicited by deviants at all speaker locations under the three experimental conditions. This N1 was largest to deviants at the attended locations or its immediate neighbor but also had substantial amplitude at more distant locations. In all three

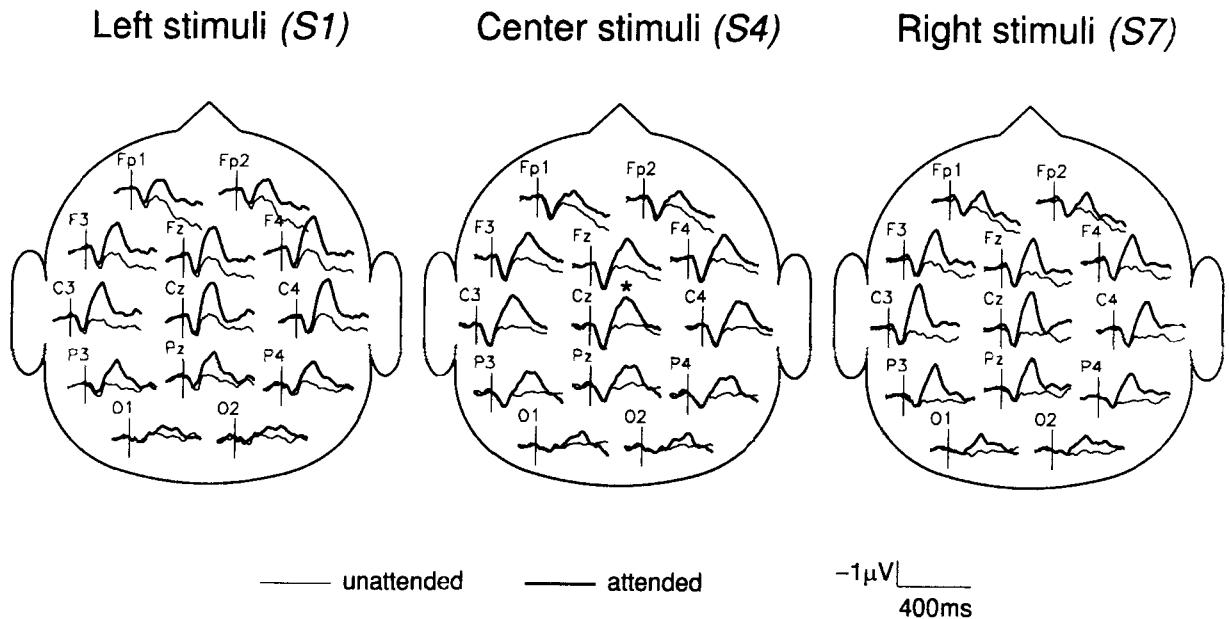


Figure 3. The distribution of grand average ERPs to the frequent standard sounds for each of the three attention conditions (for simplicity, only 13 channels out of 44 are shown). In each condition, ERPs elicited by to-be-attended tones were negatively displaced (N1/Nd component) relative to the ERPs to the same tones when unattended (i.e., when a distant speaker was attended). The N1 peak at site Cz is marked with an asterisk.

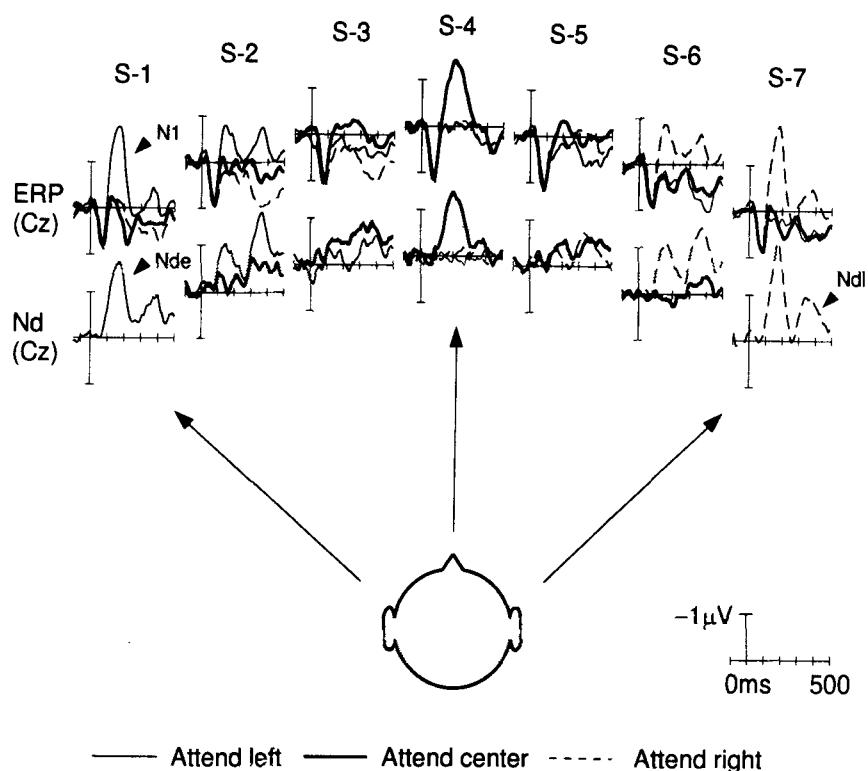


Figure 4. Top row: Grand average ERPs elicited at the vertex (Cz) by frequent standard stimuli from all speaker positions in three experimental conditions. Bottom row: Negative difference waves (Nd) obtained by subtracting the unattended ERPs from the attended ERPs.

