

Feature processing during high-rate auditory selective attention

DAVID L. WOODS and CLAUDE ALAIN
University of California, Davis, California
and V.A. Medical Center, Martinez, California

Auditory event-related brain potentials (ERPs) and reaction times were analyzed in a selective attention task in which subjects attended to tone pips presented at high rates (interstimulus intervals [ISIs] of 40-200 msec). Subjects responded to infrequent target tones of a specified frequency (250 or 4000 Hz) and location (left or right ear) that were louder than otherwise identical tones presented randomly to the left and right ears. Negative difference (Nd) waves were isolated by subtracting ERPs to tones with no target features from ERPs to the same tones when they shared target location, frequency, or both frequency and location cues. Nd waves began 60-70 msec after tone onset and lasted until 250-350 msec after tone onset, even for tones with single attended cues. The duration of Nd waves exceeded the ISIs between successive tones, implying that several stimuli underwent concurrent analysis. Nd waves associated with frequency processing had scalp distributions different from those associated with location processing, implying that the features were analyzed in distinct cortical areas. Nd waves specific to auditory feature conjunction were isolated. These began at latencies of 110-120 msec, some 30-40 msec after the Nds to single features. The relative timing of the different Nd waves suggests that auditory feature conjunction begins after a brief parallel analysis of individual features but before feature analysis is complete.

Feature conjunction is a fundamental operation underlying the perception of visual and auditory signals (Treisman & Gelade, 1980). According to the feature integration model (Treisman & Gelade, 1980), features are initially processed in parallel, followed by a stage of feature integration. Feature integration involves serial processing and focal attention (Briand & Klein, 1987). Although feature integration has been studied largely in the visual modality, certain similar effects—for example, conjunction errors—have been observed for auditory signals (Treisman, 1988). For example, subjects may erroneously combine the frequency of one tone with the location of another (Yund & Efron, 1974).

Little is known about the neurophysiological substrates of feature conjunction. Although it is tempting to associate the analysis of individual features with the activation of distinct, feature-specific cortical sensory fields, feature dimensions do not always correspond with known cortical specializations (Treisman & Sato, 1990). Specialized cortical fields can code biologically significant fea-

ture conjunctions in certain species (Suga, Niwa, Taniguchi, & Margoliash, 1987), and there has been some speculation that speech signals might benefit from such specialized coding in humans (Lieberman & Mattingly, 1989). However, fields for the conjunction of particular features cannot account for the more general case in which arbitrary groups of features can become conjoined depending on instructions given to the subject.

Information about the timing of feature conjunction operations and their location in the brain is available through the recording of event-related brain potentials (ERPs). Previc and Harter (1982; see also Harter, Aine, & Schroeder, 1982) were the first to use ERP techniques to investigate feature conjunction. Subjects responded to visual targets defined by the conjunction of the orientation and spatial frequency of bar gratings (e.g., small vertical bars). Previc and Harter found small ERP attention effects when they used stimuli with a single relevant feature (e.g., large vertical bars or small horizontal bars). However, attention effects were larger, and of longer duration, when the stimuli had both relevant features. When the small effects of isolated features were subtracted from the larger target effects, a difference wave remained that the authors concluded was conjunction-specific. However, in Previc and Harter's study, N2b and P3 components associated with target detection and premotor potentials linked to the response occurred following presentation of the targets. Hence, "conjunction-specific" components might have reflected these postfeature conjunction operations rather than feature conjunction itself.

Special thanks to John Lackey and Greg Shenaut for software development that made these studies possible, to Diane Swick for critical readings of earlier versions of this manuscript, and to Alain Algazi, Oren Zaidel, and Diego Covarrubias for assistance in various phases of the experiment. This research was supported by research grants from the NIDCD and the V.A. Research Service to D.L.W., and by a FCAR postdoctoral fellowship to C.A. Correspondence should be addressed to D. L. Woods, Chief, Clinical Neurophysiological Laboratory, Neurology Service (127), V.A. Medical Center, 150 Muir Road, Martinez, CA 94553 (e-mail: dlwoods@ucdavis.edu).

Studies of auditory feature conjunction have revealed shorter latency conjunction-specific effects. These can be isolated as negative difference (Nd) waves (Hansen & Hillyard, 1980) by subtracting ERPs obtained with the use of nonattended sounds from the ERPs obtained when the same sounds were attended. This subtraction procedure reveals at least two distinct Nd components, an early negative difference wave (Nd_e) beginning at latencies of 60 msec or less, and a late negative difference wave (Nd_l) that can persist for hundreds of milliseconds (for reviews, see Woods, 1990, and Näätänen, 1990). In studies of auditory feature conjunction, subjects attend to tones of a particular location and frequency (L+F+) in order to respond to hard-to-detect targets differing in intensity or duration (Hansen & Hillyard, 1983). Nd waves are isolated by subtracting ERPs elicited by standards sharing no features with the target (L-F-) from ERPs elicited by nontarget L+F+ tones, isolating an Nd-both wave associated with the analysis of both cues. However, Nd waves can also be extracted when tones share only location (Nd-locat to L+F- tones) or frequency (Nd-freq to L-F+ tones) features with the target (Alho, Sams, Paavilainen, Reinikainen, & Näätänen, 1989). Nd-freq and Nd-locat waves have different scalp distributions (Woods, Algazi, & Alho, 1991), implying that frequency and location cues are analyzed in different cortical regions.

Conjunction-specific Nd (Nd-conj) waves can be isolated by subtracting Nd-freq and Nd-locat waves from Nd-both waves. This process reveals conjunction-specific attention effects in ERPs elicited by nontarget tones at latencies as short as 120 msec (Woods et al., 1991). The short onset latencies of Nd-conj waves argues against models in which feature conjunction occurs only after the complete analysis of individual features.

ERP experiments also allow feature conjunction to be studied at rates of stimulus delivery too high to permit subjects to respond overtly to each stimulus. Results from high-rate experiments might be different from results obtained with more traditional approaches. For example, in visual search tasks, the interval between target repetitions may be 10 sec or more. This interval may be too long to permit the maintenance of a clear sensory template. The automatic processing of individual features might be an adaptive strategy, invoked because targets with conjoined features occur so infrequently. In the present experiment, stimuli were presented at high rates so that tones with conjoined (L+F+) features recurred at relatively short interstimulus intervals (ISIs).

The long duration of Nd-freq and Nd-locat waves in previous ERP studies (Woods et al., 1991) is also surprising. Indeed, the duration of Nd waves elicited by individual features has sometimes been longer than the ISIs between stimuli. Insofar as the duration of Nd waves reflects the duration of feature processing, this implies that the processing of the features of one tone might continue even after the delivery of the next tone. The possibility that successively presented tones would undergo

concurrent processing was assessed more explicitly in the current experiment.

A subsidiary aspect of the present study was the examination of the attentional modulation of middle-latency auditory evoked potentials (MAEPs). Previous reports have suggested that MAEPs may be modulated by attention when difficult tasks and highly discriminable stimuli are used (Woldorff & Hillyard, 1991). In the present experiment, we used demanding tasks and highly discriminable location and frequency cues to determine whether MAEP modulation could be associated with the processing of either frequency or location features.

