Psychophysiology, 35 (1998), 413–419. Cambridge University Press. Printed in the USA.

Automatic change detection: Does the auditory system use representations of individual stimulus features or gestalts?

DIANA DEACON,[a] JO MANETTE NOUSAK,[b] MAURA PILOTTI,[a] WALTER RITTER,[c] AND CHIEN-MING YANG[a]

[a]Department of Psychology, The City College of the City University of New York, NY, USA
[b]Department of Speech and Hearing, The Graduate Center of City University, New York, NY, USA
[c]Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY, USA

Abstract

The effects of global and feature-specific probabilities of auditory stimuli were manipulated to determine their effects on the mismatch negativity (MMN) of the human event-related potential. The question of interest was whether the automatic comparison of stimuli indexed by the MMN was performed on representations of individual stimulus features or on gestalt representations of their combined attributes. The design of the study was such that both feature and gestalt representations could have been available to the comparator mechanism generating the MMN. The data were consistent with the interpretation that the MMN was generated following an analysis of stimulus features.

Descriptors: Mismatch negativity, Preattentive storage

Human scalp recordings of event-related potentials (ERPs) provide evidence of an early, pre-attentive negative component, referred to as the *mismatch negativity* (MMN), when an auditory stimulus deviates in virtually any discriminable, physical dimension from preceding tones that are homogeneous in some aspect (for a review, see Näätänen, 1992). The MMN is usually not dependent on attention, as it may even be recorded during sleep (Campbell, Loewy, Bastien, & Bell, 1991; Csépe, Karmos, & Molnár, 1987). Depth recordings in cats (Csépe et al., 1987), monkeys (Javitt, Steinschneider, Schroeder, Vaughan, & Arezzo, 1994), and humans (Kropotov et al., 1995), have established that the MMN is generated in or near auditory cortex, at least for tones that deviate from standard stimuli on simple acoustic features.

The MMN has been used in a number of studies to investigate the transient memory upon which it depends (for a review, see Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995). Among the issues addressed in these studies is how information is stored in the memory. One possibility is that stimulus features are stored independently and another is that the features of given stimuli are stored together as a unit, that is, as a gestalt. Considerable physiological evidence suggests that stimulus features are initially processed independently, especially in the visual system (Zeki, 1993). Treisman (1992) has proposed that the integration of features requires attention. Because the operations underlying the MMN are considered to be mainly pre-attentive, we have the opportunity to determine whether the integration of features occurs pre-attentively in the auditory system.

Evidence in support of independent processing and/or storage of features has been reported in several MMN studies (Giard et al., 1995; Gomes, Ritter, & Vaughan, 1995; Levänen, Hari, McEvoy, & Sams, 1993; Schröger, 1995). Giard et al., for example, found that the location of the MMN generator varied as a function of the feature by which the deviant that elicited the MMN differed from the standard. Winkler et al. (1990), based on an interesting post hoc interpretation of their MMN data, suggested that information is stored on a gestalt basis. Using a direct test of a corollary of the gestalt hypothesis, Gomes, Bernstein, Ritter, Vaughan, and Miller (1997) obtained evidence that information about the conjunction of two features is stored in the memory. Three standards were presented, each of which had different values of intensity and frequency, and a deviant that had the intensity of one of the standards and the frequency of one of the other standards elicited a MMN. Because the MMN was associated with a difference between the deviant and standards in the combination of the values of two features, the result is consistent with, but does not necessarily establish, gestalt storage. Some features may be stored as conjunctions but other features are not.

The present report is an extension of a talk given at the 1991 EPIC X meeting and a subsequent publication (Deacon, Pilotti, & Tinsley, 1995) that were based on preliminary data. The basic idea was to present stimuli in such a way that deviant stimuli could elicit the MMN using either stored information about individual features or gestalt representations. Three kinds of infrequent deviants

This research was supported by USPHS grant RO1DC00895 to D. Deacon and USPHS grant RO1NS30029 to W. Ritter.

Address reprint requests to: Dr. Diana Deacon, Department of Psychology, City College, 138th Street and Convent Avenue, New York, NY 10031, USA.

were presented in the same run, each of which differed from the standard on a different feature (but was identical to the standard on all other features) and occurred on 10% of the trials. Each deviant, therefore, differed from the standard with regard to the value of a specific feature, but also with regard to the combination of the values of its features. The global probability of deviant tones was 30%, but the probability of deviant tones that differed from the standard on a given feature was 10%.

