Signal clustering modulates auditory cortical activity in humans

CLAUDE ALAIN and DAVID L. WOODS

University of California, Davis, California and VA Medical Center, Martinez, California

Auditory streaming and its relevance to attentional processing was examined using event-related brain potentials (ERPs) in situations facilitating perception of one or two streams of sounds. Subjects listened to sequences of brief tones of three different frequencies presented in random order. In evenly spaced (ES) conditions, the three frequencies were equidistant on the musical scale. In clustered, easy (CE) conditions, the attended frequency was distinct, while the middle and extreme distractor tones were clustered together. In clustered, hard (CH) conditions, the attended frequency was clustered with one of the distractors. The subjects pressed a button in response to occasional target tones of longer duration at a prespecified frequency. The subjects were faster and more accurate in CE conditions than they were in ES conditions, and ERP attention effects were enhanced in amplitude in CE conditions. Conversely, the subjects were slower and less accurate in CH conditions and ERP attention effects were delayed in latency and decreased in amplitude. Clustering effects suggest that the processing of stimuli belonging to the attended stream was promoted and the processing of those falling outside the stream was inhibited. The timing and scalp distribution of clustering-related changes in ERPs suggest that clustering modulates early sensory processing in auditory cortex.

The clustering by frequency of successive acoustic events influences our ability to distinguish auditory stimuli from one another. For example, order judgments of target tones embedded in short, frequently repeated sequences are improved when distractor tones are clustered together (Bregman & Rudnick, 1975; Jones, Kidd, & Wetzell, 1981). The clustering of distractors has also been shown to improve the speed as well as the accuracy of target detection, even when the tones are presented in random order (Alain & Woods, 1993). Conversely, the clustering of attended tones with distractors degrades performance and slows reaction times (RTs; Alain, Achim, & Richer, 1993). These performance variations can be explained by the clustering-based segregation of tone sequences into distinct perceptual streams. Improved discriminability is observed when the distractors are segregated in a perceptual stream distinct from the targets. On the other hand, discriminability is decreased when targets and distractors are grouped together.

Bregman (1990) proposed that auditory stream segregation is the result of preattentive perceptual grouping

mechanisms. That is, individual events are analyzed automatically and assigned to the correct stream on the basis of Gestalt principles such as grouping by tonal similarity and temporal proximity. This implies that the ability to selectively attend to one stream of sounds will vary with the stimulation context, improving in situations that promote the organization of sounds into a coherent auditory scene. From this perspective, preattentive mechanisms facilitate grouping of sounds that arise from the same source, whereas attentional processes select one stream from several competitive sources of sounds for further processing. Despite extensive investigation of the psychophysical parameters underlying the organization of sounds into a coherent auditory scene, little is known of the neurophysiological basis of auditory streaming or of its interaction with attentional processing.

The current study was designed to evaluate the effect of clustering on sensory processing using event-related brain potentials (ERPs)¹ recorded during an auditory selective attention task. The ERP methodology is well suited to evaluate the role of perceptual grouping in selective processing because it allows the study not only of variables that can affect the processing of target stimuli, but also of those that can have an influence on the processing of nontarget stimuli. Moreover, appropriate analysis of the spatial and temporal information in the ERPs can lead to inferences regarding the brain region(s) involved in perceptual and cognitive operations.

In auditory selective attention tasks, ERPs to attended tones show a negative displacement relative to ERPs to the same tones when they are nonattended (see Näätä-

This research was supported by grants from the NIDCD and the VA Research Service to D.L.W., and by a postdoctoral fellowship provided by the Fonds de la Recherche en Santé du Québec to C.A. Our special thanks to Greg Shenaut and John Lackey for software development. We would also like to thank K. H. Ogawa, K. Alho, R. Näätänen, M. G. Woldorff, A. S. Bregman, L. Krueger, and an anonymous reviewer for helpful comments. Correspondence should be addressed to Claude Alain, Department of Neurology, Northern California Systems of Clinics (612/127), 150 Muir Road, Martinez, CA 94553 (e-mail: marva4@calain@ucdavis.edu).

nen, 1992; Woods, 1990, for reviews). These attention effects can be isolated by subtracting the ERPs to nonattended stimuli from the ERPs to the same stimuli when they are attended. The resulting slow-negative difference wave (Nd) has at least two partially overlapping portions—namely, an early component (Nd_e) and a late component (Nd_l)—that peak, respectively, at approximately 200 msec and 400 msec following tone onset (Hansen & Hillyard, 1980). The Nd_e may begin at latencies as short as 50 msec following tone onset and has been associated with an early selection of sensory inputs on the basis of their physical characteristics (Hillyard & Picton, 1988; Näätänen, 1990, 1992). The Nd_e appears to include contributions from enhanced sensory processing (gating) of attended tones (Hillyard, Hink, Schwent, & Picton, 1973; Woldorff et al., 1993; Woldorff & Hillyard, 1991; Woods, Alho, & Algazi, 1994) and from an endogenous processing negativity (Näätänen, 1992; Teder, Alho, Reinikainen, & Näätänen, 1993; Woods et al., 1994).

In selective attention tasks, ERPs to nonattended tones may also show an enhanced positivity when compared with ERPs in “neutral” conditions. This increased positivity occurs only when the attended and nonattended tones are easily distinguishable, suggesting that it may reflect an inhibition of processing of nonattended tones (Alho, Töttölä, Reinikainen, Sams, & Näätänen, 1987; Alho, Woods, & Algazi, 1994; Donald, 1987; Michie, Solowij, Crawford, & Glue, 1993).

These ERP attention effects have been shown to be influenced by auditory streaming. For example, Nd waves begin earlier when the attended and nonattended streams of sounds are made easily distinguishable (Alho et al., 1987; Alho, Paavilainen, Reinikainen, Sams, & Näätänen, 1986; Alho, Töttölä, et al., 1987; Hansen & Hillyard, 1980). Recently, Alain et al. (1993) examined the relationship between clustering and ERP attention effects in a task in which subjects focused attention on one frequency in tone sequences of four different frequencies. Two clustering conditions were used, namely, (1) a baseline condition in which the four frequencies were evenly spaced on the musical scale, and (2) a grouped condition in which the two middle frequencies were closer to the extreme frequencies. An enhancement of Nd amplitude was found at 80–120 msec from stimulus onset when the frequencies were clustered into two streams. This clustering effect on Nd was mainly mediated by an increased positivity to irrelevant tones, and as such is compatible with an active inhibition of distracting events. This is consistent with the behavioral result showing that distractor clustering reduces interference on target detection (Alain & Woods, 1993).

However, it is conceivable that signal clustering can also facilitate the processing of attended stimuli. For example, clustering may function like an active bandpass filter, promoting the processing of relevant signal frequencies and inhibiting the processing of irrelevant fre-

quencies. This facilitatory–inhibitory mechanism would accentuate the contrasts between the relevant signals and background stimuli. In the current study, we aimed to evaluate facilitatory and inhibitory effects of clustering on auditory processing, to estimate the latency of clustering effects, and to evaluate the brain regions involved by sampling electrical activity from a widely distributed array of scalp electrodes.