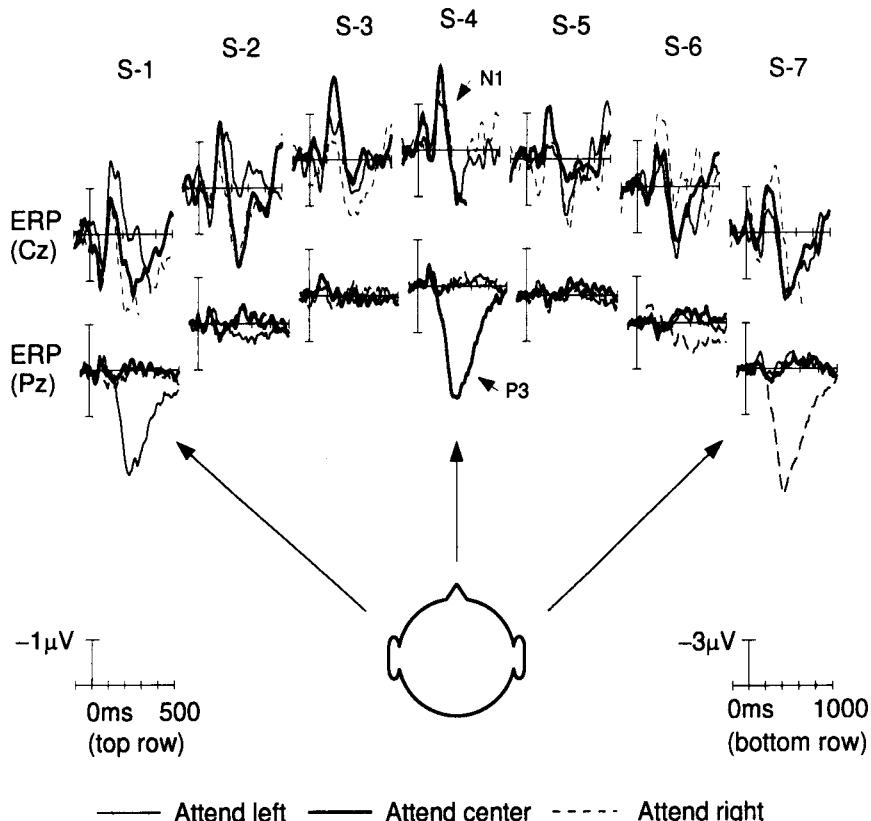


Figure 5. Top row: Grand average ERPs elicited at the vertex (Cz) by infrequent target/deviant stimuli from all speaker positions in the three experimental conditions. The P3s in the top waveforms were truncated at 350 msec so as not to overlap the lower tracings. Bottom row: Grand average ERPs (site Pz) elicited by infrequent target/deviant stimuli across all speaker positions in three different experimental conditions.

experimental conditions, the amplitudes of the N1s to standards and to deviants at the attended location did not differ significantly. The peak latencies of the N1 to deviants were significantly earlier than the peaks of the N1s elicited by standard tones, however, in the lateral (by 13 msec) [$t(11) = 4.02, p < .002$] and central conditions (by 60 msec) [$t(11) = 6.65, p < .0001$].

The lower tracings in Figure 5 show that the P3 component peaking at 350–550 msec was elicited mainly by target stimuli appearing at the designated locations (S1, S4, or S7), with a maximum amplitude at site Pz. Physically identical deviant stimuli occurring at neighboring speaker locations (nontarget deviants) elicited hardly any positive deflection within the P3 latency range. Neither the P3 peak latencies nor the P3 mean amplitudes to attended targets showed significant differences across conditions.

Comparison of Spatial Gradients of ERP Amplitudes and Detection Response Rates

The spatial gradients of ERP amplitudes within the N1, Nd, and P3 latency ranges in each experimental condition are depicted in Figure 6. The top row illustrates

the N1 and Nd gradient to standards and the N1 gradient to deviants. Due to the fact that the ERPs elicited by tones coming from unattended spatial positions were very small in amplitude, the evoked N1 amplitudes were very similar to those of the Nd obtained by subtracting the unattended from the attended ERPs to the same stimuli. The correlations between the N1 and the Nd measures to standard stimuli were .99 ($p = .009$) in the attend-left condition, .99 ($p = .009$) in the attend-center condition, and .98 ($p = .02$) in the attend-right condition. Given this similarity and for reasons described in Method section, the N1 measures were preferred for defining the attentional gradients.

In the attend-left condition, the falloff of the N1 amplitude to standards across speaker locations was significant [$F(6,11) = 33.7, p < .0001$], mainly due to a significant difference between positions S1 and S2 ($q = 7.3, p < .001$) and between positions S2 and S3 ($q = 7.3, p < .001$). In the attend-center condition, the falloff was also significant [$F(6,11) = 11.3, p < .0001$], mainly because of differences between positions S3 and S4 ($q = 5.3, p < .01$) and between positions S4 and S5 ($q = 6.6, p < .001$). Similar results were obtained for the attend-right condi-

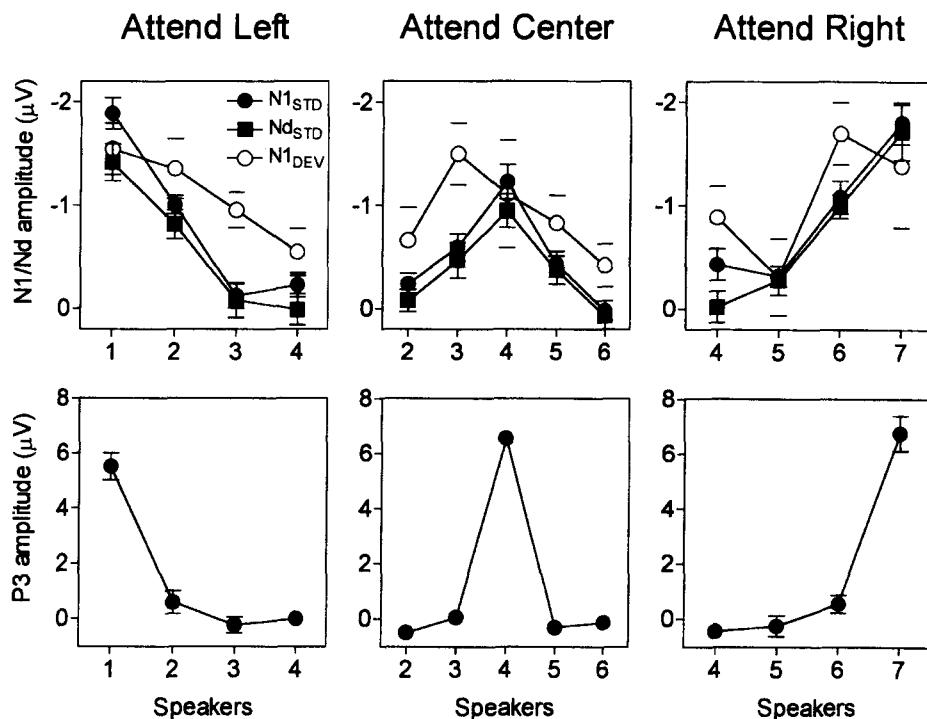


Figure 6. Grand average mean amplitudes of N1 and Nd to standards and N1 to deviants (top row; site Cz, window 120–200 msec) and P3 (bottom row; site Pz, window 350–550 msec) in three different conditions across attended and neighboring speaker positions. The error bars in the P3 graph for the attend-center condition do not exceed the size of the dots.

tion: A significant falloff [$F(6,11) = 21.9, p < .0001$] was mainly due to significant differences between positions S6 and S7 ($q = 11.7, p < .001$) and between positions S5 and S6 ($q = 6.0, p < .01$).

