

bumps surrounding the N1 component of the event-related potential (ERP) into a coherent system. His description of these phenomena is so lucid that I get the feeling these are phenomena that I myself would really know something about. Thus Näätänen's system enables me to pose these questions: (1) Is there no top-down mechanism in the process that is responsible for the mismatch negativity (MMN)? (2) How does the attentional trace hypothesis handle the fast-same effect? (3) At what stage of processing do we hear something?

(1) Top-down mechanism in MMN generation. From Figure 4 of the target article we learn that MMN is generated when the standard stimulus is 1000 Hz and the deviant tone is 1016 Hz, but there is no MMN with deviant tones of 1008 Hz. Now suppose some subjects have been trained for a week or so to discriminate 1008 Hz from 1000 Hz tones. Would there not be an MMN in these subjects with 1008 Hz in the "ignore" condition?

Another example: The target article cites Nordby et al.'s (1988b) finding that MMN is elicited even by repeated tones within a sequence of alternating tones. This may still be consistent with Näätänen's view that MMN reflects the processing of physical features only, but what happens if this simple sequence of alternating stimuli is further modified to become a series of well-known melodies, for example, children's songs? Would an MMN be evoked by occasionally deviant tones, with the deviances recognized only by people who knew those particular songs? For example, would Finnish subjects display some MMN when Finnish children's songs are incorrectly played? How would Indian subjects perform with these same melodies?

If trained 1008 Hz tones or deviances in known songs would indeed elicit MMNs then it might still be correct to state that "this processing . . . is automatic . . . and is not influenced by the direction of attention" (sect. 6) because selective attention is perhaps not necessary to detect these mismatches. However, it would be difficult to maintain that "basic sensory analysis occurs without the participation of memory" (ibid). Instead, it would seem that the "basic analysis" – a bottom-up process – is performed by analysers that are installed and tuned by top-down settings, based on previous learning, which is a matter of long-term memory.

(2) The attentional trace hypothesis and the fast-same effect. "A sensory input to the attentional trace initiates a self-terminating matching or comparison process (7), which lasts longer the more similar the eliciting stimulus is to the one represented by the attentional trace; it lasts until a 'match,' when the two stimuli are identical" (sect. 7). If overt "same"- "different" responses were required in response to each tone (actually this is hard to accomplish due to the rapid presentation rate), two predictions follow from this hypothesis for the reaction times (RTs) of these overt responses: (i) "different" RTs will be larger the more similar the stimulus is to the one represented by the trace, and (ii) "same" RTs will be largest. Indeed, the evidence available from matching tasks supports hypothesis I. However, it strongly disconfirms hypothesis II. As Farell (1985, p. 423) states in his review: "The reverse is true . . . 'same' judgments are substantially faster than 'different' RT data would lead one to expect."

There is a voluminous literature on this fast-same effect (see Czinger & Szenthe 1988, for ERP research dealing with this effect), but data have mostly been collected only in the visual modality. So one could simply argue in defense of the attentional trace hypothesis that things are different with the auditory modality and still more different with tones (cf. Farell's 1985, remark on the disappearing fast-same effect with uncodable stimuli, p. 424). The state of affairs remains unconvincing, however, because Näätänen's slow-same argument is exactly the same as the empirically disconfirmed argument in those matching tasks and because there is no chance of generalizing his attentional trace hypothesis to all those stimuli where a fast-

same effect has been demonstrated. In any case, the "matching" assumption of the attentional trace hypothesis might be tested more directly, by observing both processing negativity and RTs in matching tasks using tones as stimuli.