Two further aims of the current study were to evaluate both the effect of signal clustering on middle-latency auditory evoked potentials (MAEPs) and the role of auditory clustering in the gating of nonattended inputs. Previous reports have suggested that MAEPs may be modulated by attention, particularly when stimuli are highly discriminable and presented at rapid rates (Woldorff et al., 1993; Woldorff & Hillyard, 1991); this suggests that MAEP attention effects may be sensitive to clustering manipulations. The experiment also allowed the evaluation of the effect of clustering on the mismatch negativity (MMN), a broad negative wave peaking at around 180 msec poststimulus (Näätänen, 1992; Näätänen, Gaillard, & Mäntysalo, 1978). The MMN has been suggested to reflect the automatic processing of stimulus deviance. However, recent reports suggest that the MMN may be modulated by attention (Alho, Woods, Algazi, & Näätänen, 1992; Näätänen, Paavilainen, Titiinen, Jiang, & Alho, 1993; Woods, Alho, & Algazi, 1992). Attention effects on the MMN have been particularly prominent at high rates of stimulus delivery (Woldorff, Hackley, & Hillyard, 1991). This rate-dependence suggests that MMN attentional modulation might depend critically on perceptual grouping.

The stimulus sequences used in the current study are shown in Figure 1. Three clustering conditions were used, in which the extreme frequencies were alternately attended in successive blocks. In the evenly spaced (ES) conditions, the attended and distractor tones differed by 8 and 16 semitones, respectively. In clustered, easy (CE) conditions, the two irrelevant frequencies were clustered together and the attended frequency differed from the distractors by 8 and 9 semitones. In clustered, hard (CH) conditions, the attended frequency was clustered with the middle frequency, and the attended frequency differed from the distractors by 1 and 8 semitones. In all conditions, the subject's task was to respond to infrequent tones of slightly longer duration at a prespecified frequency (target) mixed with frequent tones of shorter duration (distractors). Target discriminability was adjusted in pilot studies to encourage a hierarchical selection, whereby the subjects would first pay attention to the frequency (the easy dimension) and then make judgments on the duration (the more difficult dimension). If clustering effects on performance are mediated only through decreased interference of distractors, we might expect differences in ERP components elicited by those unattended stimuli, but not on those elicited by attended stimuli. On the other hand, if streaming involves a facilitatory-

Figure 1. Examples of the stimuli presented in the three different clustering conditions. Arrows indicate the frequency to be attended in each clustering condition. A, B, and C refer to frequency of tones [A = high frequency (2641 or 1763 Hz); B = middle frequency (1664 Hz); note that these "B" tones were presented on every block; and C = low frequency (1048 or 1570 Hz)]. * Targets of longer duration.

inhibitory mechanism, differences in amplitude for ERPs to both attended and nonattended stimuli might be expected.

METHOD

Subjects

Twelve subjects (6 males and 6 females aged 20–33 years) participated in this study. Seven additional volunteers were excluded from the experiment because of either poor performance during training or excessive eye blinks during electroencephalogram (EEG) recording. The subjects were paid for their participation and gave informed consent according to Veterans Administration Medical Center and University of California, Davis guidelines.

Stimuli and Task

The stimulus sequences consisted of tones of three different frequencies presented in random order (Figure 1). These sequences were constructed such that each frequency occurred on 33.33% of the trials. Within each frequency, both short tones (50 msec, 27.8%) and long tones (90 msec, 5.5%) were presented. Overall, long tones occurred in 16.5% of the trials. All stimuli were shaped with 5-msec rise/fall times and delivered binaurally through TDH-39 headphones at 78 dB (SPL) over 55 dB (SPL) broadband noise. Stimuli were presented in random order at variable stimulus onset asynchrony (SOA; 200–400 msec with 10-msec step, mean = 300 msec) in blocks of 1,542 tones.

The subjects were presented with three different clustering conditions: (1) ES conditions, in which the three frequencies were equally spaced, 8 semitones apart (1048-, 1664-, and 2641-Hz

tones; Figure 1, top); (2) CE conditions, in which the extreme distractor tone was paired with the middle distractor, these two tones differing by one semitone (e.g., 2641-, 1664-, and 1570-Hz tones; or 1773-, 1664-, and 1048-Hz tones; Figure 1, center); and (3) CH conditions, in which tone sequences were identical to the CE conditions, except that the subjects were instructed to focus their attention on the extreme frequency adjacent to the middle tone (Figure 1, bottom). The middle frequency remained constant in all clustering conditions.

The subjects were seated in a sound-attenuated room and were instructed to focus their attention on either the lowest pitched (1048 or 1570 Hz) or the highest pitched (1763 or 2641 Hz) tone and to ignore the two other frequencies. They were told to press a button as quickly as possible in response to the occasional tones of longer duration of the designated frequency (targets, 5.5% of all stimuli), while trying to make as few errors as possible. They were informed that occasional long-duration tones would also occur at irrelevant frequencies (deviants; i.e., tones of the same duration as the target but differing in frequency). The frequency to be attended was counterbalanced across blocks. Before beginning the study, each subject practiced listening to the tone sequences and detecting the occasional targets. Following the practice blocks, the subjects performed in six conditions, defined by all possible combinations of attended frequency (high or low) and experimental condition (ES, CE, or CH). The order of conditions was counterbalanced across subjects.

EEG Recording

The EEG (bandpass 0.1–300 Hz) was recorded from 28 electrodes over the scalp (Fp1, Fpz, Fp2, nose, left preauricular, right preauricular, T1, F7, F3, Fz, F4, F8, T2, left mastoid, T3, C3, Cz, C4, T4, right mastoid, T5, P3, Pz, P4, T6, O1, Oz, and O2). Vertical and horizontal eye movements were recorded from electrodes placed laterally and below the left eye. All electrodes were referenced to four interconnected noncephalic electrodes at the base of the neck (Woods & Clayworth, 1985).

Data Analysis

The EEG was continuously digitized at 833 Hz/channel. The MAEPs were averaged directly from original data, but long-latency ERPs were analyzed after the EEG had been decimated by averaging together three successive sample points to reduce the effective sampling rate to 278 Hz. ERPs were extracted off-line by computer. Trials contaminated by horizontal or vertical eye movements, excessive peak-to-peak deflection (over 100 μ V), or amplifier clippings were excluded from the average. The ERPs were then averaged separately for each stimulus type and for each condition. In all cases, ERPs to targets included only the individual trials that were followed by a correct response (defined below). In the current study, individual trials that were preceded and/or followed by a false alarm (FA) were not excluded from the averages. Before analysis, long-latency ERPs were low-pass digitally filtered to eliminate frequencies above 40 Hz.