For the N1 to the deviants, the spatial gradients were shallower than for the N1 to standards. The overall deviant–N1 amplitude falloff was significant in the attend-left condition [$F(3,11) = 3.19, p < .04$], mainly due to the significant difference between S1 and S4 ($q = 4.02, p < .05$). A similar analysis of the attend-center and attend-right condition revealed no reliable falloff of the deviant–N1 amplitudes with increasing distance from the attended locations.

The bottom row in Figure 6 illustrates the corresponding P3 gradients. In the attend-left condition, the falloff (in microvolts) was significant [$F(6,11) = 36.6, p < .0001$], mainly because of the significant difference between positions S1 and S2 ($q = 13.63, p < .001$). No significant differences were obtained between any other speaker locations. In the attend-center condition, the substantial amplitude differences between positions S3 and S4 ($q = 16.8, p < .001$) and between positions S4 and S5 ($q = 17.8, p < .001$) accounted for an overall significant falloff [$F(6,11) = 45.3, p < .0001$] of the P3 amplitude. In the attend-right condition, the overall falloff was also significant [$F(6,11) = 54.4, p < .0001$], almost entirely due to the substantial amplitude difference between positions S6 and S7 ($q = 17.5, p < .001$).

Figure 7 shows normalized N1 and P3 mean amplitudes and detection response rates across the speaker array in the three experimental conditions. Across all conditions, the mean N1 amplitudes to standards were strongly correlated ($r = .89, p < .0001$) with the detection response rates. The correlation between mean P3 amplitudes and detection response rates was even greater ($r = .99, p < .0001$), however, and the difference between these two correlations was highly significant [$t(9) = 6.3, p < .0001$], using a test developed by Williams (1959, cited and modified by Howell, 1992, p.253f). The amplitudes of the N1 to standards and deviants were correlated with $r = .80 (p < .001)$ across conditions.

Gradient Differences for Central and Lateral Conditions

Since the falloffs of N1 and P3 amplitudes and detection response rates did not differ substantially between attend-left and attend-right conditions, left and right falloff data were collapsed into attend-lateral measures. Similarly, in the attend-center condition, there were no significant differences between the falloff of ERP amplitudes and detection response rates between the left and right flanking speaker. Therefore, data from left and right half-fields were collapsed. These values were then normalized with respect to the attended values, which were set to 100% (Figure 8). The differences in falloff between the attended (ATT) location values and those of

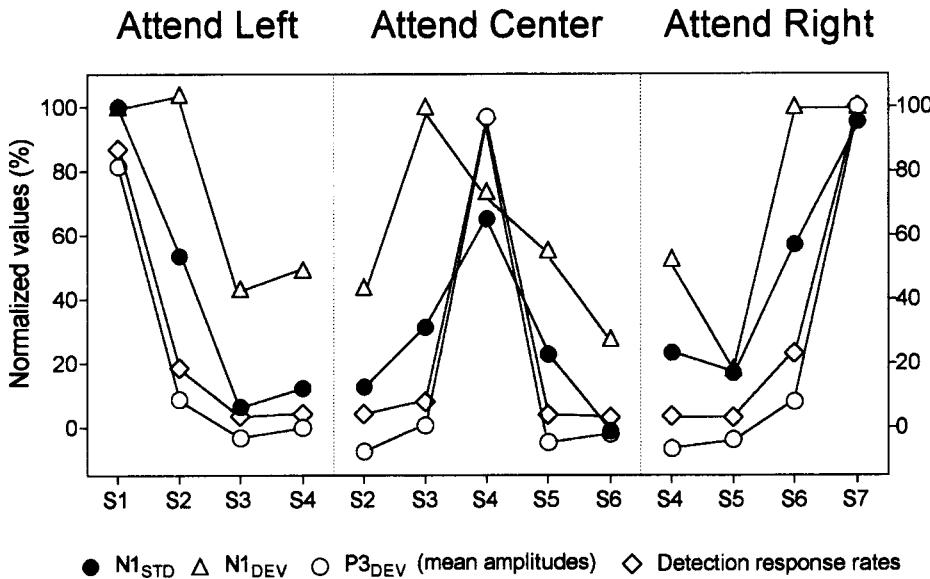


Figure 7. Distribution of normalized N1 and P3 mean amplitudes and detection response rates (percent correct hits for attended locations, false alarms to the nearest neighbors) across the speaker array for each experimental condition.

the nearest neighbor (NN) and second-nearest neighbor (SN) locations for the various measures were tested with *t* tests (paired, two-tailed).

For the attend-lateral conditions, the falloff of N1 amplitudes to standards was highly significant [$F(2,22) = 54.74, p < .0001$], with significant differences between the ATT and NN locations ($q = 7.55, p < .001$) and between the NN and SN locations ($q = 7.24, p < .001$). In contrast, for the N1 amplitudes to deviants, the overall falloff showed only a marginal trend [$F(2,21) = 3.26, p = .06$]. The falloff of P3 was highly significant [$F(2,22) = 127.30, p < .0001$], mainly due to a significant difference between the ATT and NN locations ($q = 18.10, p < .001$). The amplitude difference between the NN and SN locations did not reach significance. (Table 1).

For the attend-lateral conditions, the falloff of P3 between ATT and NN locations was markedly steeper than the falloff of the N1 amplitude to standards [$t(11) = 7.10, p < .0001$]. The detection response rates showed also a steeper falloff than the N1 amplitude [$t(11) = 3.59, p < .005$]. The P3 and detection response rates differed only slightly in their falloffs, but this difference was significant [$t(11) = 3.56, p < .005$]. However, as is evident in Figure 8, the difference in falloff between detection response rates and N1 to standards was significantly greater than the difference in falloff between detection response rates and P3 [$t(11) = 7.101, p < .0001$].

For the attend-center condition, the significant falloff of normalized N1 amplitudes [$F(2,22) = 22.89, p < .0001$] was mainly due to the difference between the ATT and NN locations ($q = 6.05, p < .001$), whereas the small difference between NN and SN locations did not reach significance. Similar results were obtained for the falloff of the P3 in the

central condition [$F(2,22) = 65.40, p < .001$], in which only the difference between the ATT and NN locations reached significance ($q = 13.81, p < .001$). The falloff of N1 amplitudes to deviants did not reach significance.