(3) At what stage of processing do we hear something? Näätänen's detailed discussions of the relationship between the negative components on the one hand and information processing and subjective perception on the other hand (e.g., in sect. 3.2.2: "numerous dissociations between N1 and specific contents of perception") are very welcome because all too often in ERP research, ERPs and psychological events have been regarded simply as "two sides of the same coin." The relation between two levels of psychological processing should perhaps also be elaborated more: In discussing MMN, Näätänen distinguishes between "basic sensory analysis" and "conscious perception" (sect. 6). It remains unclear whether this distinction is between "basic analysis" and *perception* or between "basic analysis" and *consciousness* or whether he indeed wants to say that perception is always conscious. To put the question simply: Do we *hear* standard tones under "ignore" conditions, that is, does the "basic analysis" produce perception? If the answer is yes, then the distinction is between two kinds of perception, for example, unconscious versus conscious or unattended versus attended. This question may arise because Näätänen likes to keep the "basic analysis," as he calls it, on a very "basic" level, implying no contact with long-term memory (leading to question 1, above) and reducing as much as possible its perception-like characteristics (leading to the present question).

Attentional influence on the mismatch negativity

Marty G. Woldorff and Steven A. Hillyard

Department of Neurosciences, University of California at San Diego, La Jolla, CA 92093

Electronic mail: mwoldorff@ucsd.bitnet

Näätänen presents a scholarly review of the many physiological studies investigating mechanisms of auditory selective attention; the model he proposes to account for this wide-ranging evidence is both creative and thoughtful. Central to Näätänen's theoretical position is the assertion that the physical features of auditory stimuli are automatically and completely analyzed, even in unattended channels. This is based largely on evidence that the event-related potential (ERP) wave known as the mismatch negativity (MMN) is elicited by slight deviations in a variety of stimulus dimensions and is not affected by attention. Näätänen further proposes that when the deviant sounds are attended and task-relevant, they also elicit a longer-latency negative wave called the N2b, which overlaps the latter part of the MMN and is closely associated with a subsequent positive wave, P3a.

It appears to us that there are a number of weak links in the chain of evidence taken to support these propositions. To begin with, we may consider Näätänen's own studies of the MMN. In particular, his claim that the MMN is not affected by attention is predominantly based on studies in which the deviance-related negativity was reported to be equivalent in the attended and unattended channels (e.g., Näätänen et al. 1978; 1980; Sams et al. 1984). Thus, in those cases, the deviance difference waves in the attended versus unattended conditions differed only in that in the attended case the deviance-related negativity was followed by a P3 wave. In the context of the Näätänen framework, however, it is difficult to understand why identical negativities would be elicited under both conditions. If the equivalent negativities elicited by the attended and unattended deviants both represent an automatic MMN that is not affected by attention, this would imply that essentially no N2b at all was

elicited by the task-relevant, attended deviants. Such a conclusion would appear difficult to reconcile with the Näätänen framework, however, which specifies that when deviant tones are attended and task-relevant, they *should* elicit an N2b in addition to an MMN.

If no N2b's were elicited by the attended deviants in these experiments, one might hypothesize that the amount of attention directed towards them was perhaps insufficient to elicit an N2b. On the other hand, if attention was not focused very *selectively* in these experiments, and if eliciting an N2b requires only a minimal allotment of attention, another possibility to consider is that an equivalent amount of N2b was elicited by both the attended and unattended deviants, summing with whatever "automatic" MMN might have been elicited. Either way, these considerations suggest that the independence of the MMN from attentional influences was not very strongly tested. Such a conclusion is consistent with the view that the conditions in these experiments (typically, reading or dichotic listening at slow stimulus presentation rates) were probably not particularly conducive to the selective focusing of attention. (For further discussion, see Hansen & Woldorff, in press; Woldorff, Hackley & Hillyard, submitted; also see sect. 5.1. of the target article.)