METHOD

Subjects

Twelve subjects (6 females, 6 males; age range, 18–28) participated in the experiment. All were right handed, with normal hearing. Two additional subjects were excluded because of excessive EMG and EOG artifacts. The subjects were paid \$10.00/h for their participation, and they gave informed consent according to UC Davis and V.A. Medical Center guidelines.

Stimuli and Task

A dichotic attention task was performed, with four different standard and four different deviant tones presented in a randomized sequence at a high rate (Figure 1). The standard tones were 250 or 4000 Hz, with 10-msec durations and 2-msec linear rise/fall times. They were presented in the right or left ear at 86/87 dB SPL over broadband binaural masking noise (58 dB SPL). The loudness levels



Figure 1. Subjects attended to monaural tone pips and responded to occasional target tones of a specified frequency and location that were slightly louder than the corresponding standards. Representative stimulus sequences are shown during the condition "attend low frequency right ear." h, high frequency (4000 Hz); l, low frequency (250 Hz); H, high-frequency deviant; L, low-frequency deviant; L*, attended target; RT, buttonpress response.

of the two tones had been matched prior to the study by four listeners with normal hearing. After the experiments, more extensive psychophysical studies indicated that the 4000-Hz tone had been louder (5–6 dB mismatch) than the 250-Hz tone (Woods, Alain, Covarrubias, & Zaidel, in press).

The deviant tones were 3–15 dB more intense than the corresponding standards and were presented infrequently (overall probability, 2.4%). The intensity differences between the deviants and corresponding standards were adjusted on a subject-by-subject basis so that targets were detected with 80%–90% accuracy. These small intensity differences were barely perceptible. As a result, the subjects only rarely noted deviant tones of the nonattended frequency or in the nonattended ear. Detection performance was continuously monitored throughout the experiment, and target intensity increased or decreased between blocks to maintain detection performance at 80%–90% throughout the experiment. The subjects were instructed to attend to a combination of frequency and ear of delivery (e.g., "low-frequency tones on the right") and to ignore competing stimuli (see Figure 1). Their task was to respond only to occasional deviant tone pips ("targets") at the specified location and frequency but of higher intensity (L+F+I+), while withholding responses to all other stimuli.

The tones were presented rapidly: ISIs ranged from 40 to 200 msec ($M = 120$ msec) in 10-msec increments according to a rectangular distribution. Unconstrained, randomized scenarios controlled the delivery of tones to both ears so that two tones were never delivered simultaneously (i.e., all tones were monaural). To evaluate the contribution of ERPs from preceding stimuli overlapping with those of the current epoch, tones were omitted on 20% of the trials. This permitted averages to be obtained when no signal was present, so that the overlap from preceding and following trials could be directly visualized.

The subjects were first given several training blocks. The EEG was then recorded during 16 stimulus blocks, derived from 4 replications of each of the 4 combinations of 2 frequencies and 2 locations (i.e., attend 250 Hz left, attend 4000 Hz left, attend 250 Hz right, and attend 4000 Hz right). Each block lasted 6.67 min and contained 3,330 stimuli. In all, each subject listened to 53,280 tone pips over the course of the experiment. Behavior was continuously monitored, and the subjects were encouraged to be fast and accurate in their responses. Feedback about hit rate, reaction time (RT), false alarms, and misses was provided after each block.

EEG Recording

The EEG (bandpass 1.0–300 Hz) was recorded from 28 electrodes over the scalp (Fp1, 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, O2). Vertical and horizontal eye movements were recorded from electrodes lateral and superior to the left eye. All electrodes were referenced to four interconnected, EKG-balanced electrodes at the base of the neck (Woods & Clayworth, 1985).

Data Analysis

The EEG was continuously digitized at 833 Hz/channel and ERPs were extracted off line by computer. Trials contaminated by vertical or horizontal eye movements (in excess of 100–120 μ V), excessive peak-to-peak deflections, amplifier clipping, or bursts of electromyographic activity were excluded from the average. For the analysis of short-latency components, averages included 50 msec of prestimulus activity and 257 msec of poststimulus activity.

Longer latency components were evaluated after decimating the EEG by averaging together three successive sample points in the original digitized EEG to reduce the effective sampling rate to 278 Hz. For longer latency components, the analysis included 100 msec of prestimulus activity and 822 msec of poststimulus activity.

Following averaging, the ERPs were digitally smoothed to eliminate frequencies above 130 Hz for short-latency components, and 40 Hz for long-latency components. ERPs were quantified under software control. Measurements were obtained relative to the prestimulus baselines. Peak voltage and latency measures were gathered for descriptive purposes. The effects of attention were evaluated on mean voltages obtained at 5-msec intervals from 10 to 50 msec, at 10-msec intervals from 50 to 140 msec, and at 20-msec intervals from 140 to 460 msec. The results were statistically evaluated with analyses of variance for repeated measures. The effects of attention were analyzed first at the Fz and Cz electrodes only. The scalp distributions were examined from 23 electrodes (excluding eye and preauricular electrodes and the electrode on the nose).

Type I errors associated with inhomogeneity of variance were controlled by reducing the degrees of freedom associated with individual electrodes and electrode \times subject interactions according to a modified Greenhouse-Geisser procedure. Previous studies with 5–7 scalp electrodes have established that the degrees of freedom should be reduced by 40%–70% (Ruchkin, 1990). Since we used a larger number of scalp electrodes, a more stringent epsilon correction factor (0.1) was adopted. For example, scalp distribution interactions, usually with 23 scalp electrodes and 12 subjects [$F(22,242)$], would be reported as significant only if F ratios exceeded that required at 2 and 24 degrees of freedom ($F = 3.40$, at $p < .05$). When the scalp distributions were compared, the data were normalized (McCarthy & Wood, 1985) to equate voltages at the Fz electrode.

RTs were also quantified by computer. Correct responses were defined as buttonpresses occurring 200–1,000 msec after target onset. Responses outside this window were classified as false alarms. Because of the small number of deviant tones, it was impossible to obtain deviant ERPs with high signal/noise ratios. As a result these data were not examined in detail and will not be discussed further.

RESULTS

Behavioral Results

Preliminary testing showed that subjects were more sensitive in detecting intensity increments in the 4000-Hz tone pip ($M = 7.04 \pm 1.19$ dB) than in the 250-Hz tone pip [11.89 ± 1.55 dB, $t(11) = 23.14$, $p < .001$], which was consistent with previous reports (Scharf & Houtsma, 1986). However, since the intensity increments had been adjusted to equate discriminability, no difference in accuracy was found across conditions (the mean detection rates were, for 250 Hz, 86.3%, and for 4000 Hz, 84.3%). Furthermore, subjects in different conditions produced similar false alarm rates (for 250 Hz, 10.9%; for 4000 Hz, 9.7%) and comparable RTs (464 and 458 msec for 250 and 4000 Hz, respectively). The ear of stimulation had no significant effect on either RT or response accuracy.