With centrally directed attention, the falloff of P3 between ATT and NN locations was again much steeper than that of the N1 to standards [$t(11) = 7.44, p < .0001$], and the detection response rate also dropped off more than the N1 amplitude [$t(11) = 5.36, p < .0002$]. Similarly to in the lateral attention conditions, the small difference in falloff between P3 and detection response rate was significant [$t(11) = 2.31, p < .04$], but, again, the difference between detection response rates and N1 was considerably larger than the difference between detection response rates and P3 [$t(11) = 7.44, p < .0001$].

The falloff of N1 to standards (ATT-NN) was much steeper than the falloff of the N1 to deviants in both attend-center [$t(11) = 5.78, p < .0001$] and attend-lateral [$t(11) = 3.02, p < .02$] conditions.

As for differences between lateral and central conditions, the N1 amplitude falloff between ATT and NN

Table 1
N1 and P3 Difference Amplitudes (in Microvolts;
Mean Amplitudes; Site Cz) Across Adjacent Spatial Positions

Wave	Condition	ATT-NN (&V)	NN-SN
N1	Attend left	-1.10‡	-0.62‡
	Attend center	-1.00‡	-0.34*
	Attend right	-0.80‡	-0.78†
P3	Attend left	6.00‡	0.73 (n.s.)
	Attend center	7.50‡	-0.04 (n.s.)
	Attend right	6.80‡	1.50 (n.s.)‡

Note—ATT, attended location; NN, nearest neighbor; SN, second-nearest neighbor; n.s., nonsignificant. * $p < .05$. † $p < .01$. ‡ $p < .001$.

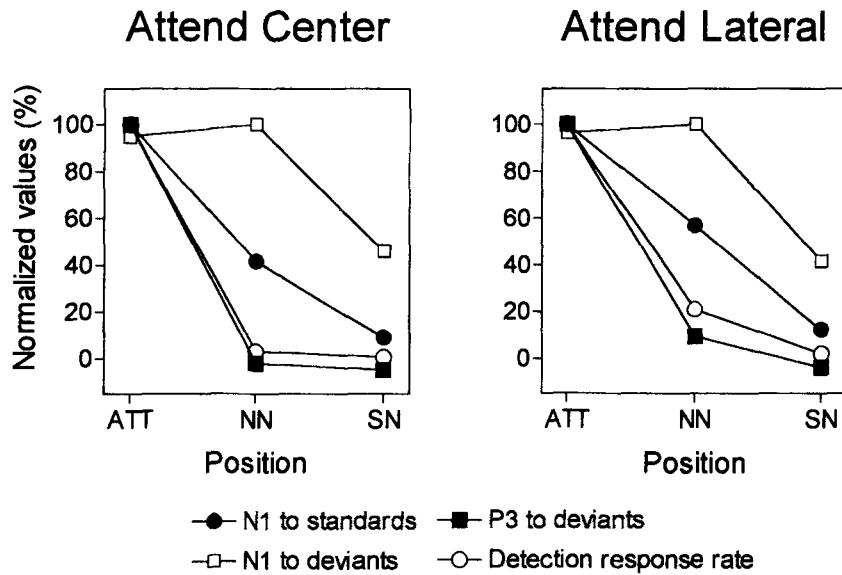


Figure 8. A comparison between falloffs of the normalized N1 and P3 mean amplitudes and detection response rates for attend-center and attend-lateral conditions (collapsed across left- and right-field speaker positions). ATT = attended location; NN = nearest neighbor; SN = second-nearest neighbor.

locations was marginally steeper for centrally directed attention [$t(11) = 2.15, p < .06$]. The falloff of the P3 amplitudes, however, was significantly steeper for central conditions than for lateral conditions [$t(11) = 2.65, p < .02$], as was the falloff for detection response rate [$t(11) = 3.85, p < .003$].

The Late Nd

Figure 9 shows ERPs and the biphasic morphology of the Nd to standard stimuli for all three experimental conditions for the attended and adjacent locations. The late Nd mean amplitudes over the interval 300–420 msec were larger for stimuli at NN speaker positions than at either ATT [$t(11) = 2.78, p < .01$] or SN [$t(11) = 8.23, p < .0001$] positions across all three experimental conditions. Similar effects could not be determined for the late Nd to deviants due to the superimposition of the P3.

Time Course Analysis

To examine the time course of the ERP selective attention effects, N1 amplitudes to standards at consecutive 30-msec intervals and P3 amplitudes to deviants at consecutive 40-msec intervals were plotted and statistically analyzed for attended locations. This analysis was not carried out for the N1 to deviants due to the relatively low signal-to-noise ratio of those waveforms and their nearest neighbors (Figure 10).

For the N1 component, the point of maximal selectivity (attended vs. unattended amplitude) was reached at the peak latency range (about 160 msec for the lateral stimuli, and 200 msec for the central stimuli) and rapidly declined thereafter. It can also be seen that stimuli at the NN locations received substantial processing relative to

those at SN locations. In the attend-left condition, the differences ATT–SN and ATT–NN became significant during the interval 70–100 msec [$t(11) = 6.80, p < .0001$, and $t(11) = 2.92, p < .014$, respectively]. In the attend-center condition, the differences ATT–SN and ATT–NN became significant during the interval 130–160 msec [$t(11) = 8.60, p < .0001$, and $t(11) = 4.57, p < .001$, respectively]. In the attend-right condition, the difference ATT–SN became significant during the interval 70–100 msec [$t(11) = 3.73, p < .004$], and the difference ATT–NN became significant during the time window 130–160 msec [$t(11) = 3.37, p < .006$].

The bottom row in Figure 10 shows the time course of the P3 amplitudes, which were much larger for stimuli at the target locations relative to stimuli presented at adjacent positions. In the attend-left condition, the differences ATT–SN and ATT–NN both became significant during the interval 280–320 msec [$t(11) = 2.33, p < .04$, and $t(11) = 2.40, p < .04$, respectively]. In the attend-center condition, the difference ATT–SN became significant during 280–320 msec [$t(11) = 4.79, p < .001$] and the difference ATT–NN reached significance earlier during 200–240 msec [$t(11) = 2.42, p < .04$]. In the attend-right condition, both differences became significant during the same time interval (240–280 msec) [ATT–SN, $t(11) = 2.79, p < .02$; ATT–NN, $t(11) = 5.04, p < .001$].