After publication of the Deacon et al. (1995) preliminary report, we realized that a confound may have been present in the design. Deviants that differed from the standards in duration, intensity, and frequency were used. The standard tones were 50 ms in duration and the deviant tones that differed in duration were 75 ms in duration. Consequently, the 75-ms deviant tones could have sounded louder than the 50-ms standards, producing a confound with the intensity deviants. In the present study, deviants that differed from the standards in duration were replaced by deviants that differed from the standards in stimulus onset asynchrony (SOA). An entirely new set of subjects participated in the present study.

As is well known, the amplitude of the MMN is inversely related to the probability of deviant events (Näätänen, 1992). If the standard is stored as a gestalt containing integrated information about all of its features, and the MMN reflects deviation from the gestalt as stored, then no distinction should be made by the system as to the particular feature by which a given deviant happens to differ from the standard. In other words, the amplitude of the MMN should reflect the global probability of the deviant tones.

Another way of conceptualizing the issue pertains to Näätänen's (1992) suggestion that the memory may entail the anticipation or prediction of future events on the basis of past events (see also Winkler, Karmos, & Näätänen, 1996). To avoid any implication of conscious expectation, as for example is associated with N2 and P3 (and not intended by these authors), we will use the terms *state* and *preparedness* for the same notion. Using these terms it can be said that, at any given moment, the state of the MMN system is such, or it is prepared in such a way, that a particular stimulus event will not elicit a MMN. The MMN would be elicited when a stimulus does not match what the system is prepared for, in which case the state of the system would be altered such that, if the same stimulus event is immediately repeated, it will either elicit a smaller MMN or none at all. Winkler et al. (1996) suggested that the degree of certainty of the "prediction" (what we would call the degree of preparedness) is positively correlated with the size of the MMN elicited by deviant stimuli. Using this conceptualization, what is prepared for depends on whether the system operates on the basis of gestalts or features. If the MMN system is operating on the basis of gestalt information, then the event for which the system is prepared would be the standard with its combination of features. When three different kinds of deviants are used, each of the deviants would have a similar effect on the system in that each would alter the degree of preparedness for the standard. Moreover, each deviant should reduce the preparedness for the standard by an equivalent degree. Therefore, the effect on preparedness for the standard should be identical in a circumstance where there are three different deviants that occur on a given percentage of the trials and another circumstance where there is only one deviant that occurs on the same percentage of trials. Put in empirical terms, when three kinds of deviant tones each differ from the standard on a different feature, and each kind of deviant has a probability of 10%, the MMN should have an amplitude comparable to a single deviant that differs from the standard on one feature and occurs on 30% of the trials. Hence, the amplitude of the MMN would be associated with the global probability of deviance.

However, if the MMN is operating on the basis of features rather than gestalts, and the MMN reflects deviation from stored representations of given features, then the system should make a distinction with regard to the feature by which a given deviant differs from the standard. In other words, the amplitude of the MMN should reflect the probability by which a given feature differs from the standard and not the global probability of deviance. Note that in the condition in which three different deviants were used in the present study, the specific features by which they differed from the standard were identical on 90% of the trials. Whereas a given deviant differed from the standard on one feature, its other features had values identical to the standard. Consequently, the amplitude of the MMN elicited by each of the three deviants should be similar to what is obtained in a simple oddball condition in which a single deviant has a 10% probability of occurrence.

Stated with regard to the hypothesis that the memory entails preparedness, there would be a separate preparation for each feature of the standard. In this circumstance, different kinds of deviants would not have similar effects on the system, because a deviant that differed on one feature from the standard would have no effect on the preparation for other features.

To establish which of these possibilities applied to the condition in the present study in which three different kinds of deviants were tested, conditions were employed in which a single deviant was presented on either 10% or 30% of the trials. The critical question was whether the condition with three deviants would yield MMNs that were more like those obtained in the condition that had a single deviant that occurred on 30% of the trials or the condition in which a single deviant occurred on 10% of the trials. Because deviants were used that differed from the standards on three different features, assessment of the pattern of results along three dimensions was possible.

Methods

Subjects

Subjects were three male and seven female healthy adults, aged 25–40 years ($M = 30.6$ years), with normal hearing and no history of neurological disorder. Some of the subjects were graduate students working in the laboratory and others were paid volunteers who were recruited from introductory psychology classes.

Procedure

The subjects sat in a comfortable chair in a quiet room and read the material that they had brought with them while the stimuli were delivered. The stimuli were pure tones that were delivered through headphones. Subjects were told to ignore the stimuli.