Figure 2 shows the effects on RTs produced by tones preceding the target. In these analyses, the effects of preceding tones were obtained by averaging over the different intervening stimuli that preceded the target. For example, stimuli in the –400 to –200 msec range typically occurred either two or three positions before the target (T–2 or T–3). These tones produced a pattern similar to previous results (Woods et al., 1991): RTs were fastest following L+F+ and L–F+ tones, intermediate follow-

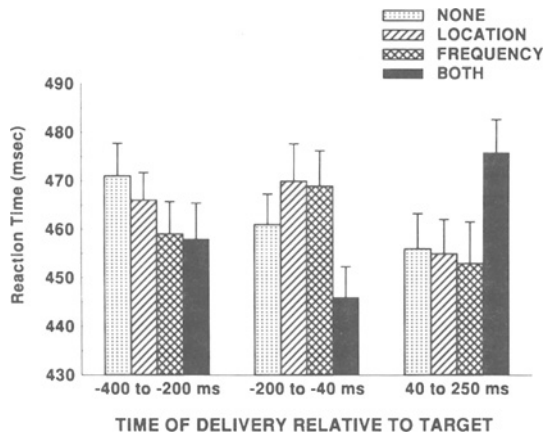


Figure 2. Reaction times to correctly detected targets are shown as a function of the features shared by the target and the preceding and following standard tones. Latency ranges -400 to -200 msec (left); -200 to -40 msec (center); and $+40$ to $+250$ msec (right). NONE = L-F-; LOCATION = L+F-; FREQUENCY = L-F+; BOTH = L+F+. Priming effects at a given interval were averaged over random variations in tone pips that preceded and followed the priming tone. Error bars show standard errors.

ing L+F- tones, and slowest following L-F- stimuli (although the results failed to reach statistical significance). Tones in the -200 to -40 msec range produced more complex changes: RTs were speeded following L+F+ stimuli, in comparison with RTs following L+F-, L-F+, and L-F- tones [$F(3,33) = 5.04, p < .01$]. Specific comparisons showed that only the L+F+ tones caused significant priming. Tones at -120 to -40 msec before the target did not produce significant RT effects. Nor were significant effects seen as a function of stimuli at -800 to -600 msec, or -600 to -400 msec.

The standard tones that followed the target ($+40$ to $+250$ msec) also had a significant effect on RT. Attended standards (L+F+) slowed the RTs to preceding targets in comparison with RTs following L+F-, L-F+, and L-F- tones [$F(3,30) = 4.09, p < .02$; 1 subject, who lacked RTs in one category, was excluded from this analysis].

ERPs: Sensory Responses

Figure 3 shows the ERPs elicited by 250- and 4000-Hz standard tones sharing neither frequency nor location (L-F-) with the attended target. ERPs to nonattended tones were characterized by polyphasic short-latency deflections including wave V of the brainstem ERP (mean peak latency over tone pips of both frequencies = 7.1 msec) and Na (20.0 msec), Pa (29.5 msec), and Nb (38.2 msec) components of the middle latency auditory ERP. These were followed by long-latency P1 (52.6 msec) and N1 (101 msec) components. Peak latencies of ERP components elicited by the 250-Hz tones were longer than those elicited by 4000-Hz tones (Figure 3),

and the differences increased with component latency (see Woods, Alain, et al., in press, for a further discussion). Na and Pa components were symmetrically distributed over the frontocentral scalp. The P1 was also symmetrical over the two hemispheres (e.g., at ipsilateral and contralateral central sites, $C_i = 0.62 \mu V$, $C_c = 0.62 \mu V$). Unlike the middle-latency components, the P1 showed a significant electrode \times frequency interaction [$F(22, 242) = 5.13, p < .02$], due primarily to the fact that the P1 to 4000-Hz tones was more parietally distributed than the P1 to 250-Hz tones. The N1 was larger contralateral to the stimulated ear [$F(1, 11) = 42.2, p < .01$] and tended to be more frontally distributed for 4000-Hz than for 250-Hz tones, but the tonotopic differences failed to reach statistical significance (cf. Woods et al., 1991).

ERPs: Short-Latency Effects of Attention

Attention exerted systematic effects on ERPs that depended on the features shared between the attended target and the ERP-evoking stimulus. These effects are shown in Figure 4 for 4000-Hz standard tones.

Figures 5 and 6 show the effects of attention on ERPs. There was no evidence of attentional modulation during the MAEP latency range, regardless of the features shared with the attended target tone. Specific comparisons of peak amplitudes and peak latencies for each of the MAEP components showed no attention-related differences. Nor were there significant differences in mean voltages between L-F- tones and tones with either or both attended features, in 10 different measures obtained at 5-msec intervals from 5 to 50 msec, or in mean voltage measures from 20 to 50 msec.

The effects of attention began at about 60 msec and continued until 300-350 msec. Attention effects were isolated by subtracting ERPs to nonattended standards (L-F-) from ERPs to same stimuli when one or both of the features were shared with the target. Three sets of difference waves were formed: Nd-freq (L-F+ minus L-F-, Figure 5), Nd-locat (L+F- minus L-F-, Figure 5), and Nd-both (L+F+ minus L-F-, Figure 6). A conjunction-specific difference wave (Nd-conj) was also isolated by subtracting Nd-freq plus Nd-locat from Nd-both waves (Figure 6).

Attention exerted its earliest effects at Fz [at 60-70 msec, $F(1,11) = 4.80, p < .05$; see Figures 5 and 7). Early attention effects were evident for tones (L-F+ and L+F+) sharing the frequency of the attended target. Nd-freq waves reached significance at 70-80 msec [$F(1,11) = 8.44, p < .02$] and continued at statistically significant levels from 80 to 160 msec [$F(1,11)$ between 14.9 and 28.2, $p < .01$, for all comparisons]. The negative phase of Nd-freq waves was followed by a significant positive enhancement at 200-220 msec [$F(1,11) = 9.56, p < .01$]. No changes were found in the distribution of the Nd-freq waves as a function of the frequency of the attended tone.

Nd-locat waves were smaller in amplitude, more central in distribution, and delayed in latency in comparison with Nd-freq waves (Figure 7). Nd-locat waves ap-