DISCUSSION

Under the free-field conditions of the present study, deviant sounds occurring at the attended location (i.e., targets) were responded to most frequently, with a progressive decline in the proportion of detection responses

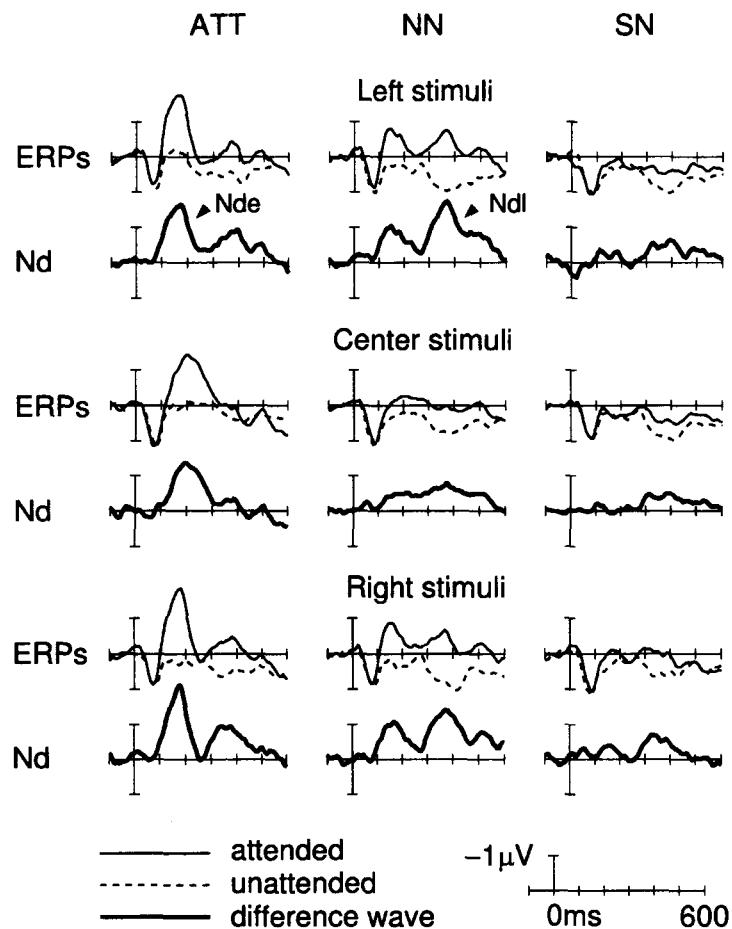


Figure 9. Grand average ERPs (site Cz) and attend minus unattend difference waves (Nd) obtained in the three conditions for target (to-be-attended) locations and the two adjacent speaker locations. ATT = attended location; NN = nearest neighbor; SN = second-nearest neighbor. The arrows indicate the early and late Nd peaks (Nd_e , Nd_l).

made to deviants at sources increasingly distant from the attended speaker. The principal ERP components triggered by both standard and deviant stimuli were similarly distributed as gradients across the array of speakers, with maximal amplitudes typically elicited by sounds at the attended location. Thus, both the behavioral and the ERP data were in accord with Mondor and Zatorre's (1995) proposal that auditory attention to a sound source in free field is distributed as a gradient that tapers off as a function of distance from the attended point. In the present experiment, however, discriminative RT was not a particularly sensitive index of attentional allocation to location. This may have been due to our lack of emphasis on response speed in instructing the subjects, thereby making it likely that RT was determined principally by factors other than the allocation of attention.

The gradients of detection performance and ERPs around the attended location in the present study were considerably steeper than the gradients of RT observed

by Mondor and Zatorre (1995). This difference might be attributed to the substantial differences in stimuli and task design between the two studies. In the present study, the subjects had to respond only to stimuli coming from a single attended location within a closely spaced (9° separation) loudspeaker array, with stimuli presented at a rapid rate (5–6 per second). Since the spatial cues associated with the attended and unattended locations were repeated at short intervals in a continuous sequence, the subjects were able to use those cues effectively to sustain attention rather narrowly on the designated sound source. In contrast, Mondor and Zatorre used a trial-by-trial cuing design in which subjects had to respond to sounds coming from any of several possible locations that were more widely spaced in location (30° or 45° separation between speakers) and in time. This requirement to respond to all locations (even though sounds were most likely to occur at the cued location) and the long intervals between stimuli would likely be conducive to the deployment of a