Both the standard and deviant tones were 50 ms in duration and had rise and fall times of 10 ms. Table 1 displays the stimulus parameters used. In all runs the standard stimuli were 90-dB SPL, 1000-Hz tones, which were delivered at a rate of one every 780 ms. The three conditions are delineated in Table 2. Two runs (420 stimuli each) were presented in Condition 1 and two runs were presented for each deviating feature in Conditions 2 and 3. In all conditions, the standard and deviant tones were sequenced randomly. The order in which the different types of conditions and runs were presented was randomized within and across subjects.

Condition 1. Standard tones occurred on 70% of the trials and three types of deviant tones were presented on 10% of the trials

Table 1. *Stimulus Parameters for Standards and Deviants*

	Standard	Intensity deviant	Frequency deviant	SOA deviant
Intensity (dB SPL)	90	80	90	90
Frequency (Hz)	1000	1000	1050	1000
SOA (ms)	780	780	780	600

Note: SOA = stimulus onset asynchrony.

each. Deviant tones that differed from the standards, and the other deviants, in intensity (these tones were 80 dB SPL) were delivered on 10% of the trials. Deviant tones that differed from the standards, and the other deviants, in frequency (1050 Hz) were presented on another 10% of the trials. Deviant tones that had a shorter SOA (600 ms) than all of the other tones occurred on another 10% of the trials. This condition is referred to as the 70-10-10-10 condition.

Table 2. *Summary of Experimental Conditions*

Condition 1	Standard = 70%
	Intensity deviant = 10%
	Frequency deviant = 10%
	SOA deviant = 10%
Condition 2	
A	Standard = 70%
	Intensity deviant = 30%
B	Standard = 70%
	Frequency deviant = 30%
C	Standard = 70%
	SOA deviant = 30%
Condition 3	
A	Standard = 90%
	Intensity deviant = 10%
B	Standard = 90%
	Frequency deviant = 10%
C	Standard = 90%
	SOA deviant = 10%

Note: SOA = stimulus onset asynchrony.

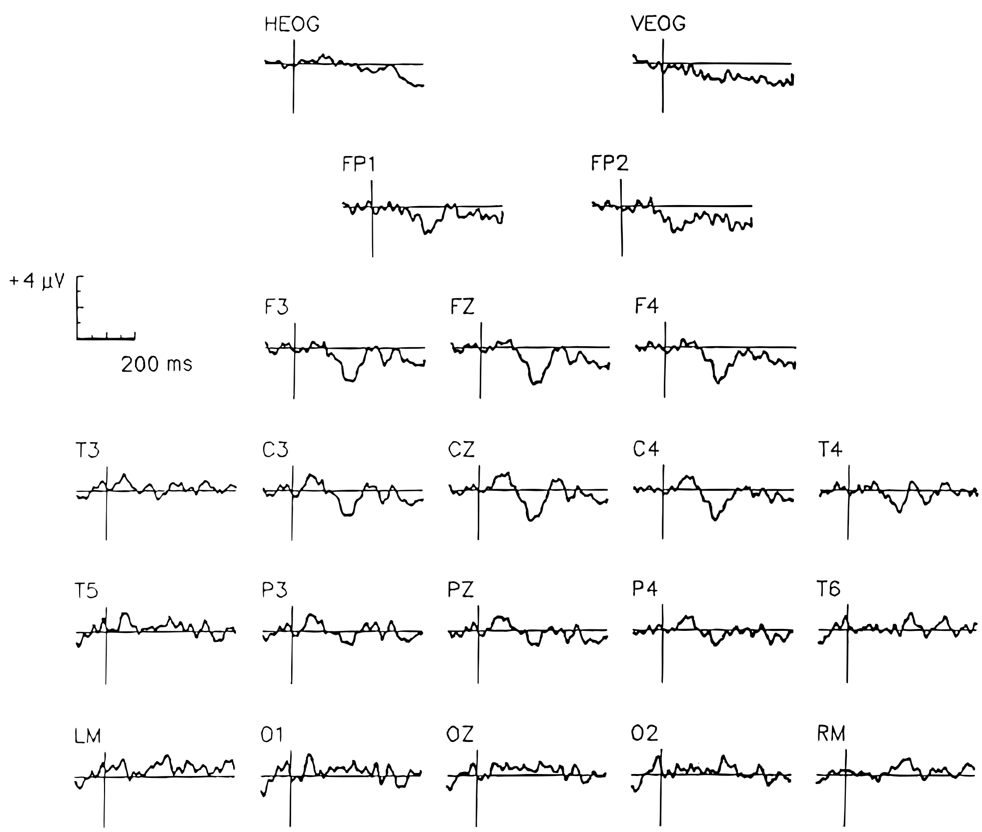

Figure 1. Scalp topography of the grand mean mismatch negativity difference waveforms obtained at all recording sites in the 70-10-10-10 condition by subtracting the event-related potential (ERP) elicited by the frequent, p (probability of occurrence) = .70 standard stimulus from the ERP elicited by the rare, p = .10, deviant stimulus that differed from the standards in intensity. The vertical line indicates stimulus onset, and positive is up, in this and all subsequent figures.