MAEP (Na, Pa, Nb, and P20–50) measurements were obtained relative to a 50-msec prestimulus baseline, with the effects of attention evaluated on mean voltages obtained at 5-msec intervals from 10 to 50 msec. Long-latency ERPs (P1, N1, P2, Nd, Nd, P3b, and MMN) were measured relative to a 200-msec prestimulus baseline. The effects of attention and clustering on long-latency ERPs were tested on successive mean voltages obtained at 10-msec intervals from 50 to 150 msec, at 20-msec intervals from 150 to 310 msec, and at 50-msec intervals from 310 to 700 msec poststimulus. The results were statistically evaluated using repeated measures analyses of variance (ANOVAs).

Traditionally, ERP studies of auditory selective attention have used balanced stimulus sequences in which the task difficulty of different conditions is equated. While this was true for the ES conditions of the current experiment, in the CE and CH conditions,

task difficulty and stimulus factors could not be simultaneously equated because of the stimulus asymmetry associated with clustering. Thus, there were two ways to compare the data in the latter conditions. First, physical stimulus factors (but not task difficulty) could be equated by comparing ERPs to identical stimuli when they were attended and nonattended in different conditions. For example, in ES conditions, ERPs to Tone A when it was attended were compared with those recorded for Tone A when Tone C was attended (see Figure 1, top). In clustered conditions, however, ERPs to attended tones in CE or CH conditions were compared with those obtained in the complementary CH or CE conditions. For example, in CE conditions, ERPs to Tone A when it was attended were compared with those recorded in CH conditions when Tone C was attended (see Figure 1, middle and bottom). Second, task difficulty (but not physical stimulus factors) could be equated by comparing ERPs to attended tones with those recorded for nonattended tones within the same condition (e.g., ERPs to Tone A vs. ERPs to Tone B in CE condition; see Figure 1, middle). We report the results of both analyses.

Several precautions were taken to control possible Type I errors: (1) Main effects were first tested at the Fz electrode and were only reported if they reached significance for several related measures (peak and mean voltage, or several adjacent mean voltage measures); (2) scalp distributions were only reported if they were significant after a strict Greenhouse-Geisser correction, in which the degrees of freedom for the electrode factor were reduced to two—for example, scalp distribution interactions with 21 electrodes and 12 subjects [$F(20,220)$] were only reported if F ratios exceeded those required for significance at $F(2,22)$; and (3) in comparing the scalp distributions of different components, the data were first normalized to control for component $\times$ electrode interactions resulting simply from differences in component amplitude (McCarthy & Wood, 1985).

The RTs were analyzed for correct trials only (i.e., those trials with RTs occurring between 200 and 800 msec after target presentation). Responses at other times were classified as FAs. Mean RTs for each subject were compared across conditions using two-way repeated measures ANOVAs, with attended frequency (A or C) and clustering condition (ES, CE, or CH) as factors. The number of misses and the number of FAs were also analyzed using two-way ANOVAs.

For behavioral as well as ERP data, target frequency did not produce a significant main effect or an interaction with attention and/or clustering conditions. All the subsequent analyses, tables, and figures will therefore be collapsed across frequency.

RESULTS

Task Performance

The mean RTs and percentages of misses and FAs for the three clustering conditions are presented in Table 1. The ANOVAs revealed a main effect of condition on the number of misses [$F(2,22) = 94.79, p < .001$]. Pairwise comparisons showed that the subjects detected more targets in CE conditions than they did in ES conditions

Table 1
Mean Performance for Each Condition Lumped
Over Attended Frequency

Conditions	Misses %	False Alarms %	Reaction Time (msec)
Evenly spaced	21.8	7.2	454
Clustered, easy	18.9	6.3	444
Clustered, hard	38.7	20.8	482

Table 2
Mean RT (in Milliseconds) to Correctly Detected Target Preceded by a Short-Duration Tone, for Each Condition

Condition	Preceding Tone Type		
	Distant	Middle	Identical
Evenly spaced	451	467	445
Cluster, easy	453	446	431
Cluster, hard	505	490	472

[$F(1,11) = 5.31, p < .05$], and more targets in ES than in CH conditions [$F(1,11) = 130.08, p < .001$].

Clustering also affected the number of FAs [$F(2,22) = 7.82, p < .01$]. Subjects made more FAs in CH conditions than they did in ES or CE conditions [$F(1,11) = 6.7$ and 9.7 , respectively, $p < .05$]. No difference between ES and CE conditions was found in terms of the number of FAs.

Clustering also had significant effects on RTs [$F(2,22) = 17.05, p < .001$]. Pairwise comparison showed that subjects responded more rapidly to target tones in CE conditions than to those in ES conditions [$F(1,11) = 8.87, p < .05$] and that they were faster in ES conditions

than they were in CH conditions [$F(1,11) = 10.79, p < .01$]. No other main effects or interactions reached significance.

Previous research has shown that the features of a distractor tone preceding a target influence the speed of target detection (Woods, 1990). The effect of a preceding tone on target detection (i.e., the priming effect) seems sensitive to clustering manipulation (Alain & Woods, 1993). To further evaluate the interaction between physical clustering and priming, the effect of a preceding distractor on target detection was compared across conditions. The effect of preceding-tone-type on RT for the three conditions is presented in Table 2. Mean RTs were faster when the target was preceded by a tone of identical frequency than they were when it was preceded by a middle or distant distractor tone [450 msec vs. 468 and 470 msec, respectively, $F(2,22) = 7.16, p < .01$]. These priming effects were somewhat reduced by clustering. However, in contrast to previous studies (Alain & Woods, 1993), the interaction between the preceding-tone-type and clustering factors failed to reach significance [$F(4,44) = 2.08$].

Figure 2. Grand means of 12 subjects for ERPs recorded at different scalp electrodes in ES conditions. Solid line = attended frequency; dashed line = middle distractor; dotted line = extreme distractor. ERPs have been averaged over the two attended frequencies. In this and subsequent figures, the stimulus onset is indicated by the vertical bar and the negativity is plotted upwards. *Insert:* Enlarged ERPs from frontal and right temporal sites.

Figure 3. Grand means of 12 subjects for ERPs recorded at different scalp electrodes in CE conditions. Solid line = attended frequency; dashed line = middle distractor; dotted line = extreme distractor. ERPs have been averaged over the two attended frequencies. *Insert*: Enlarged ERPs from frontal and right temporal sites.

Middle Latency Auditory Evoked Potentials (MAEPs)

In all conditions, short-duration tones elicited reliable MAEP components, including Na (20 msec, mean peak latency at Fz), Pa (31 msec), and Nb (38 msec). In CE conditions, ERPs to attended tones were positively displaced from 10 to 40 msec in comparison with ERPs to the extreme distractors [$F(1,11)$ ranged from 4.13 to 7.31, $p < .05$]. However, the difference failed to reach significance when stimulus factors were equated (by comparing ERPs recorded to the attended extreme tone in CE conditions with ERPs to the same nonattended extreme tone in CH conditions). There were no significant attentional effects on MAEPs for the ES or CH conditions, regardless of whether stimulus factors or task difficulty were equated, and interactions between attention and clustering factors also failed to reach significance.