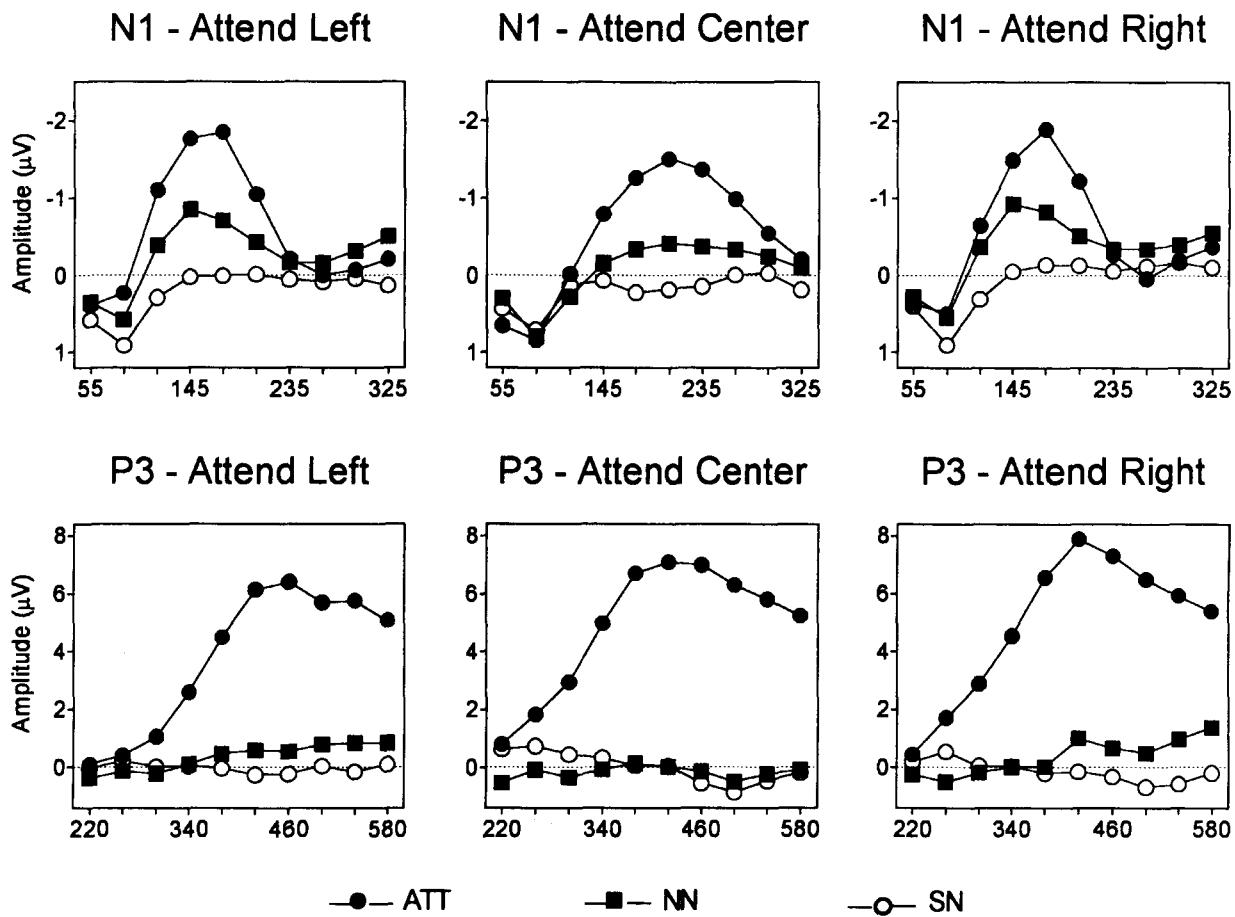


Figure 10. The time course of the N1 to standard stimuli and the P3 in the three attention conditions. The curves represent consecutive mean amplitude measurements spanning a range of 40–250 msec in 30-msec windows (for N1) and a range of 200–600 msec in 40-msec windows (for P3). ATT = attended location; NN = nearest neighbor; SN = second-nearest neighbor.

broader gradient of attention around the designated location than in the present study. Finally, the requirement for a deviant-standard discrimination (rather than a simple RT response) in the present study may have produced a narrower focusing of attention.

The gradients of detection responses and ERPs observed here were steeper around attended central stimuli than around attended peripheral stimuli. This probably is due to a more refined capability for discriminating between adjacent sound sources centrally than in the periphery. On the basis of earlier studies in cats (Jenkins & Masterton, 1982), it was found that humans can discriminate adjacent sound sources directly in front of them when separated by only 2° – 3° , whereas discrimination between peripheral sounds (e.g., at 40° azimuth) was markedly inferior (Perrott, Constantino, & Cisneros, 1993; see also Oldfield & Parker, 1984). Clearly, the sharpness of attentional allocation to a particular sound source among nearby competing sources depends in part on the ability to discriminate the location of the attended source from those of its neighbors.

The early N1 component elicited by the infrequent target/deviant stimuli showed a shallow gradient distribution around the attended location. The N1 to the frequent standard stimuli showed a steeper gradient distribution around the attended location, but it was not so narrowly focused as the gradient of behavioral detection responses. Considerable evidence supports the hypothesis that the N1/Nd represents stimulus-evoked neural activity in auditory cortex that has passed through an early attentional filter (Hillyard et al., 1995; Woldorff et al., 1993). Accordingly, we propose that N1/Nd amplitude variations across speaker locations may represent the gradient of spatial tuning of this early attentional filter. Previous studies of auditory attention have shown that the steepness of spatial gradients as indexed by the N1/Nd can vary considerably depending on stimulus and task factors. For example, Hink et al. (1978) observed much shallower gradients of N1 amplitude across their five sound source locations than in the present study, which is most likely attributable to their use of more widely spaced sound locations (presented through headphones rather than in

free field), a much slower rate of stimulus presentation, and a less demanding discrimination task. On the other hand, Teder and Näätänen (1994) reported a very fine tuning of auditory spatial attention in a task in which subjects had to listen to one of two competing streams of natural speech presented from closely spaced speakers in free field. Under these conditions, subjects were able to focus attention selectively on one of the messages at speaker separations of only 3° of angle as indexed by N1/Nd amplitudes elicited by probe stimuli superimposed on each message and matched to them in pitch. Possibly, the selective focusing of attention was enhanced in this case because the speech messages could be distinguished by pitch, rhythm, and amplitude modulation cues, as well as by location (see Teder, 1994).

The finding of steeper attentional N1 gradients for the standards than for the deviants in the present study indicates that the selection mechanism reflected in the N1/Nd is not based on stimulus location alone. This difference in tuning may be explained with reference to the "attentional trace" model proposed by Näätänen (1982), wherein the early selection process depends on a comparison of auditory inputs against a representation (trace) of the attended stimulus attributes held in short-term memory. To the extent that a stimulus matches the trace, it is accepted for further processing, which is indexed by an increased N1/Nd amplitude. In the present experiment, if we assume that the attentional trace includes information about the pitch (bandwidth) of the attended targets and their locations, it would be expected that the deviant stimuli sharing the pitch attribute with the targets would elicit a relatively larger N1 than would standards at corresponding nonattended locations. Such a situation could well result in the observed shallower N1/Nd gradients for the deviants at the non-attended locations. A similar pattern would result if separate attentional traces or filters were maintained, one for the attended location and the other for the deviant frequency of the target.

Other possible mechanisms could also be suggested to account for the broader spatial tuning of the N1 to the deviants. For example, it might be the case that the initial selection is based primarily on the location (or pitch-location) attributes of the attended standards, since those sounds are presented much more frequently than the deviants and hence have a more strongly reinforced attentional trace. In this case, the broader N1 gradients for the deviants may result from the activation of additional N1 subcomponents specific to those mismatching stimuli. It has been pointed out that the N1 consists of multiple subcomponents (Näätänen & Picton, 1987), including a "mismatch negativity" (MMN) that is a specific index of the mismatch between a deviant stimulus and the memory trace of the repeating standard. Also, a deviant stimulus typically activates additional N1 generators that are not as refractory as those activated by the standards. It may be the case then that these mismatch-evoked components to the deviants are less susceptible to modula-

tion by the attentional trace than is the N1 to the more frequent standard sounds.⁶

The amplitudes of the P3 to deviant stimuli were also maximal for attended-location stimuli and were distributed as a gradient around that location. The P3 gradient was significantly steeper than the earlier N1/Nd gradients, however, and its slope corresponded almost perfectly to that of the gradient of behavioral detection response rates. This suggests that sounds from the different competing sources were selected in two distinct stages, with an initial, less precise selection by location at the level of the auditory cortex (indexed by N1/Nd) being followed by a more sharply tuned selection (indexed by P3) based on further analyses of stimulus properties that also form the basis for discriminative target detection performance. A consideration of the time course of the N1/Nd and P3 attention effects reinforces this proposal of two distinct levels of spatial selectivity. The early N1/Nd index of selectivity increased progressively over the interval 80–180 msec, reaching a maximum at the peak of the N1 and declining thereafter. The more spatially discrete P3 index began at around 250 msec and reached a maximum over the period 400–500 msec, just prior to the emission of the behavioral detection responses. These results are congruent with studies of visuospatial attention showing that early sensory evoked components (P1 and N1) show relatively broad gradients to stimuli around an attended location, whereas the later (N2, P3) endogenous components reflect a more sharply tuned selection process (Mangun & Hillyard, 1988; Mulder et al., 1994).