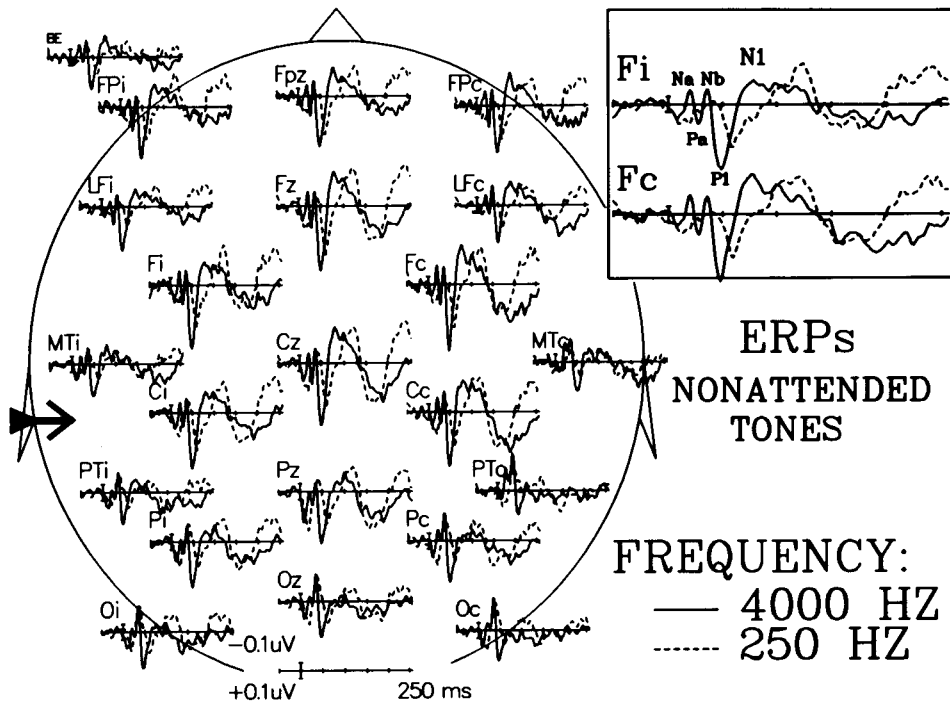


Figure 3. ERPs to 250-Hz (dotted) and 4000-Hz (solid) tone pips sharing neither the location nor the pitch with the attended target (L-F-). BE, below eye; FP, frontopolar; LF, lateral frontal; F, frontal; MT, mid-temporal; C, central; PT, posterior temporal; P, parietal; O, occipital. Averaged over ear of stimulation with electrodes transposed so that those on the right of the figure were contralateral to the stimulated ear. The electrodes have been relabeled accordingly: i, ipsilateral; c, contralateral; z, midline. ERPs to null stimuli have been subtracted from the ERPs shown to correct for ERP overlap. Grand mean of 12 subjects. Insert: Enlarged ERPs from ipsilateral and contralateral frontal sites.

proached statistical significance at 90–100 msec and 100–110 msec [at Cz, $F(1,11) = 3.11$, $p < .15$ and 3.94 , $p < .10$, respectively], reached significance at 110–120 msec [$F(1,11) = 6.14$, $p < .03$, tended to significance at 120–130 msec [$F(1,11) = 4.45$, $p < .06$], and were again significant at 130–140 msec [$F(1,11) = 8.92$, $p < .01$]. The late positive phase of the location effect also reached significance at 180–200 msec [$F(1,11) = 5.8$, $p < .05$]. Like Nd-freq waves, no changes in the distribution of the Nd-locat waves were found as a function of the frequency of the attended tone.

A specific comparison of the scalp distributions of Nd-locat and Nd-freq waves was performed at the latency range at which their mean voltages were maximal (130–140 msec). As in previous studies (Woods et al., 1991), the distributions differed—Nd-locat waves had central distributions, posterior to those of Nd-freq waves [$F(22,242) = 4.31$, $p < .03$; see Figure 5].

Nds to tones with both attended features (Nd-both) were larger in amplitude than the Nds to either individual cue (Figure 7). Like Nd-freq waves, Nd-both waves had onset latencies of 60–70 msec, reached maximal negativity at 130–160 msec over frontocentral (Fz and Cz) areas, and showed a subsequent positive phase between 180 and

260 msec. No changes in the distribution of the Nd-both waves were found as a function of the frequency of the attended tone. Nd-conj waves (Figure 6) were obtained by subtracting Nd-freq and Nd-locat waves from Nd-both waves. Nd-conj waves had central distributions not significantly different from the average of Nd-freq and Nd-locat distributions.

The time course of the processing of the individual and conjoined features is shown in Figure 7. Nd-freq and Nd-both waves began earliest, followed by Nd-locat waves. Initially, Nd-freq and Nd-both waves had nearly identical amplitudes. As a result, Nd-conj waves were delayed by 30–40 msec relative to the onset of Nd-freq and Nd-both waves. The Nd-conj peak was also delayed (to 150–180 msec). Nd-conj waves reached significance from 100 to 140 msec [$F(1,11)$ ranged from 6.70 to 16.5, $p < .05$ for all comparisons], and tended toward significance at 140- to 160- and 160- to 180-msec intervals [$F(1,11) = 4.65$ and 4.49, $p < .06$, in both cases].

ERPs: Long-Latency Effects of Attention

Long-latency Nd waves are shown in Figures 8 and 9, derived from decimated EEG data. Although Nd_i waves were reduced because of the bandpass of the recordings

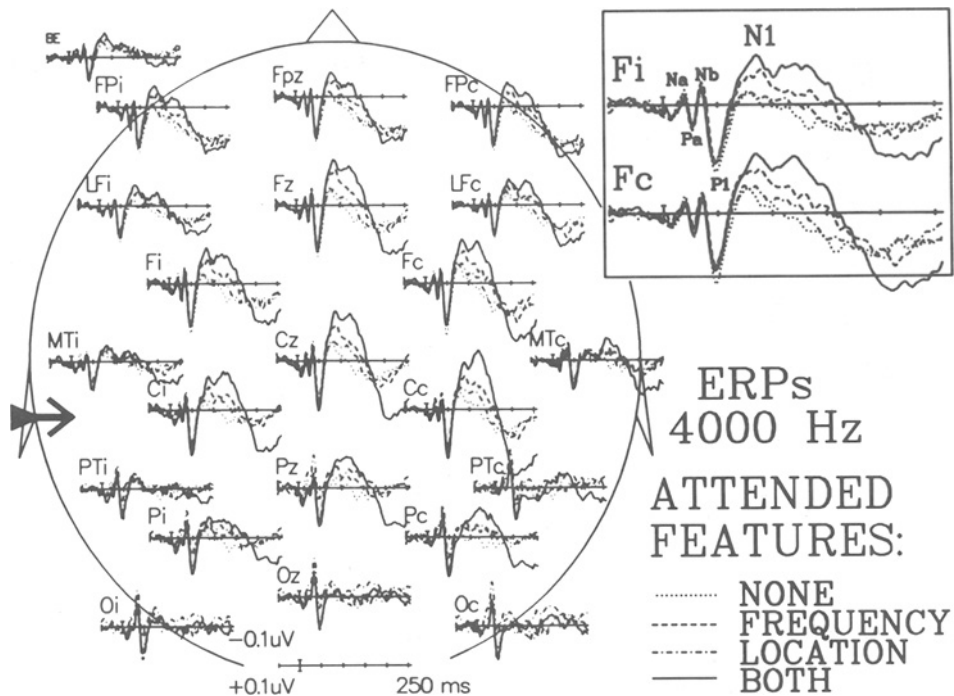


Figure 4. ERPs to 4000-Hz tone pips sharing no cues (dotted), frequency (dashed), location (dot-dash), or both frequency and location (solid) cues with attended targets. Averaged over ear of stimulation. The electrodes have been transposed so that the electrodes on the right of the figure were contralateral to the stimulated ear, and relabeled accordingly: i, ipsilateral; c, contralateral. Grand mean of 12 subjects.

(optimized for MAEP components), attention produced several significant effects on ERPs at latencies of 300 msec or more. Both Nd-locat and Nd-freq waves (Figure 8) showed positivities at 230–260 msec, negativities at 280–350 msec, and subsequent positivities. The late negative portion of the Nd-locat wave reached significance at 280–300 msec [mean voltage, $-0.12 \mu\text{V}$, $F(1,11) = 7.84$, $p < .02$], and tended toward significance at 300–320 msec [$F(1,11) = 3.81$, $p < .08$]. The positive phase of Nd-freq waves reached significance at 360–380 msec [at Fz $0.18 \mu\text{V}$, $F(1,11) = 6.56$, $p < .03$] and at 380–400 msec [$F(1,11) = 13.02$, $p < .005$], with a trend also evident in the 400- to 420-msec interval [$F(1,11) = 4.23$, $p < .07$]. Long-latency Nd-both and Nd-conj waves are shown in Figure 9. These also contained a late positive-negative-positive deflection. Nd-both showed an enhanced positivity at 240–260 msec [$F(1,11) = 9.39$, $p < .02$], while Nd-conj waves showed a positive phase at 240–260 msec [$F(1,11) = 6.52$, $p < .03$], followed by a later negativity at 360–380 msec [$F(1,11) = 5.09$, $p < .05$].