To obtain further information on this point, we recently investigated the MMN in two dichotic listening experiments in which conditions were optimized for the selective focusing of attention (Woldorff, Hackley & Hillyard, 1989; submitted). These conditions included: (1) easily discriminable channels of tones that were distinguished by both ear of entry and pitch, (2) a rapid rate of stimulus presentation, and (3) a fairly difficult target-detection task within the attended channel (responding to infrequent, slightly fainter deviant tones). The effects of attention on the ERPs elicited by the standard (i.e., nondeviant) tones in these experiments indicated that a highly selective focusing of attention had been induced that affected processing as early as 20 msec post-stimulus (the "P20-50"); in addition, there was strong evidence for direct amplitude modulation by attention of several of the exogenous subcomponents of the N1 and later waves (see Hackley, Woldorff & Hillyard, 1987; Woldorff, Hansen & Hillyard 1987; Woldorff and Hillyard, submitted). These experiments accordingly provided a strong test of Näätänen's hypothesis that the MMN is unaffected by attention.

In Experiment 1, with interstimulus intervals (ISIs) of 120–320 msec, the fainter deviant tones in both the attended and unattended channels elicited negative waves consistent in waveshape, latency, and distribution with previously described MMNs. In sharp contrast with Näätänen's results, however, and as shown in Figure 1, the unattended-channel MMN was drastically reduced (peak amplitude less than 1 uV) relative to the corresponding negative wave in the attended channel (~3 uV). In the second experiment, with ISIs of 65–205 msec, the MMN elicited by the deviant fainter tones in the unattended ear was practically abolished, whereas the corresponding deviance-related negative wave in the attended channel was again around 3–4 uV. These results strongly suggest that the MMN elicited by the deviants in the unattended channel was highly suppressed, presumably a result of attention being strongly focused on the attended channel.

In order to reconcile these data with Näätänen's view that the MMN is independent of attentional influence, one would presumably have to argue that equivalent "automatic" MMNs were, in fact, elicited by both the attended and unattended deviant tones, and that the entire additional negativity elicited by the attended deviants consisted of a separate component, such as an N2b, that was associated with their task relevance. There are a number of problems with such an explanation, however. First, according to the Näätänen framework, the short ISIs and the moderate intensity deviations (~15 dB decrements) that were used in these two experiments should have been conducive to eliciting relatively large "automatic" MMNs.

Deviance Difference Waves

(Deviant ERP – Standard ERP)

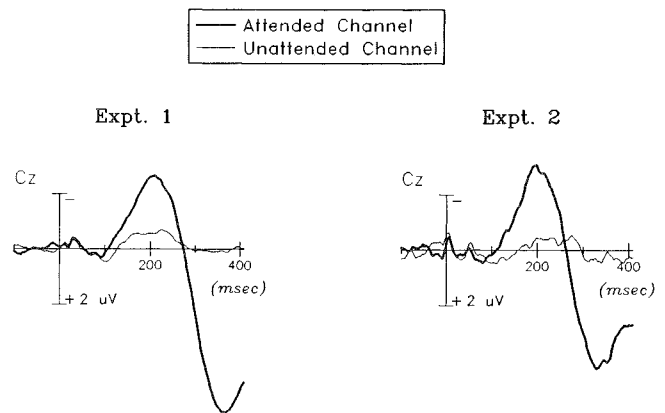


Figure 1(Woldorff & Hillyard). Deviance-related ERPs (Cz site) in the attended and unattended channels in two dichotic listening experiments optimized for the selective focusing of attention. Deviance difference waves were derived by subtracting the ERP to the standard tones from the ERP to the deviant tones (occasional intensity decrements) in the same channel. Traces are collapsed across ear of stimulation and grand averaged across subjects (Experiment 1: N=16; Experiment 2: N=12). Note that the deviance-related negativity in the attended channel was of similar amplitude in the two experiments, but the corresponding negativity in the unattended channel was considerably reduced in Experiment 1 and practically abolished in Experiment 2. Data from Woldorff, Hackley and Hillyard (submitted).