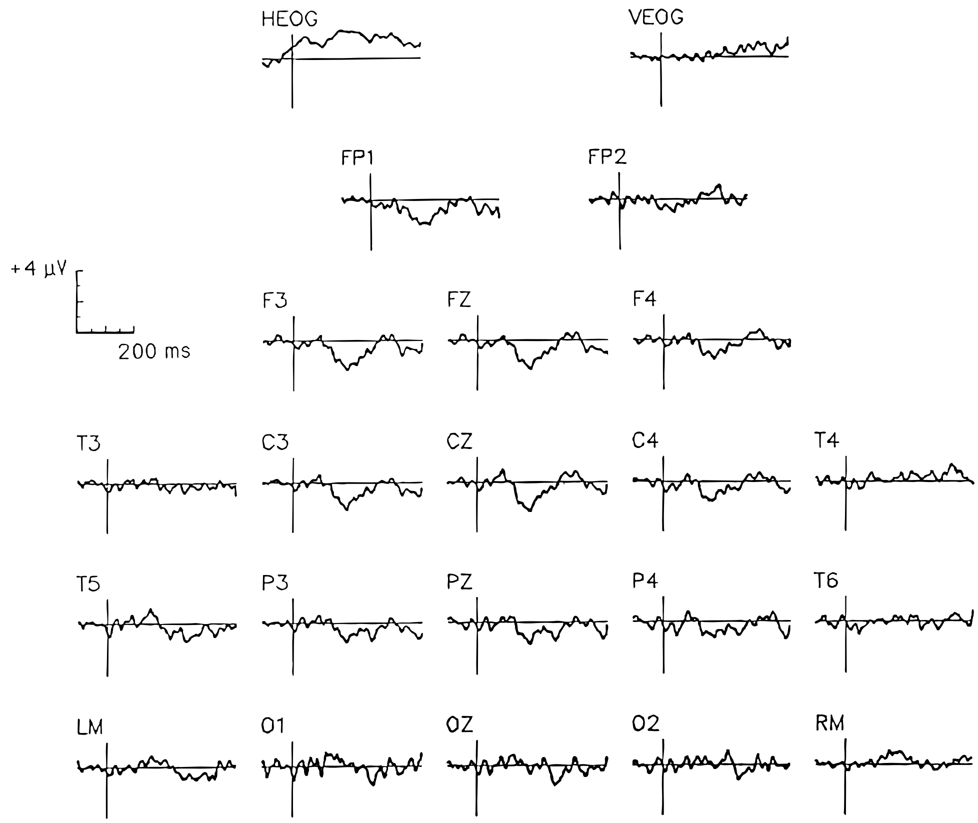

Figure 2. Same as Figure 1 except that the deviant differed from the standards in frequency.