DISCUSSION

Sensory Components of Auditory ERPs

The present experiment elucidated several characteristics of middle and long-latency auditory evoked poten-

tials. In addition to wave V of the brainstem auditory evoked potential, at least three components, the Na, Pa, and Nb, were reliably recorded during the first 40 msec after tone delivery. These were followed by more prominent P1 (mean latency, 52.6 msec), N1, and P2 waves.

Auditory evoked potentials were small in amplitude, suggesting that the masking noise had been effective in limiting excitation to frequency-specific regions of the cochlea. MAEPs and the N1 showed frequency-related differences in component latency. Frequency-related latency differences are a reliable finding when masking noise is used to limit the frequency splatter in the cochlea (Woods, Alho, & Algazi, in press-a). They reflect frequency-related differences in the speed of auditory processing: high-frequency tones are processed more rapidly than low-frequency tones (see Woods, Alain, et al., in press, for a further discussion).

The P1 was more posteriorly distributed after 4000-Hz tones than after 250-Hz tones. This suggests that the P1 was generated in a tonotopically organized structure. However, unlike what was observed in our previous studies, the rising phase of the N1 failed to show tonotopic differences in distribution (Woods et al., 1991). There are two possible explanations. First, generators contributing to the N1/P90 may have long refractory periods, and so, they might have been suppressed at the short ISIs used in the present experiment. Alternatively, they may

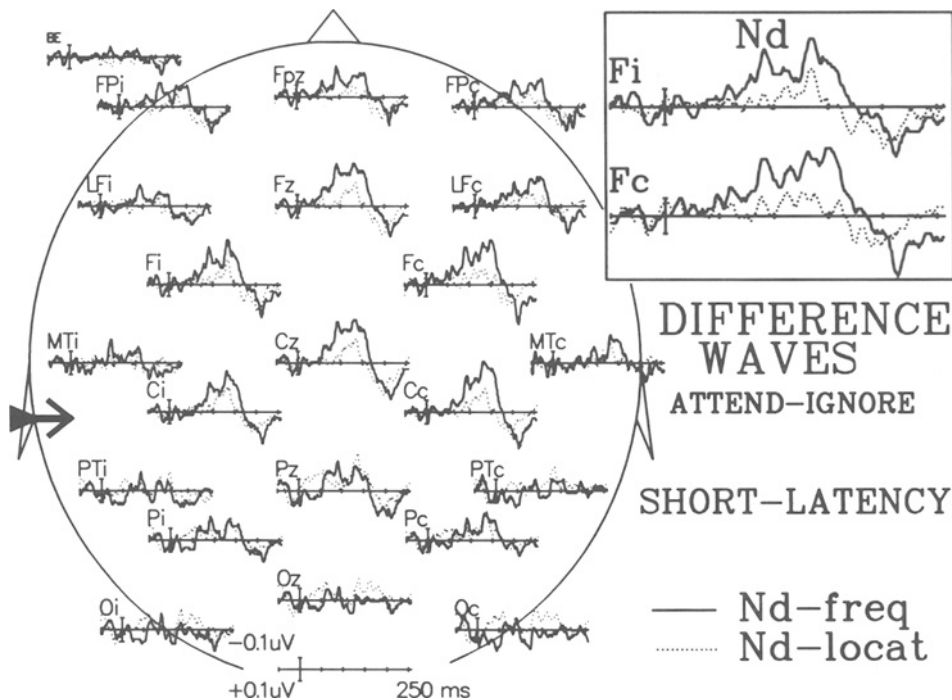


Figure 5. Difference waves associated with the processing of frequency (solid) and location (dotted) cues. Attention effects were isolated by subtracting ERPs to tone pips sharing no cues with the target (L-F-) from ERPs to tone pips sharing frequency (L-F+) or location (L+F-). Averaged over tone frequency and ear of stimulation. The electrodes have been transposed so that the electrodes on the right of the figure were contralateral to the stimulated ear and relabeled accordingly: i, ipsilateral; c, contralateral. Grand mean of 12 subjects.

have long temporal integration times and thus may have required tone bursts with durations longer than 10 msec. Both explanations are consistent with the small N1 amplitudes in the present study. Mean N1 voltages were reduced by about 80% in comparison with amplitudes in previous studies (Woods et al., 1991; Woods et al., in press-a).

Finally, previous studies have shown that attention can produce small positive enhancements in MAEPs (the P20-50; Woldorff & Hillyard, 1991). In the present experiment, despite the distinctiveness of attended and nonattended inputs and the high signal/noise ratio of the ERPs, attention showed no tendency to alter either the amplitudes or latencies of any MAEP component. Nor were MAEP mean voltages (including those from 20 to 50 msec) influenced by attention, regardless of whether tones shared frequency, location, or both frequency and location cues with the designated target. The P20-50 has been recorded when attended and nonattended tones differ in redundant frequency and location features—that is, a sequence of high-frequency tones in one ear is mixed with a sequence of low-frequency tones in the other. In these cases, a single discrete “stream” of nonattended signals can be identified. Such perceptual grouping may facili-

tate the rejection of nonattended signals as a group (Alain, Achim, & Richer, in press; Alain & Woods, in press). In the present experiment, subjects were faced with the more complex task of ignoring at least three different streams of auditory input at the same time.

The Effects of Stimulus Sequence on Reaction Times

In previous studies, a speeding of reaction times (RT priming) has been observed with the use of targets that follow tones sharing target features (Alain & Woods, in press). For example, in auditory selective attention tasks, RTs are speeded with targets that follow tones with the frequency, location, or both frequency and location features of the targets in comparison with tones with no target features (Woods, Alho, & Algazi, in press-b). One explanation is that a template against which target tones are compared is reinforced by repetitions of any target feature. In previous studies, priming effects have been evident for tones occurring from 200 to 400 msec before target delivery (-200 to -400 msec); but not for tones at -400 to -800 msec. Similar but nonsignificant effects were noted in the present experiment, and RT priming was also seen following tones delivered at -200 to