The unattended-channel MMN in Experiment 1, however, was markedly smaller than previously reported MMNs (which were typically 3–7 uV) in experiments that used similar or even smaller deviations (see Näätänen 1986 for comparison; Figure 5 of the target article shows the same data but the calibration scale is missing). Moreover, in Experiment 2, the "automatic" MMN did not even rise above noise levels. In contrast, the amplitude of the corresponding deviance-related negative wave in the attended channel in both experiments was in line with that of previously reported MMNs. Such results are clearly difficult to explain by postulating that the attended deviants elicited an additional negative component (such as N2b) that summated with an automatic MMN.

Another problem with such a hypothesis is that no evidence for two separate negative components was observed in our data. In several previous studies reporting equivalent MMNs in the attended and unattended channels, the waveforms did suggest that an additional negative component, termed N2b, was elicited by the attended-channel deviants. In those cases, the "automatic" MMN was *defined* as the negativity in the deviance difference wave (deviant-tone ERP minus standard-tone ERP) for the unattended channel, and the N2b was said to be distinguishable in the attended-channel difference wave as a second negative deflection that had a later onset and a more posterior distribution (Näätänen & Gaillard 1983; Sams et al. 1985). In our data, however, no such distinction between two components was evident. In Experiment 1 (in which the deviance-related negative wave for unattended tones was large enough to permit comparison), the attended-channel negativity began at about the same time as the unattended-channel negativity, it was significantly larger than the unattended onset in the earliest window in which the two appeared, and it peaked at about the same latency. Furthermore, the deviance-related negativities had very similar scalp distributions, with fronto-central maxima.

Thus, to argue that all the additional negativity in the attended channel was N2b one would also have to explain why such an N2b-dominated wave did not display the previously described N2b characteristics and was essentially indistinguishable in waveshape, latency, and distribution from the unattended-channel MMN in the same experiment, as well as previously reported MMNs.

These recent data from our laboratory, together with the ambiguities in the results presented by Näätänen and colleagues, suggest that the framework and interpretations proposed by Näätänen for the MMN and N2b waves may need to be reconsidered. Moreover, the waveforms illustrated in Figure 1 strongly challenge Näätänen's claim that feature analyses and mismatch detection of auditory stimuli are unaffected by attention. On the contrary, our data would appear to be more consistent with the view that auditory selective attention can attenuate processing of elementary stimulus features in unattended channels.

ACKNOWLEDGMENTS

This work was supported by an NSF Graduate Fellowship to M. Woldorff and by NIMH grant MH 25594 to S. A. Hillyard. The authors would like to express their appreciation for the helpful comments of J. C. Hansen and G. R. Mangun.

Selective auditory attention: Complex processes and complex ERP generators

David L. Woods

Department of Neurology (127), University of California, Davis, VAMC, Martinez, CA 94553
Electronic mail: dlwoods@ucdavis.edu

PN – a single component or family of components? While Näätänen's theory emphasizes the role of auditory cortex in generating the processing negativity (PN), selective attention also requires other brain regions. For example, lesions affecting frontal or parietal lobes, the cingulate gyrus, basal ganglia, or thalamus produce auditory inattention (Mesulam 1981). Modulation of neuronal activity with auditory attention during the latency range of the PN occurs in structures outside primary and secondary auditory cortex, including the dorsolateral frontal lobe, basal ganglia and thalamus (reviewed in Woods 1989). Moreover, unless the time course of neuronal discharges was identical in these different auditory fields, relative activity would change in different cortical areas over the latency range of the PN. Hence, it is plausible that the PN (even its short latency component) is actually a family of subcomponents, whose distribution should change as a function of latency.

If the PN reflects early stimulus selection then it might be expected to arise in different cortical fields if selections occurred on the basis of different stimulus features (e.g., Merzenich & Brugge 1973). For example, while intermodal auditory selection might be based on the output of any auditory field, fine discriminations of stimulus location or pitch would require modulation of processing in auditory fields with sharp tuning curves for those features. Hence, intermodal attention effects might be expected to differ in waveshape, amplitude, and distribution from PN effects found with dichotic tone discriminations (e.g., Woods 1989).