The present results are consistent with gradient models of spatial attention (e.g., Downing, 1988; LaBerge & Brown, 1989; Yund et al., 1990), which specify that stimuli are selectively processed to an extent that varies with their distance from the attended location. These data further affirm that attention is allocated as a gradient in auditory and visual modalities and in sustained attention tasks and in trial-by-trial cuing tasks (Hink et al., 1978; Mondor & Zatorre, 1995). However, there is still some uncertainty about the nature of the stimulus selection processes underlying the gradients. One possibility is that the ERP and behavioral gradients are determined by attentional filters or allocation mechanisms that have sloping spatial tuning functions that are maintained over multiple stimulus presentations (e.g., LaBerge & Brown, 1989). An alternative possibility, however, is that the attentional "spotlight" has a fixed diameter with a sharp border but moves around from one stimulus presentation to the next; in such a case, the gradient function would be determined by the proportion of time that attention is centered on the various locations (e.g., Eriksen & Hoffman, 1972). Since both the behavioral and the ERP gradients here were based on data averaged over many trials, it is difficult to distinguish between these alternative mechanisms. In any case, it is clear that the steepness of the slopes of the attentional gradients differ sharply between early and late processing stages.

The late phase of the Nd component (300–400 msec) was prominent in response to standard stimuli at the attended location, but its amplitude was actually largest for the standard sounds at locations immediately adjacent to the attended speaker. Evidence from previous studies has suggested that the late Nd is an index of further processing of more detailed stimulus features (Hansen & Hillyard, 1983; Okita, 1979), and its amplitude is reportedly diminished when stimuli are presented at rapid rates that preclude detailed processing of each individual stimulus (Näätänen, 1992; Teder, Alho, Reinikainen, & Näätänen, 1993). In the present study, however, substantial late Nds were elicited despite the rapid stimulus delivery rate of 5–6/second. A possible explanation is that the necessity for fine spatial tuning of attention among closely spaced sound sources requires more extended processing of stimuli arising from sources adjacent to the attended location in order to verify their relative positions.

In sum, the combined behavioral and ERP data of the present study indicate that auditory spatial attention is deployed as a finely tuned gradient around an attended sound source in a free-field array of closely spaced sources simulating a cocktail party. These data provide additional support for gradient models of spatial attention and, together with evidence from prior studies, make it clear that the slopes of additional gradients may vary widely according to stimulus and task factors. The ERP data suggest that the spatial focusing of attention is achieved in two distinct stages, with an early, more broadly tuned filtering of inputs occurring over the first 80–180 msec after stimulus onset (indexed by N1/Nd), followed by a more narrowly focused selection of attended-location deviants that begins at around 250 msec (indexed by P3 amplitude) and closely resembles the shape of the behavioral gradient of target detections. By combining behavioral and electrophysiological measures in this way, it was possible to observe both spatial and temporal properties of the fine tuning of auditory attention to location.

REFERENCES

- CHEAL, M., LYON, D. R., & GOTTLÖB, L. R. (1994). A framework for understanding the allocation of attention in location-precued discrimination. *Quarterly Journal of Experimental Psychology*, **47A**, 699–739.
- CHERRY, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *Journal of the Acoustical Society of America*, **25**, 975–979.
- CLARK, V. P., FAN, S., & HILLYARD, S. A. (1995). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping*, **2**, 170–187.
- DOWNING, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 188–202.
- DOWNING, C. J., & PINKER, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Martin (Eds.), *Attention and performance XI* (pp. 171–188). Hillsdale, NJ: Erlbaum.
- ERIKSEN, C. W., & HOFFMAN, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, **12**, 201–204.
- HANSEN, J. C., & HILLYARD, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 1–19.
- HANSEN, J. C., & HILLYARD, S. A. (1988). Temporal dynamics of human auditory selective attention. *Psychology*, **25**, 316–329.
- HAWKINS, H. L., HILLYARD, S. A., LUCK, S. J., MOULOUA, M., DOWNING, C. J., & WOODWARD, D. P. (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 802–811.
- 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., MANGUN, G. R., WOLDORFF, M. G., & LUCK, S. J. (1995). Neural systems mediating selective attention. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 665–681). Cambridge, MA: MIT Press.
- HINK, R. F., FENTON, W. H., PFEFFERBAUM, A., TINKLENBERG, J. R., & KOPELL, B. S. (1978). The distribution of attention across auditory input channels: An assessment using the human evoked potential. *Psychophysiology*, **15**, 466–473.
- HINK, R. F., VOORHIS, S. T., VAN, HILLYARD, S. A., & SMITH, T. (1977). The division of attention and the human auditory evoked potentials. *Neuropsychologia*, **15**, 597–605.
- HOWELL, D. C. (1992). *Statistical methods in psychology*. Boston: PWS-Kent.
- JENKINS, W. M., & MASTERTON, R. B. (1982). Sound localization: Effects of unilateral lesions in the central auditory system. *Journal of Neurophysiology*, **47**, 987–1016.
- LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101–124.
- LUCK, S. J., HILLYARD, S. A., MOULOUA, M., WOLDORFF, M. G., CLARK, V. P., & HAWKINS, H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 887–904.
- MANGUN, G. R., & HILLYARD, S. A. (1988). Spatial gradients of visual attention: Behavioral and electrophysiological evidence. *Electroencephalography & Clinical Neurophysiology*, **70**, 417–428.
- MONDOR, T. A., & ZATORRE, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, No. 2, 387–409.
- MULDER, G., WIJERS, A. A., BROOKHUIS, K. A., SMID, H. G. O. M., & MULDER, L. J. M. (1994). Selective visual attention: Selective cuing, selective cognitive processing, and selective response processing. In H.-J. Heinze, T. F. Münte, & G. R. Mangun (Eds.), *Cognitive electrophysiology*. Boston: Birkhäuser.
- MÜLLER, H. J., & HUMPHREYS, G. W. (1991). Luminance-increment detection: Capacity-limited or not? *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 107–124.
- NÄÄTÄNEN, R. (1982). Processing negativity: An evoked-potential reflection of selective attention. *Psychological Bulletin*, **92**, 605–640.
- NÄÄTÄNEN, R. (1992). *Attention and brain function*. Hillsdale, NJ: 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.
- OKITA, T. (1979). Event-related potentials and selective attention to auditory stimuli varying in pitch and localization. *Biological Psychology*, **9**, 271–284.
- OKITA, T. (1981). Slow negative shifts of the human event-related potential associated with selective information processing. *Biological Psychology*, **12**, 63–75.
- OKITA, T., & WEI, J.-H. (1993). Effects of eye position on event-related potentials during auditory selective attention. *Psychophysiology*, **30**, 359–365.
- OLDFIELD, S. R., & PARKER, S. P. A. (1984). Acuity of sound localization: A topography of auditory space: I. Normal hearing conditions. *Perception*, **13**, 581–600.
- PERROT, D. R., CONSTANTINO, B., & CISNEROS, J. (1993). Auditory and visual localization performance in a sequential discrimination task. *Journal of the Acoustical Society of America*, **93**, 2134–2138.