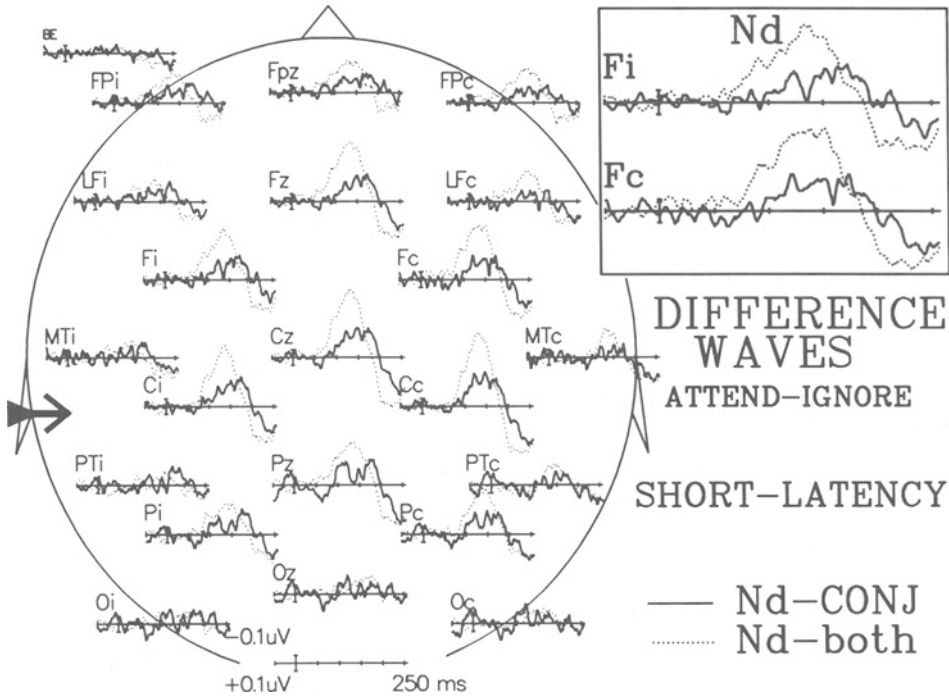


Figure 6. Dotted lines: Selective attention effects of tone pips with conjoined features, isolated by subtracting ERPs to L-F- standards from ERPs to L+F+ standards. Solid lines: Conjunction-specific selective attention effects (Nd-conj) derived by subtracting the sum of Nd-freq and Nd-locat waves (Figure 5) from Nd-both waves. Averaged over tone frequency and ear of stimulation. The electrodes have been transposed so that the electrodes on the right of the figure were contralateral to the stimulated ear and relabeled accordingly: i, ipsilateral; c, contralateral. Grand mean of 12 subjects.

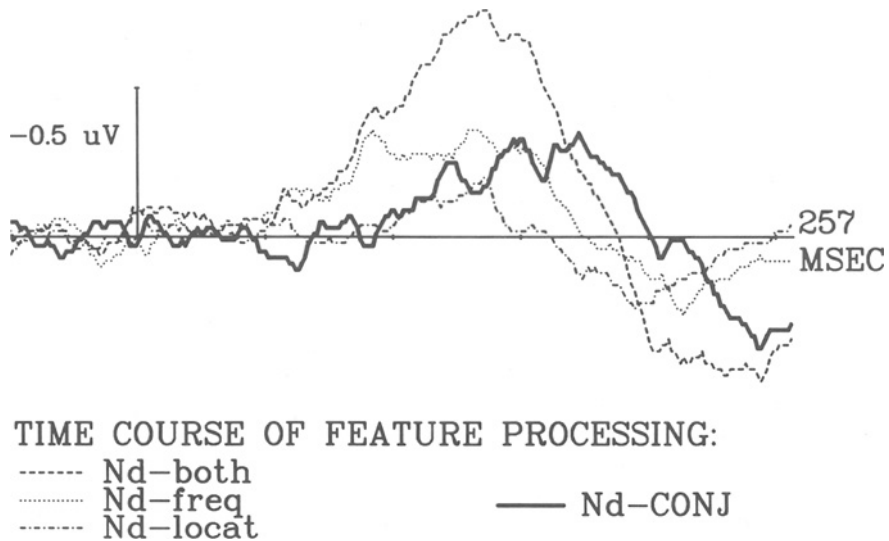


Figure 7. Nd waves at the Fz electrode showing the time course of feature conjunction effects. Averaged over tone frequency and ear of stimulation. Grand mean of 12 subjects.

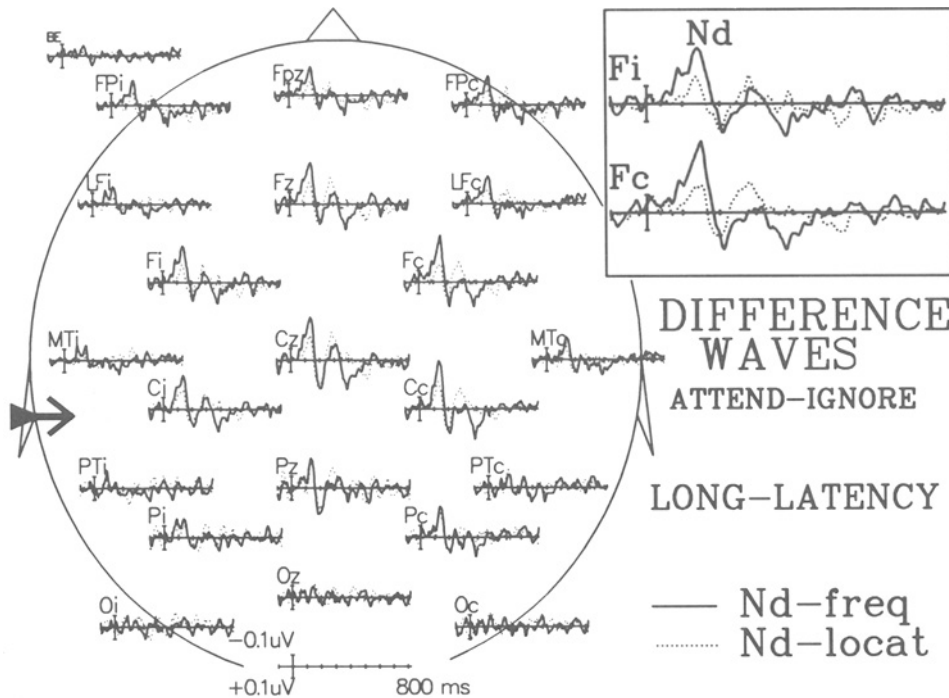


Figure 8. Nd-freq (solid) and Nd-locat (dotted) difference waves shown on a long-latency time base. Averaged over tone frequency and ear of stimulation. The electrodes have been transposed so that the electrodes on the right of the figure were contralateral to the stimulated ear and relabeled accordingly: i, ipsilateral; c, contralateral. Grand mean of 12 subjects.

–120 msec with respect to target onset. However, no priming was seen during the 120 msec that immediately preceded target delivery. This suggests that cortical elaboration of signals is required before priming develops. As in previous studies, no priming was noted from –800 to –400 msec. Apparently, priming decays over time or is subject to interference from stimuli that intervene between the prime and the target.

The nature of priming varied with the tone–target interval. At –200 to –120 msec, priming was seen only after L+F+ tones. As in previous studies, the results suggest that the processing of individual tone features is modulated when attention is focused on tones with conjoined features. However, the magnitude of priming was smaller in the present study (maximum 15 msec) than in the previous experiment (52 msec). This may reflect backward and forward masking effects from tones preceding and following the priming stimulus. The short duration of the tone pips (10 msec) in the present study may also have reduced priming.