The possibility that the PN is generated in different auditory fields in different tasks might also explain some curious data about "attentional leakage." For example, when attention is directed to a tone sequence characterized by pitch and location cues, small PNs are seen for tones that share one cue with the attended sequence but differ in the other. For example, small "leakage effects" are noted for tones that share spatial location with the attended sequence but differ in pitch (Alho et al., 1989;

Hansen & Hillyard 1983). In other paradigms, however, no such "leakage" effects are observed. For example, processing negativities are not seen for tone probes that share location cues with continuous speech but differ in frequency and timbre (Woods et al. 1984), or for novel sounds sharing the location of the attended sequence but distinguished by pitch (Woods 1989).

Why do attentional effects lead to the contralateral ear in some experiments but not others? Location leakage typically occurs when differences between attended and nonattended sequences are small. With small differences in pitch between attended and nonattended signals, selection would depend on cortical fields with sharp frequency tuning curves. If these fields had crude location tuning (as EE cells in primary cortex do, Imig & Brugge 1978), modulating their outputs would concurrently modulate processing for tones of similar frequency in the nonattended ear. Similarly, leakage might occur between different frequencies presented to the attended ear in cortical fields with sharp location tuning but crude frequency tuning. With larger pitch and location differences, modulation could take place in cortical fields with cruder frequency and location tuning, to permit selection without leakage. This reasoning implies that PN distributions should differ for tasks using different cues, or even for tasks of differing difficulty that use the same cues.

Does auditory selection involve inhibition as well as facilitation? While Näätänen stresses the facilitated processing of attended inputs, there is substantial neurophysiological evidence that selective attention also involves inhibition of the processing of nonattended stimuli (Wurtz et al. 1984). For example, Moran and Desimone (1985) suggest that such inhibition is the basic mechanism of visual selective attention. They found neurons that responded briskly to one stimulus and moderately to another. When both stimuli were presented together responses continued at high rates provided that attention was directed to the effective stimulus. However discharges were inhibited when attention was directed to the ineffective stimulus. Thus, the response to the effective stimulus (still present in the visual field) appeared to be "gated" when attention was directed to the ineffective stimulus.

In the human ERP literature, indirect evidence for inhibition of the processing of nonattended inputs can be noted in P3 probability effects. When subjects perform in oddball tasks, P3s are produced by infrequent stimuli regardless of whether frequent or infrequent stimuli are attended (Squires et al. 1977). In selective attention tasks this probability function is changed in two ways. First, the amplitude of P3s in attended channels is determined exclusively by the relative probabilities of standards and targets in the attended channel, not by nonattended inputs. Second, no P3s are produced by rare standard tones (even with probabilities of 25% or less) in dual cue tasks; nor do deviant "target" stimuli produce P3s in nonattended sequences.

These results imply that sensory processing of nonattended stimuli is inhibited before P3 generation and before evaluations of stimulus probability that may modify P3 amplitude. The inhibitory process may function like an attentional "parasol," inhibiting the full analysis of nonattended stimuli and of stimuli that resemble them acoustically. Hence, it would inhibit differential processing of deviant stimuli, but only deviant stimuli that closely resembled rejected standards.

Recent studies provide evidence about two other aspects of this inhibitory process. First, like Näätänen's attentional trace, it may require stimulus repetition for its maintenance. Thus, deviant stimuli in nonattended channels produce P3s early in a stimulus sequence (Hansen & Hillyard 1988), or when nonattended inputs are repeated infrequently (Alho et al. in press). Second, the processing of sounds that deviate acoustically from other stimuli in the to-be-ignored channel are not inhibited: Equivalent P3as are elicited by novel stimuli in attended and nonattended channels (Woods 1989). This suggests that the inhibitory parasol, like the putative attentional spotlight, may have a narrow aperture. Novel stimuli that deviate markedly