Long-Latency Auditory ERPs to Frequent Short Duration Tones

Figures 2, 3, and 4 show group average ERPs in ES, CE, and CH conditions, respectively. The ERPs were

characterized by polyphasic deflections, including P1 (53 msec, mean peak latency at Fz), N1 (127 msec), and P2 (221 msec). In general, ERPs to attended tones were more negative than ERPs to distractors. Further, ERPs to attended tones in ES and—particularly—in CE conditions, were characterized by an enhancement of positivity around 100 msec poststimulus (P1'; Woods, in press), visible mainly at posterior inferior temporal sites. The peak latency of the P1' deflection occurred during the rising phase of the N1 deflection.

In ES conditions, ERPs were characterized by an attention-related gradient of negativity. Attended tones showed the maximum negativity, while intermediate negativity was seen for middle-frequency distractors and minimal negativity was noted for nonattended extreme distractors. This attention-related gradient of negativity can be seen in the difference waves from the ES conditions (Figure 5, top). The attention effect on ERPs increased with the frequency separation between the attended and nonattended frequencies, being larger for the attended minus extreme difference waves than for attended minus middle difference waves [from 130 to 250 msec, $F(1,11)$ ranging from 13.2 to 42.2, $p < .05$].

Figure 4. Grand means of 12 subjects for ERPs recorded at different scalp electrodes in CH conditions. Solid line = attended frequency; dashed line = middle distractor; dotted line = extreme distractor. ERPs have been averaged over the two attended frequencies. *Insert*: Enlarged ERPs from frontal and right temporal sites.

Complementary patterns emerged in the two clustering conditions. In CE conditions, attended tones showed maximum negativity, and ERPs to middle distractors resembled those of the nonattended extreme distractors (Figure 3, insert). As a result, attended minus middle difference waves and attended minus extreme difference waves were virtually identical (Figure 5, middle). This suggests that the grouping manipulation had been successful: processing appeared similar for nonattended middle and extreme pitches. The ERP pattern reflected a decrease in negativity in the ERPs to the middle tone in CE conditions relative to those in ES conditions [from 70 to 170 msec interval and from 230 to 350 msec, $F(1,11)$ ranged from 4.75 to 17.89, $p < .05$ for all comparisons]. The ERPs to nonattended extreme tones had similar amplitudes in ES and CE conditions.

In CH conditions, ERPs to middle tones resembled those to attended tones (Figure 4, insert). Both showed enhanced negativity when compared with the nonattended extreme tone. As a result, attended minus extreme difference waves and middle minus extreme difference waves were virtually identical (Figure 5, bottom). This suggests that attended and middle frequency tones had been

grouped together and underwent similar processing. The ERPs to attended extreme tones were less negative in CH than in ES conditions during the 50–170-msec and the 250–350-msec intervals, whereas ERPs to nonattended middle tones were more negative in CH than in ES conditions during the 150–250-msec and the 450–550-msec intervals [$F(1,11)$ ranged from 5.09 to 26.81, $ps < .05$].

Clustering/Nd waves. Figure 6 shows ERPs and difference waves from frontal and right posterior temporal electrodes during different clustering conditions when stimulus factors were controlled. In ES conditions, the difference waves were obtained by subtracting ERPs to extreme nonattended tones (e.g., those to Tone C when Tone A was attended; see Figure 1, top) from ERPs to the *same* extreme tones when attended (e.g., those to Tone C when it was attended). In CE conditions, the difference waves were obtained by subtracting ERPs to nonattended tones in CH conditions (e.g., those to Tone A when Tone C was attended) from ERPs to the *same* tones when attended in CE conditions (e.g., those to Tone A when they were attended). Finally, in CH conditions, the difference waves were obtained by subtract-

Figure 5. Group means of 12 subjects for ERP difference waves obtained within each condition at Fz by subtracting ERPs to nonattended extreme or middle tones from ERPs to attended tones. Difference waves have been averaged over frequency. Solid line = attended minus middle distractors; dashed line = attended minus extreme distractors; dotted line = middle distractors minus extreme distractors.

ing ERPs to nonattended tones in CE conditions (e.g., those to Tone A when Tone C was attended) from ERPs to the *same* tones when attended in CH conditions (e.g., those to Tone A when it was attended). This subtraction procedure was used to assure that the difference waves controlled for neuronal refractory effects.² The ERPs were elicited by physically identical stimuli in each case; only the direction of attention varied.

Signal clustering altered the amplitude of stimulus-controlled Nd waves [from 50 to 400 msec at Fz, $F(2,22)$ ranged from 4.60 to 20.43, $p < .05$]. Pairwise comparisons showed that the Nd was larger in CE conditions than it was in ES conditions [from 170 to 230 msec poststimulus, $F(1,11)$ ranged from 4.11 to 5.00, $p < .07$, and from 250 to 350 msec, $F(1,11)$ ranged from 8.61 to 57.26, $p < .05$]. To determine whether Nd alterations were due to enhanced negativity to attended tones and/or to enhanced positivity to nonattended tones, separate comparisons of mean voltages relative to the 200-msec prestimulus baseline were made on ERPs to attended and nonattended tones. Clustering effects on Nd reflected both enhanced positivity to nonattended tones [CH vs. ES conditions from 190 to 450 msec poststimulus,

$F(1,11)$ varied between 5.77 and 31.47, $p < .05$] and increased negativity to attended tones [CE vs. ES conditions, from 150 to 170 msec, $F(1,11) = 4.55$, $p = .054$].

On the other hand, stimulus-controlled Nd waves were reduced in amplitude and their onset was delayed by 60–80 msec in CH conditions relative to ES conditions [from 50 to 270 msec poststimulus, $F(1,11)$ varied between 5.73 and 41.45, $p < .05$]. This was due mainly to decreased negativity to attended tones [CH vs. ES conditions; during the 50–170-msec and the 250–350-msec intervals, $F(1,11)$ ranged from 5.50 to 50.26, $p < .05$], rather than to a reduced positivity to nonattended tones. N1 latencies were also delayed by 15 msec in CH conditions in comparison with ES conditions.

At posterior temporal electrodes, the attention-related enhancement of the P1' component was larger in CE than in ES conditions [from 110 to 130 msec, $F(1,11) = 9.09$ and 8.19, $p < .05$]. This enhanced positivity was followed by an increased negativity during the 150–170-msec interval for CE conditions relative to ES conditions [$F(1,11) = 8.56$, $p < .05$]. Differences between CH and ES conditions in P1' amplitude did not reach significance.

Figure 6. Group means of 12 subjects for ERPs to the same tones when attended and nonattended, and their corresponding difference waves, recorded at Fz and T6. Solid line = CE; dashed line = ES; dotted line = CH. ERPs have been averaged over frequency.