The RT was slowed by L+F+ standard tones occurring 40–250 msec after the target. This result suggests that the loudness of the target was evaluated over a brief interval following its presentation. The fact that target (L+F+I+) RTs were slowed by L+F+I– tones but were uninfluenced by lower intensity tones of other frequencies or locations suggests a hierarchical comparison in

which target intensity was evaluated only for L+F+ tones. This is unsurprising: hard-to-discriminate differences, like the small difference in intensity between standard and deviant tones, typically have slow search functions (Quinlan & Humphreys, 1987). Since the current paradigm included large differences between frequency cues and between location cues, but only small (near-threshold) differences in intensity, subjects may have adopted a strategy in which only the tones of attended location and frequency underwent further intensity analysis. Backward masking or loudness enhancement (Scharf & Houtsma, 1986) may also have contributed to slowed RTs following L+F+ tones.

ERP Signs of the Processing of Frequency and Location Features

Although subjects attended to stimuli defined by conjoined location and frequency cues, attention-related changes were also observed in ERPs elicited by stimuli sharing only frequency or location features with the target. These effects began at latencies as short as 70–80 msec and persisted for hundreds of milliseconds.

Apparently the subjects could not focus attention exclusively on tones with conjoined stimulus features (L+F+) without modulating the processing of signals with single attended features (L–F+ or L+F–). In previous studies, this might have been due to the relatively long

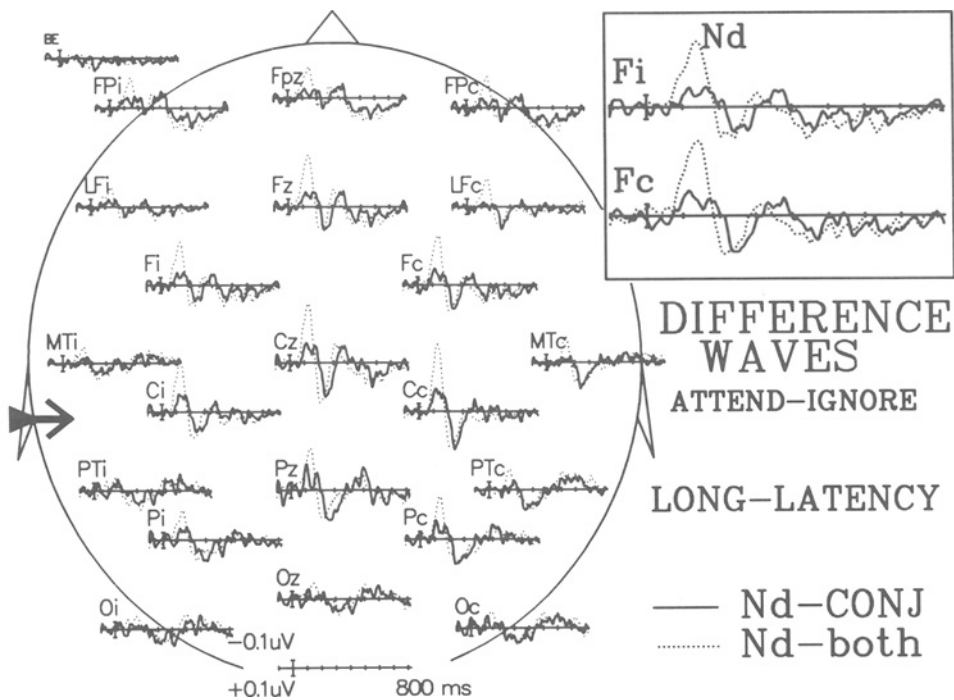


Figure 9. Nd-conj (solid) and Nd-both (dotted) difference waves shown on a long-latency time base. Averaged over tone frequency and ear of stimulation. The electrodes have been transposed so that the electrodes on the right of the figure were contralateral to the stimulated ear and relabeled accordingly: i, ipsilateral; c, contralateral. Grand mean of 12 subjects.

ISIs between L+F+ tones. At long ISIs, the “template” of L+F+ tones may have decayed, encouraging the subjects to adopt a strategy of attending to the more numerous stimuli with single relevant features. However, in the present experiment, mean ISIs were reduced to 120 msec, and only two tone frequencies were used. As a result, standard tones with conjoined location and frequency cues (L+F+) recurred at mean ISIs of 480 msec—short enough for short-latency Nd waves to emerge (Hansen & Hilliard, 1984; Näätänen, 1990).

Nd-freq effects were larger and occurred at shorter latencies than Nd-locat effects, which was consistent with previous results (Woods et al., 1991). However, in our previous experiment there were three frequency cues and only two location cues, so that the frequency conveyed more task-relevant information than did location. Although the information conveyed by location and frequency cues was similar in the present study, Nd-freq waves were still larger and of shorter latency than Nd-locat waves. The results imply that frequency cues may play a central role during auditory selective attention.

ERP Signs of the Concurrent Analysis of Successive Stimuli

Selective attention resulted in both negative and positive modulations of the ERPs to tones with relevant frequency, location, or both frequency and location features.

Nd waves showed a general resemblance: the short-latency averages revealed an initial negative phase beginning at 60 to 100 msec and peaking at 130 to 180 msec, followed by a positive phase persisting from 190 to 300 msec, a negativity of varying amplitude from 300 to 360 msec, and subsequent positivity lasting up to 440 msec. Throughout this period, Nd-freq and Nd-both waves emerged at somewhat shorter latencies than Nd-locat and Nd-conj waves. Previous experiments have shown polyphasic Nd-freq and Nd-locat waves with similar time courses (Woods et al., 1991).

The durations of Nd waves were considerably longer than the ISIs between stimuli. If stimuli were processed serially, the Nd waves would have terminated, on the average, at 120 msec, and should have terminated in every case before the longest ISI (200 msec). Our results show that the Nd waves lasted much longer. If the duration of the Nd wave is assumed to represent the duration of processing, the results suggest that successively presented stimuli may be processed concurrently.

The long duration of Nd-freq and Nd-locat waves supports the hypothesis that the analysis of single features, once started, continued to completion. The latency of target discrimination can be estimated by subtracting the motor delay from the RT. Motor delay, the delay between the decision to respond and buttonpress closure, is influenced by the task and the construction of the button. Pre-

vious studies in our laboratory using the same button suggest a motor delay of 140–180 msec (Knight, Singh, & Woods, 1989; Woods et al., in press-b). In the present experiment, this implies that the decision to respond occurred 280–320 msec after target delivery. Since Nd-freq and Nd-locat waves lasted 300 msec or more, the decision to respond occurred at roughly the same latency as the termination of Nd-locat and Nd-freq waves. Had stimulus features been analyzed serially, the duration of Nd-freq and Nd-locat waves would presumably have been much shorter.

ERP Signs of Auditory Feature Conjunction

Conjunction-specific effects (Nd-conj) had onset and peak latencies that were delayed by 30–40 msec in comparison with selective attention effects to L+F+ and L–F+ tones. The similarity of this estimate to that obtained in our previous study at longer interstimulus intervals (Woods et al., 1991) implies that the time required for the conjunction of location and frequency features—30–40 msec—may be a fundamental constant of the auditory system.

Although Nd-conj waves began 30–40 msec after Nd-freq waves, both had long durations. The onset of the Nd-conj occurred at latencies much shorter than those of the termination of Nd waves associated with frequency and location features. It should also be noted that in the present experiment, unlike that of Previc and Harter (1982), Nd-conj waves were extracted from ERPs to nontarget tones that did not cue responses. Thus, Nd-conj waves cannot be related to N2 and P3 components associated with target detection, or to premotor or motor potentials linked to the motor response. The short latency of onset of Nd-conj waves suggests that feature conjunction begins early during stimulus analysis. There was only a brief period (30–40 msec) during which the features of the L+F+ tones were seemingly processed in parallel without evidence of conjunction effects.