- PICTON, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, **9**, 456-479.
- SCHRÖGER, E., & EIMER, M. (1993). Effects of transient spatial attention on auditory event-related potentials. *NeuroReport*, **4**, 588-590.
- SCHRÖGER, E., & EIMER, M. (1997). Endogenous covert spatial orienting in audition: "Cost-benefit" analyses of reaction times and event-related potentials. *Quarterly Journal of Experimental Psychology*, **50A**, 457-474.
- SCHWENT, V. L., & HILLYARD, S. A. (1975). Evoked potential correlates of selective attention with multichannel auditory inputs. *Electroencephalography & Clinical Neurophysiology*, **38**, 131-138.
- SCHWENT, V. L., SNYDER, E., & HILLYARD, S. A. (1976). Auditory evoked potentials during multichannel selective listening: Role of pitch and localization cues. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 313-325.
- SHULMAN, G. L., WILSON, J., & SHEEHY, J. B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics*, **37**, 59-65.
- SPENCE, C., & DRIVER, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1005-1030.
- SPERLING, G. (1984). A unified theory of attention and signal detection. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 103-181). New York: Academic Press.
- TEDER, W. (1994). *Selective attention and auditory event-related potentials*. Doctoral dissertation, University of Helsinki.
- TEDER, W., ALHO, K., REINIKAINEN, K., & NÄÄTÄNEN, R. (1993). Interstimulus interval and the selective attention effect on auditory ERPs: "N1 enhancement" versus processing negativity. *Psychophysiology*, **30**, 77-87.
- TEDER, W., & NÄÄTÄNEN, R. (1994). Event-related potentials demonstrate a narrow focus of auditory spatial attention. *NeuroReport*, **5**, 709-711.
- VERLEGER, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, **34**, 131-156.
- WOLDORFF, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, **30**, 98-119.
- WOLDORFF, M. G., GALLEN, C. C., HAMPSON, S. A., HILLYARD, S. A., PANTEV, C., SOBEL, D., & BLOOM, F. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, **90**, 8722-8726.
- WOLDORFF, M. G., & HILLYARD, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography & Clinical Neurophysiology*, **79**, 170-191.
- WOOD, N. L., & COWAN, N. (1995). The cocktail party phenomenon revisited: Attention and memory on the classic selective listening procedure of Cherry (1953). *Journal of Experimental Psychology: General*, **124**, 243-262.
- YOST, W. A. (1992). Auditory perception and sound source determination. *Current Directions in Psychological Science*, **1**, 179-184.
- YOST, W. A. (1997). The cocktail party problem: 40 years later. In R. H. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 329-347). Mahwah, NJ: Erlbaum.
- YOST, W. A., DYE, R. H., JR., & SHEFT, S. (1996). A simulated "cocktail party" with up to three sound sources. *Perception & Psychophysics*, **58**, 1026-1036.
- YUND, E. W., EFRON, R., & NICHOLS, D. R. (1990). Detectability gradients as a function of target location. *Brain & Cognition*, **12**, 1-16.

NOTES

1. Prior to measurements, the audio amplifiers and loudspeakers were calibrated to equal gain within a small band of ± 1.0 dB averaged across five octaves ranging from 300 to 9600 Hz.

2. The N1 measure can be obtained for each speaker during each attention condition and thus can be directly associated with behavioral measures (detection response rates and RTs) obtained under the same condition. The Nd measure, on the other hand, requires subtraction of ERPs taken under two different attention conditions.

3. Between the experimental conditions, the peak latencies of the auditory N1 differed significantly [$F(2,11) = 18.0, p < .0001$]. A Tukey-Kramer multiple comparisons posttest revealed that the N1 (site Cz) in the attend-left condition peaked earlier (at 152 msec) than in the attend-right condition (at 173 msec) ($q = 3.6, p < .05$). In the attend-center condition, the N1 peak latency (201 msec) was significantly later than during either the attend-left ($q = 8.4, p < .001$) or attend-right ($q = 4.7, p < .01$) condition. The peak latency of the early Nd did not differ between the attend-left condition (160 msec) and the attend-right condition (170 msec) but was significantly later (200 msec) in the attend-center condition than in either the attend-left ($q = 4.5, p < .05$) or the attend-right ($q = 3.7, p < .05$) condition.

4. An analysis of differences in N1 mean amplitudes to standard stimuli between conditions yielded a significant result [$F(2,11) = 4.8, p < .02$], mainly due to the fact that the mean amplitude in attend-center condition was on the average somewhat smaller than in the attend-left condition ($q = 4.1, p < .05$). The mean amplitudes of the early Nd were similar for the attend-left and attend-center conditions, whereas the amplitude in the attend-right condition was significantly larger ($q = 4.1, p < .05$) than in the other two conditions.

5. There were no significant latency differences for the late Nd peak between attend-left (370 msec), attend-center (390 msec), and attend-right (370 msec) conditions. However, there was an amplitude difference between the conditions [$F(2,11) = 3.8, p < .004$], due to the late Nd in the attend-center condition being significantly smaller than in either the attend-left ($q = 3.7, p < .05$) or the attend-right ($q = 4.4, p < .05$) condition.

6. However, the finding that the N1 latencies to the attended deviants were actually shorter than to the attended standards argues against the presence of a mismatch-negativity (MMN) contribution here, since the MMN typically has a longer latency than the N1. A full analysis of mismatch-specific components is beyond the scope of this paper.

(Manuscript received June 19, 1997;
revision accepted for publication November 2, 1997.)