The distribution of the attention effect on ERPs varied with signal clustering. In ES and CE conditions, the early and late phases of Nd (Nd_e , 100–250 msec; Nd_l , 250–400 msec) showed maximum voltages at fronto-central sites, whereas the Nd waves obtained in CH conditions were maximal over central areas. This produced significant clustering $\times$ electrode interactions [from 170 to 250 msec, $F(20,220)$ ranged from 3.30 to 10.73, $p < .05$; from 250 to 400 msec, $F(20,220)$ ranged from 3.40 to 15.04, $p < .05$].

ERPs to Target Tones

The ERPs to correctly detected targets were characterized by N1 and P3b deflections, the latter peaking on average at 503 msec following target onset (Figure 7, left panel). The ERPs to long deviant tones showed an enhanced early negativity in comparison with short-duration tones, due in part to contributions from the overlapping MMN (see below). At fronto-central sites, difference waves between the attended (target) and nonattended deviant stimuli were enhanced in comparison with Nd waves to frequent short tones, but showed identical patterns to the latter as a function of grouping. As with the short tones, stimulus-controlled target dif-

ference waves were larger in CE than in ES conditions [from 270 to 400 msec, $F(1,11)$ varied between 6.60 and 33.51, $p < .05$]. Conversely, target difference waves were smaller in CH than in ES conditions [from 120 to 170 msec and from 250 to 300 msec, $F(1,11)$ varied between 5.58 and 15.18, $p < .05$]. As with the short tones, the enhanced difference waves in CE conditions reflected both enhanced negativity to attended targets [from 110 to 140 msec, $F(1,11)$ ranged from 6.56 to 10.32, $p < .05$] as well as, at longer intervals, enhanced positivity to non-attended deviants [from 300 to 500 msec, $F(1,11)$ ranged from 7.59 to 16.94, $p < .05$; Figure 7, top center panel].

The latency of the P3b, measured at Pz, varied with clustering conditions [$F(2,22) = 4.73$, $p < .05$]. The P3b was delayed in CH conditions in comparison with ES and CE conditions [520 vs. 498 and 489 msec, respectively; $F(1,11) = 4.14$ and 8.29, $p < .07$]. While no difference was found between CE and ES conditions in the peak latency of P3b [$F(1,11) = .80$], there was a significant early P3b amplitude enhancement in CE conditions [from 400 to 450 msec, for CE vs. ES and CH conditions, $F(1,11) = 6.01$ and 12.11, respectively, $p < .05$].

Clustering/priming. As noted above, RTs were significantly faster when target tones were preceded by

Figure 7. Group means of 12 subjects for ERPs to the *same* deviant tones when attended and nonattended, and their corresponding difference waves, recorded at Fz and Pz. ERPs to attended deviant (target) included only hit trials. Solid line = CE; dashed line = ES; dotted line = CH. ERPs have been averaged over frequency.

tones of identical frequency (see Table 2). Paralleling this was the effect of preceding-tone-type on P3b latency and amplitude for the three clustering conditions (Figure 8). Enhanced early positivity (300–500 msec) was seen when the target was preceded by a short tone of identical frequency, compared with when it was preceded by a tone of different frequency [$F(1,11)$ ranged from 5.49 to 25.80, $p < .05$]. The peak latency of the P3b also varied as a function of the preceding stimulus. Consistent with RT measures, the peak latency of the P3b was earlier when the targets were preceded by a tone of identical frequency (484 msec) than it was when they were preceded by an extreme (508 msec) or middle-frequency (506 msec) tone [$F(1,11) = 9.80$ and 9.32 , $ps < .02$]. As with RTs, no interaction was found between P3b latency as a function of preceding tone type and clustering conditions.

Mismatch Negativity

Figure 9 shows the ERPs for deviant and frequent short tones, and the difference between them, at the mid-frontal (Fz) electrode. The mismatch negativity (MMN)³ began at about 100 msec and peaked at about 200 msec

in all conditions. The amplitude of both the MMN (quantified between 130 and 190 msec) and the N2b (190–250 msec) varied with attention, being larger for tones of the attended frequency than it was for either the middle or the extreme distractors [from 130 to 250 msec, $F(1,11)$ ranged from 5.63 to 20.93, $p < .05$]. No difference was found between the amplitude of the MMN obtained in the middle and the extreme nonattended tones, and interactions between clustering conditions and attention also failed to reach significance.

DISCUSSION

Tonal clustering produced systematic and consistent effects on ERPs and behavioral performance. When distractors were clustered, ERP attention effects were enhanced and performance improved. Conversely, when attended stimuli were clustered with distractors, ERP attention effects decreased and performance deteriorated. These results suggest that the experimental manipulations resulted in effective stimulus grouping. Facilitatory effects were observed when the two distractor frequencies were grouped together into a perceptual stream

Figure 8. Group means of 12 subjects for target ERPs as a function of the relationship between the preceding tone and the target. ERPs to targets included only hit trials. Solid line = same frequency, dashed line = middle distractor; dotted line = extreme distractor. ERPs have been averaged over frequency.

distinct from the attended frequency. On the other hand, when the attended and the middle distractor frequency were grouped together, target discriminability was decreased. In the current study, clustering manipulations were expected to facilitate the organization of sounds into one or two different streams. That is, all stimuli in ES conditions should have been processed as if they belonged to the same stream of sounds. Although we cannot exclude the possibility that the subjects were able to segregate the stimuli into three different streams on the basis of frequency, the gradient of negativity in ES conditions suggests that the stimuli underwent a depth of processing that was mainly determined by their physical similarity to the target. In clustered conditions, the picture was somewhat different: the ERP attention effect did not depend on the physical similarity between the attended and nonattended stimuli, but instead varied according to whether the stimuli belonged to the attended stream of sounds or not.

Ignoring the notion of auditory streaming for a moment, what alternative explanations could account for the ERP results? One possibility is that the differences

between conditions reflect altered prestimulus activity. At short SOAs, ERPs overlap with those from preceding stimuli. As a result of slow-positive and, more particularly slow-negative, waves (e.g., the processing negativity, PN) elicited by preceding stimuli, this overlap could artificially increase or decrease ERP negativity. For example, in CH conditions, the frequency of two thirds of the stimuli was similar to that of the target and therefore those stimuli were more likely to produce a PN of long duration. Thus, ERPs to distractors that were not clustered with the target might have been positively shifted because of overlap with a resolving PN elicited by the majority of the previous stimuli. The enhanced negativity to attended tones in CE conditions could similarly be explained if distractor frequencies produced a PN that was smaller and/or of shorter duration in CE than in ES conditions, thereby producing a less negative prestimulus baseline.