Implications for Feature Conjunction Theories

The results of the present experiment clarify several aspects of auditory feature conjunction. First, the modulation of processing of individual frequency and location features occurred automatically when attention was focused on tones with conjoined attributes. This result has been observed in previous studies when tones were presented at longer ISIs (Woods et al., 1991). The results suggest that the attentional modulation of individual feature processing is an essential aspect of attention to tones with conjoined features.

As in previous studies (Woods et al., in press-b), Nd-locat effects had a more central distribution than did Nd-freq waves. The different scalp distributions of Nd-freq and Nd-locat waves indicate that frequency and location cues were initially processed in distinct cortical fields, or in spatially distinct groups of neurons within a common field. The generators of the Nd-conj waves remain to be clearly defined. Nd-conj distributions were intermediate

between those of Nd-freq and Nd-locat waves, suggesting that feature conjunction may reflect enhanced processing in the fields analyzing individual stimulus features. However, various cortical and subcortical generators could produce voltage distributions similar to those seen for the Nd-conj waves, including sources along the superior temporal plane or in the thalamus, basal ganglia, cingulate gyri, or parietal lobes (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991; LaBerge, 1990; Mesulam, 1990; Posner, Petersen, Fox, & Raichle, 1988). The conjunction of auditory stimulus features may thus have been associated with increased activity in one of these brain regions.

Finally, the results are consistent with the parallel processing of stimulus features on three levels. First, there was evidence of parallel processing of stimuli in time: the long durations of Nd-freq and Nd-locat waves suggest that several successive stimuli underwent concurrent processing. Second, as in previous studies (Woods et al., 1991), there was a short (30–40 msec) period in which Nd-both waves were identical to the sum of Nd-freq and Nd-locat waves. This suggests that the analysis of frequency and location cues in the L+F+ tones initially involved the simultaneous, parallel activation of location and frequency processing fields in the same manner as did the analysis of tones with single attended features. Finally, Nd-conj, Nd-freq, and Nd-locat waves had short onset latencies and long durations. This argues against the hypothesis that feature conjunction took place only after the analysis of individual features was complete. Rather, feature conjunction appears to occur in parallel and concurrently with the analysis of individual stimulus features.

REFERENCES

- ALAIN, C., ACHIM, A., & RICHER, F. (in press). Perceptual context and the auditory selective attention effect on event-related brain potentials. *Psychophysiology*.
- ALAIN, C., & WOODS, D. L. (in press). Auditory streaming modulates the effects of pretarget tones on detection speed and accuracy during selective listening. *Perception & Psychophysics*.
- ALHO, K., SAMS, M., PAAVILAINEN, P., REINIKAINEN, K., & NÄÄTÄNEN, R. (1989). Event-related brain potentials reflecting processing of relevant and irrelevant stimuli during selective listening. *Psychophysiology*, *26*, 514–528.
- BRIAND, K. A., & KLEIN, R. M. (1987). Is Posner's "beam" the same as Treisman's "glue"? On the relation between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception & Performance*, *13*, 228–241.
- CORBETTA, M., MIEZIN, F. M., DOBMEYER, S., SHULMAN, G. L., & PETERSEN, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed: Functional anatomy by positron emission tomography. *Journal of Neuroscience*, *11*, 2383–2402.
- HANSEN, J. C., & HILLYARD, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography & Clinical Neurophysiology*, *49*, 277–290.
- 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. (1984). Effects of stimulation rate and attribute cuing on event-related potentials during auditory attention. *Psychophysiology*, *21*, 394–405.

- HARTER, M. R., AINE, C., & SCHROEDER, C. (1982). Hemispheric differences in the neural processing of stimulus location and type: Effects of selective attention on visual evoked potentials. *Neuropsychologia*, **4**, 421-438.
- KNIGHT, R. T., SINGH, J., & WOODS, D. L. (1989). Pre-movement parietal lobe input to human sensori-motor cortex. *Brain Research*, **498**, 190-194.
- LABERGE, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience*, **2**, 358-372.
- LIBERMAN, A. M., & MATTINGLY, I. G. (1989). A specialization for speech perception. *Science*, **243**, 489-494.
- 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.
- MESULAM, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology*, **28**, 597-613.
- 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.
- POSNER, M. I., PETERSEN, S. E., FOX, P. T., & RAICHEL, M. E. (1988). Localization of cognitive operations in the human brain. *Science*, **240**, 1627-1631.
- PREVIC, F. H., & HARTER, M. R. (1982). Electrophysiological and behavioral indicators of selective attention to multifeature gratings. *Perception & Psychophysics*, **32**, 465-472.
- QUINLAN, P. T., & HUMPHREYS, G. W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception & Psychophysics*, **41**, 455-472.
- RUCHKIN, D. S. (1990). Comments on the editorial policy on analysis of variance. *Psychophysiology*, **24**, 476-477.
- SCHARF, B., & HOUTSMA, A. J. M. (1986). Audition II. Loudness, pitch, localization, aural distortion, pathology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception & psychophysics: Vol 6. Sensory processes and perception* (pp. 15:1-60). New York: Wiley.
- SUGA, N., NIWA, H., TANIGUCHI, I., & MARGOLIASH, D. (1987). The personalized auditory cortex of the mustached bat: Adaptation for echolocation. *Journal of Neurophysiology*, **58**, 643-654.
- TREISMAN, A. M. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, **40**, 201-237.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration model of attention. *Cognitive Psychology*, **12**, 97-136.
- TREISMAN, A., & SATO, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.
- WOLDORFF, M., & HILLYARD, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography & Clinical Neurophysiology*, **79**, 170-191.
- WOODS, D. L. (1990). The physiological basis of selective attention: Implications of event-related potential studies. In J. W. Rohrbaugh, R. Johnson, & R. Parasuraman (Eds.), *Event-related brain potentials: Issues and interdisciplinary vantages* (pp. 178-209). New York: Oxford University Press.
- WOODS, D. L., ALAIN, C., COVARRUBIAS, D., & ZAIDEL, O. (in press). Frequency-related differences in the speed of human auditory processing. *Hearing Research*.
- WOODS, D. L., ALGAZI, A., & ALHO, K. (1991). Brain potential signs of feature processing during auditory selective attention. *NeuroReport*, **2**, 189-192.
- WOODS, D. L., ALHO, K., & ALGAZI, A. (in press-a). Intermodality selective attention: Evidence for processing in tonotopic and nontonotopic auditory fields. *Psychophysiology*.
- WOODS, D. L., ALHO, K., & ALGAZI, A. (in press-b). Stages of auditory feature conjunction: An event-related brain potential study. *Journal of Experimental Psychology: Human Perception & Performance*.
- 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.
- YUND, E. W., & EFRON, R. (1974). Dichoptic and dichotic micropattern discrimination. *Perception & Psychophysics*, **15**, 383-390.

(Manuscript received June 1, 1992;
revision accepted for publication September 17, 1992.)