However, for three reasons, alterations in the prestimulus activity cannot adequately account for the results we obtained. First, at the SOAs used here, baseline voltage shifts would have produced maximal effects

Figure 9. Grand means for 12 subjects for ERPs to deviants (top) and to frequent short-duration tones (middle), and the corresponding difference wave (bottom), recorded at Fz in ES, CE, and CH conditions. Solid line = attended extreme tones; dashed line = middle distractors; dotted line = extreme distractors. ERPs have been averaged over frequency.

near the time of stimulus onset, dominated by monophasic change in voltage. In the current experiment, clustering often produced effects of different polarities at different scalp sites. For example, at posterior sites, distractor clustering enhanced the positivity of the P1' and at the same latencies a negative enhancement was seen for the rising phase of the N1 at frontal sites. Second, clustering effects were often polyphasic. At frontal sites, clustering the attended and the middle distractor tones enhanced the amplitude of both the N1 and the P2 deflections of the nonattended extreme tones. Moreover, at posterior sites, distractor clustering enhanced both the positivity of the P1' and the negativity of the subsequent N1c component of the attended tones. Third, clustering had distinct effects on ERPs elicited by different tones within the same condition. For example, relative to the ES conditions, distractor clustering decreased the attention-related negativity elicited by the middle tones, but had little effect on the ERPs to the extreme tone. If the decrease in negativity was due to a baseline shift, the ERPs to extreme tones should have shown a similar enhanced positivity. However, this was not the case. ERPs

to nonattended extreme tones in ES and CE conditions were similar in amplitude.

Another possible explanation of the ERP results is that clustering might have affected ERPs through alterations in refractoriness. Previous reports have shown that the amplitude of the N1 and P2 components decreases with increasing similarity between the tones composing the sequence (Butler, 1968; see Näätänen & Picton, 1987, for a review). In the current study, consistent with an increase in refractoriness in clustered conditions, distractor clustering reduced the amplitude of the N1 deflection to middle tones. However, this variation in negativity is also compatible with a smaller and/or faster decrement of a PN elicited by middle distractors in these conditions. The fact that ERPs to nonattended extreme tones had similar amplitude in CE and ES conditions also argues against an explanation strictly in terms of a difference in the refractory period of the generators.

Auditory Streaming and Controlled Processing

Our data suggest that signal clustering can facilitate both the processing of task-relevant stimuli and the re-

jection of distracting events. The fact that ERPs to attended tones, including both the P1' and N1 deflections, were larger when the distractors were clustered together implies that clustering facilitated sensory processing. If signal clustering influenced performance solely by an active inhibition of distracting frequencies, ERPs to attended tones should have been identical in ES and CE conditions. The cluster-related inhibition of irrelevant events was reflected in the enhanced positivity seen for ERPs to nonattended tones. This enhancement of positivity is consistent with previous reports showing an increment of positivity of ERPs to distractors when the tonal pattern facilitates the segregation of sequences into distinct perceptual streams (Alain et al., 1993; Alho, Töttölä, et al., 1987; Alho et al., 1994; Donald, 1987; Michie et al., 1993). These results are compatible with the notion of an active attentional bandpass filter, in which the processing of relevant signal frequencies is promoted and the processing of irrelevant frequencies is actively inhibited.

Previous reports have shown that attention can produce a positive displacement during the MAEPs (P20-50), particularly when stimuli are highly discriminable and presented at rapid rates (Woldorff et al., 1993; Woldorff & Hillyard, 1991). In the present experiment, selective attention did produce a small, and marginally significant, positive enhancement in MAEPs when distractors were clustered together. This is consistent with the results of Woldorff et al. (1993) and suggests that perceptual grouping may play a role in the emergence of this early attention effect.

The earliest reliable ERP clustering effect was the attentional enhancement of the P1' component. In the current study, the timing and the scalp distribution of clustering effects suggests that clustering altered sensory processing in auditory cortex. In addition to showing an enhanced negativity over fronto-central sites, ERPs to attended stimuli were characterized by an increment of positivity at infero-posterior electrodes during the rising phase of the N1 deflection (cf. Giard, Perrin, Pernier, & Peronnet, 1988). The P1' has been modeled with generator sources in auditory cortex either on the lateral surface of the superior temporal gyrus (Wolpaw & Penry, 1975; Wood & Wolpaw, 1982) or within the superior temporal plane itself (Bertrand, Perrin, & Pernier, 1991; Woods et al., 1994).

The analysis of clustering effects on P3b latency and amplitude provided further evidence that perceptual grouping affected sensory processing. The P3bs were smaller and/or delayed in ES and CH conditions in comparison with CE conditions. This is consistent with the short-latency ERP changes that indicate that clustering altered the time required for the sensory evaluation of target signals.

Auditory Streaming and Automatic Processing

A further aim of our study was to evaluate whether the "automatic processing" indexed by the MMN (Näätä-

nen, 1990, 1992) was affected by attention and/or by auditory clustering. In the current study, deviant tones at nonattended and attended frequencies elicited an MMN. The MMN overlapped the N1 deflection and peaked on average at 200 msec following stimulus onset. Since deviant stimuli had longer durations than standard stimuli in this study, the early phase of the MMN included contributions from enhanced N1 activity (Onishi & Davis, 1968). In all conditions, the MMN was larger for the attended frequency than it was for the two distractors. This MMN difference between attended and nonattended frequencies cannot easily be explained by the overlap of an N2b component associated with target detection, since attentional enhancement was evident at relatively short latencies (130–170 msec). The short latency of attentional effects on the MMN also argues against a contribution of premotor activity. Our data are consistent with previous reports showing that MMN amplitude can vary with attention (Alho et al., 1992; Näätänen et al., 1993; Woldorff et al., 1991; Woods et al., 1992), and throw into question the view that the MMN represents an automatic process (Näätänen, 1990, 1992), indicating, rather, that the MMN amplitude can be modulated by attention. The amplitude of the MMN did not vary with clustering conditions, nor did the interaction between clustering and attention factors reach significance, suggesting that perceptual grouping has little effect on the processing indexed by the MMN.

Auditory Streaming and Priming Effects

The analysis of the effects of stimulus sequence on RTs and ERPs to targets provides new evidence that priming modulates sensory processing. As in previous studies, overall RTs showed a relative improvement when the target tones were preceded by a tone of identical frequency (Alain & Woods, 1993; Woods & Alain, 1993). Priming could affect RT either by facilitating motor responses (Zeef & Kok, 1993) or by improving the sensory analysis of signals. P3b latencies can provide a test of these alternatives, since P3b latency varies with the time required for the sensory analysis of signals but is little affected by demands placed on response production (Magliero, Bashore, Coles, & Donchin, 1984). We found that priming effects on RT were similar in all conditions, reducing it on average by 18 msec when targets were primed by tones of identical frequency in comparison with targets that were preceded by a middle-frequency tone. Priming effects on P3b peak latency were almost identical, averaging 23 msec. This implies that priming speeded sensory processing of targets, perhaps because targets of longer duration could be compared against the clear sensory trace of a short-duration tone. Finally, it should be noted that signal clustering did not alter priming effects on either RT or P3b. This suggests that clustering and priming affected independent processing stages, clustering altering the early sensory processing associated with the separation of attended and nonattended streams of input, and priming affecting the discrimina-

tion of frequent short tones and long-duration targets within the attended input stream.

Implications for the Attentional Trace Model

Näätänen (1990, 1992) argued that the Nd wave reflects an early selection of stimuli based upon a gradual comparison of sensory input to a memory trace (i.e., a temporary neuronal representation of the distinctive physical features of the attended tones). According to Näätänen, each incoming stimulus is compared with the memory trace of the physical features of the attended tone. The comparison process is indexed by a surface negative potential (the PN) whose duration depends upon the similarity between the stimulus and the memory trace. From this perspective, the Nd wave reflects the difference between the PNs to relevant and irrelevant stimuli. According to this model, the onset of the Nd wave indicates the time needed to classify an incoming stimulus as a distractor that can be rejected without further perceptual processing.

We found some support for Näätänen's (1990, 1992) model. For example, in ES conditions, the attention effect on ERPs decreased as a function of the frequency separation between the attended tones and the distractors. This is consistent with an increase in PN to tones that are similar in frequency to the target. The results from the CH conditions can also be interpreted in terms of the attentional trace theory. In this condition, relative to the ES condition, the onset of the Nd wave was delayed. This is consistent with previous reports showing a delay in Nd-wave onset as a result of the decrease in frequency separation between attended and nonattended stimuli (Alho et al., 1986; Alho, Töötölä, et al., 1987).

Although some of our results can be interpreted in terms of Näätänen's (1990, 1992) trace model, the current version of the model has difficulty in explaining three of our findings. First, large variations in ERP negativity were observed as a function of clustering conditions (e.g., between CE and ES conditions). In particular, the attention-related negativity to the middle tones differed between ES and CE conditions even though the frequency separation remained constant. The attentional trace theory would have predicted that the constant pitch difference of the middle tones from the targets should have led to the same amount of ERP negativity for the middle stimuli. Moreover, Nd waves varied systematically as a function of clustering conditions. For example, in CE conditions, the difference between the extreme nonattended tone and the target was diminished by almost 50%, yet Nd difference between the ERP to this tone and the ERP to tones at the attended frequency were enhanced in comparison with ES conditions. Second, clustering effects included an enhancement of negativity to attended tones. In the attentional trace model, however, it is assumed that the PN to the attended tone is maximal and constant, whereas the PN to nonattended stimuli varies according to the similarity between attended and nonattended stimuli. Third, signal clustering enhanced the ERP positivity to distractors that formed a stream dis-

tinct from the attended frequency. This enhanced positivity occurred for middle-frequency tones in CE conditions in comparison with those in ES conditions. However, this positivity was particularly evident in CH conditions for nonattended extreme tones, even though in comparison with the extreme distractors in ES conditions, the former were only half as distant in frequency from targets. Thus, opposite effects would have been predicted if physical similarity had an exclusive influence on the attention-related negativity, as proposed by the attentional trace model.

The results obtained in CH conditions are consistent with the possibility that subjects divided their attention between the relevant frequency and its nearest neighbor. Adopting this strategy would have reduced Nd_e amplitude and delayed Nd_e onset (Hink, Van Voorhis, Hilliard, & Smith, 1977; Okita, 1979; Parasuraman, 1980). In the present study, evidence for a division of attention between the attended frequency and the adjacent middle frequency in CH conditions is also compatible with the large number of FAs and the slowed RTs in CH conditions. Moreover, the ERPs to occasional deviant middle tones in the CH conditions showed a late positivity at Pz. Woods (1990) proposed that P3b in nonattended channels could be used as a converging method for evaluating channel separation. In the current study, the late positivity elicited by the deviant middle tones in CH conditions suggest that these tones received additional processing. Lastly, the scalp distribution of the attention effect in CH conditions differed from that in the ES and CE conditions. This is consistent with the hypothesis that the subjects used different cognitive strategies in CH conditions than they did in ES and CE conditions. Thus, it appears that divided attention could explain the difference in Nd onset observed in the present data.

The current data are consistent with previous reports (Alain et al., 1993) showing that attention effects on ERPs are sensitive to contextual information. Our results suggest that the attentional trace, like perception itself, must be sensitive to the relationships among and between auditory stimuli.

Implication for Theory of Auditory Perceptual Organization

Jones (1976; Jones & Boltz, 1989) proposed a theory for auditory perceptual organization founded upon rule-based predictability. This theory assumes that the auditory system automatically extracts regularities or patterns from a sequence of sounds. An attentional process then uses these regularities as rules to predict the next sounds in the sequence. From this perspective, auditory streaming occurs because the attentional mechanism cannot "move" rapidly enough in a multidimensional auditory space to catch tones that violate good-sequence rules. Although certain types of regularity may contribute to organizing the sounds into distinct perceptual units, this theory has difficulty explaining the effect of signal clustering when the auditory stimuli are presented in random order and at variable SOAs. The random tone

presentation made the recognition of a pattern very difficult, providing little information about the subsequent incoming sounds. The fact that clustering by frequency affects sensory processing of tones presented in random order implies the existence of a mechanism that allows random events to be grouped on the basis of their frequency similarity. This also suggests that neither temporal nor frequency predictability are necessary for auditory streaming. Our results appear to be consistent with the auditory scene analysis proposed by Bregman (1990), who proposed that auditory stimuli are grouped into distinct perceptual units on the basis of their similarity in location, frequency, and other characteristics.

Conclusions

In summary, our results show that signal clustering produces a coherent pattern of changes in performance and ERP signs of selective attention. The effect of clustering by frequency on both the behavioral and the physiological data is consistent with the segregation of the sequence into two distinct perceptual streams. The auditory stream segregation might provide additional perceptual cues that could be used concurrently with the physical features of the stimulus to guide further perceptual analysis of each incoming event. This may allow for a rapid categorization of the tones as being low or high in pitch and, hence, for both the early selection of relevant stimuli and the rejection of unwanted stimuli from further processing (Kahneman & Henik, 1981). The physiological data revealed that clustering can influence sensory processing at latencies as short as 100 msec following tone onset. Both the timing and the scalp distribution suggest that the clustering effect on sensory processing depends in part on modulation of early sensory ERPs generated in auditory cortex.

REFERENCES

- ALAIN, C., ACHIM, A., & RICHER, F. (1993). Perceptual context and the auditory selective attention effect on event-related brain potentials. *Psychophysiology*, **30**, 572-580.
- ALAIN, C., & WOODS, D. L. (1993). Distractor clustering enhances detection speed and accuracy during selective listening. *Perception & Psychophysics*, **54**, 509-514.
- ALHO, K., DONAUER, N., PAAVILAINEN, P., REINIKAINEN, K., SAMS, M., & NÄÄTÄNEN, R. (1987). Stimulus selection during spatial attention as expressed by event-related potentials. *Biological Psychology*, **24**, 153-162.
- ALHO, K., PAAVILAINEN, P., REINIKAINEN, K., SAMS, M., & NÄÄTÄNEN, R. (1986). Separability of different negative components of the event-related potential associated with auditory stimulus processing. *Psychophysiology*, **23**, 613-623.
- ALHO, K., TÖTTÖLÄ, K., REINIKAINEN, K., SAMS, M., & NÄÄTÄNEN, R. (1987). Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalography & Clinical Neurophysiology*, **68**, 458-470.
- 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.
- ALHO, K., WOODS, D. L., ALGAZI, A., & NÄÄTÄNEN, R. (1992). Intermodal selective attention: II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography & Clinical Neurophysiology*, **82**, 356-368.
- BERTRAND, O., PERRIN, F., & PERNIER, J. (1991). Evidence for a tonotopic organization of the auditory cortex observed with auditory evoked potentials. *Acta Oto-laryngologica*, **491**, 116-123.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA and London: MIT Press.
- BREGMAN, A. S., & RUDNICKY, A. I. (1975). Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 263-267.
- 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.
- DONALD, M. W. (1987). The timing and polarity of different attentional-related ERPs changes inside and outside of the attentional focus. In R. J. Johnson, Jr., J. W. Rohrbaugh, & R. Parasuraman (Eds.), *Current trends in event-related potentials research* (pp. 81-86). New York: Elsevier.
- GIARD, M. H., PERRIN, F., PERNIER, J., & PERONNET, F. (1988). Several attention-related wave forms in auditory areas: A topographic study. *Electroencephalography & Clinical Neurophysiology*, **69**, 371-384.
- HANSEN, J. C., & HILLYARD, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography & Clinical Neurophysiology*, **49**, 277-290.
- 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., & PICTON, T. W. (1988). Electrophysiology of cognition. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology: The nervous system V* (pp. 519-583). Bethesda, MD: American Physiological Society.
- HINK, R. F., VAN VOORHIS, S. T., HILLYARD, S. A., & SMITH, T. (1977). The division of attention and the human auditory evoked potentials. *Neuropsychologia*, **15**, 597-605.
- JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, **83**, 323-355.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.
- JONES, M. R., KIDD, J. G., & WETZEL, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1059-1073.
- KAHNEMAN, D., & HENIK, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181-209). Hillsdale, NJ: Erlbaum.
- MAGLIERO, A., BASHORE, T. R., COLES, M. G., & DONCHIN, E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology*, **21**, 171-186.
- MCCARTHY, G., & WOOD, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography & Clinical Neurophysiology*, **62**, 203-208.
- MICHIE, P. T., SOLOWJ, N., CRAWFORD, J. M., & GLUE, L. C. (1993). The effects of between-source discriminability on attended and unattended auditory ERPs. *Psychophysiology*, **30**, 205-220.
- NÄÄTÄNEN, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral & Brain Sciences*, **13**, 201-288.
- NÄÄTÄNEN, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- NÄÄTÄNEN, R., GAILLARD, A., & MÄNTYSALO, S. (1978). The N1 effect of selective attention reinterpreted. *Acta Psychologica*, **42**, 313-329.
- NÄÄTÄNEN, R., PAAVILAINEN, P., TIITINEN, H., JIANG, D., & ALHO, K. (1993). Attention and mismatch negativity. *Psychophysiology*, **30**, 436-450.
- 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 location. *Biological Psychology*, **9**, 271-284.
- ONISHI, S., & DAVIS, H. (1968). Effects of duration and rise time of

- tone bursts on evoked V potentials. *Journal of the Acoustical Society of America*, **44**, 582-591.
- PARASURAMAN, R. (1980). Effects of information processing demands on slow negative shift latencies and N100 amplitude in selective and divided attention. *Biological Psychology*, **11**, 217-233.
- 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**, 71-81.
- WOLDORFF, M. G., GALLAN, C. C., HAMPSON, S. A., HILLYARD, S. A., PANTEV, C., SOBEL, D., & BLOOM, F. E. (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., 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 processing during selective listening to rapidly presented tones. *Electroencephalography & Clinical Neurophysiology*, **79**, 170-191.
- WOLPAW, J. R., & PENRY, J. K. (1975). A temporal component of the auditory evoked response. *Electroencephalography & Clinical Neurophysiology*, **39**, 609-620.
- WOOD, C. C., & WOLPAW, J. R. (1982). Scalp distribution of human auditory evoked potentials. II. Evidence for overlapping sources and involvement of auditory cortex. *Electroencephalography & Clinical Neurophysiology*, **54**, 25-38.
- WOODS, D. L. (1990). The physiological basis of selective attention: Implications of event-related potential studies. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson, Jr. (Eds.), *Event-related brain potentials: Basic issues and applications* (pp. 178-209). New York: Oxford University Press.
- WOODS, D. L. (in press). The component structure of the N1 wave of the human auditory evoked potentials. *Electroencephalography & Clinical Neurophysiology* (Supp.).
- WOODS, D. L., & ALAIN, C. (1993). Feature processing during high-rate auditory selective attention. *Perception & Psychophysics*, **53**, 391-402.
- WOODS, D. L., ALHO, K., & ALGAZI, A. (1992). Intermodal selective attention: I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography & Clinical Neurophysiology*, **82**, 341-355.
- WOODS, D. L., ALHO, K., & ALGAZI, A. (1994). Stages of auditory feature conjunction: An event-related brain potential study. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 81-94.
- WOODS, D. L., & CLAYWORTH, C. C. (1985). Click spatial position influences middle latency auditory evoked potentials (MAEPs) in humans. *Electroencephalography & Clinical Neurophysiology*, **60**, 122-129.
- ZEEF, E. J., & KOK, A. (1993). Age-related differences in the timing of stimulus and response processes during visual selective attention: Performance and psychophysiological analyses. *Psychophysiology*, **30**, 138-151.

NOTES

1. The ERPs are voltage changes in the ongoing electroencephalogram (EEG) that are time locked to stimulus events. The scalp-recorded ERPs to auditory signals consist of a series of positive (P) and negative (N) voltage peaks that occur with specific time delays. On the basis of their latencies, auditory ERPs are divided into three groups: (1) auditory brain stem responses occurring from stimulus onset to 10 msec poststimulus; (2) middle auditory evoked potentials occurring between 10 and 60 msec poststimulus (Na, Pa, Nb, and P1); and (3) long-latency auditory evoked potentials, including the N1 (80-120 msec), P2 (160-220 msec), N2 (160-250 msec), and P3 (250-600 msec) deflections.
2. Since this comparison involved ERP subtractions between conditions varying in difficulty, the differences may have been confounded with differences in overall arousal (see Näätänen, 1992). However, in the current experiment, arousal-related differences were likely to be small, since the speeded-RT task remained subjectively difficult and demanding even in CE conditions.
3. The MMN is obtained by subtracting ERPs to occasional tones of longer duration from ERPs to frequent tones of shorter duration for the attended tones and for the middle and extreme distractors. For example, in ES conditions, the MMN is obtained by subtracting Tone A* from Tone A (see Figure 1, top).

(Manuscript received December 6, 1993;
revision accepted for publication May 10, 1994.)