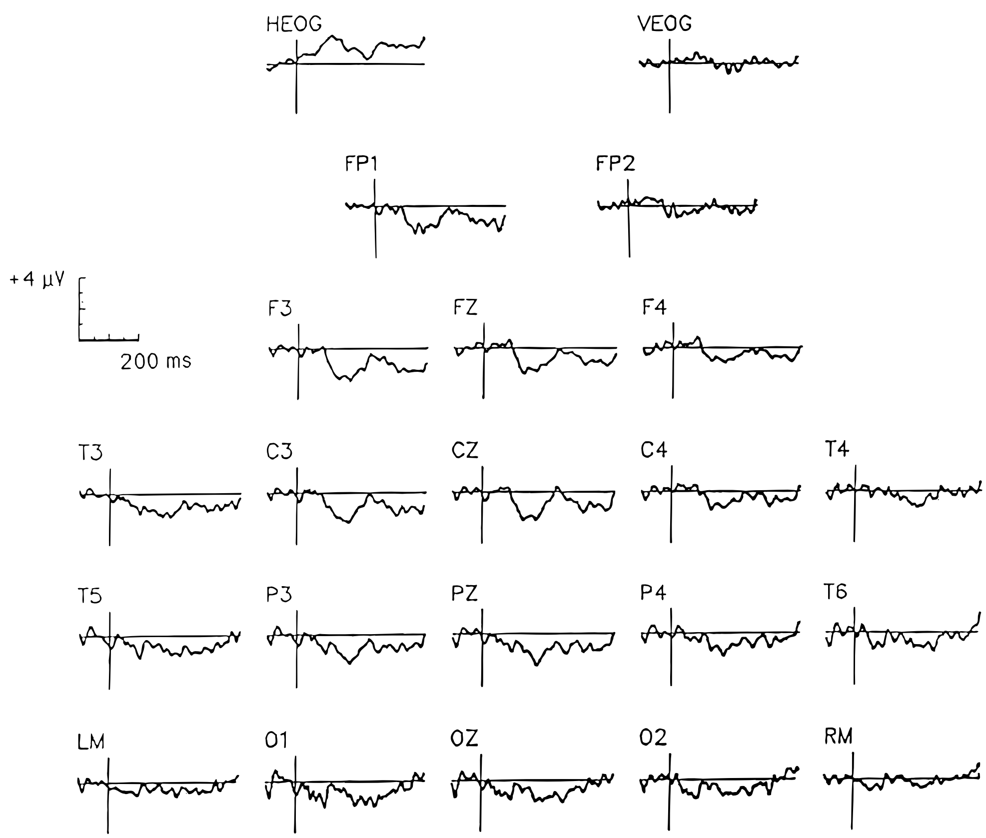

Figure 3. Same as Figure 1 except that the deviant differed from the standards in stimulus onset asynchrony.

Condition 2. As in the 70-10-10-10 condition, the standards occurred on 70% of the trials. On two runs, deviant tones that differed from the standards in intensity (80 dB SPL) were delivered on 30% of the trials. On another two runs, deviant tones that differed from the standards in frequency (1050 Hz) occurred on 30% of the trials. And on another two runs, deviant tones that had a shorter SOA than the standards (600 ms) were presented on 30% of the trials. This condition is referred to as the 70-30 condition. In Conditions 1 and 2, the global probability of deviance was identical (30%). The difference between the two conditions was whether the deviant tones differed from the standard on one or three dimensions.

Condition 3. Standard tones were delivered on 90% of the trials. On two runs, deviant tones that differed from the standards in intensity (80 dB SPL) were presented on 10% of the trials. On another two runs, deviant tones that differed from the standards in frequency (1050 Hz) occurred on 10% of the trials. On another two runs, deviant tones which had a shorter SOA than the standards (600 ms) were delivered on 10% of the trials. This condition is referred to as the 90-10 condition.

Recording Procedures

Brain electrical activity was recorded with cutoffs of 0.01 and 40 Hz. Recordings began 100 ms before stimulus delivery and continued for 450 ms after the stimulus, 256 points being sampled over the total epoch of 550 ms. Figure 1 displays the electrode placements used, which included all 10-20 sites and Oz and the left and right mastoids (LM and RM, respectively), referenced to the nose. Electrodes placed at the two outer canthi were used to monitor horizontal ocular potentials. Electrodes placed above and below one eye were used to monitor vertical ocular potentials.

The electrical activity was averaged with respect to the prestimulus baseline. Trials during which electrical activity exceeded $\pm 50\ \mu V$ for any recording were automatically rejected.

Data Analysis

The MMN was delineated by subtracting the ERP elicited by the standard tones in each condition from the ERP(s) elicited by the deviant(s) within the same condition. The peak amplitude of the MMN in the difference waveform was measured at Fz (its maximum). The MMN was identified as the largest negative peak occurring within a window of 100–300 ms. This time window was chosen on the basis of the approximate range of peak latencies of the MMN, in the grand mean waveforms, across the conditions.

Amplitude measures were submitted to analysis of variance on the factors of deviating feature (intensity, frequency, SOA) and condition of probability (70-10-10-10 condition, 70-30 condition, 90-10 condition). A significance level of .05 was adopted. The Huynh–Feldt correction procedure was used and Tukey post hoc tests were applied where appropriate.

Results

Figures 1, 2, and 3 display the grand mean difference waveforms used to delineate the MMN at all recording sites for the 70-10-10-10 condition associated with the deviants that differed from the standards in intensity, frequency, and SOA, respectively. The MMN was largest in the frontocentral region and peaked at around 190 ms for intensity deviants, 185 ms for frequency deviants, and 150 ms for SOA deviants.

Figure 4 presents the difference waveforms at Fz across all conditions and stimulus probabilities. The MMN was largest in the 70-10-10-10 condition and smallest in the 70-30 condition. A main effect of probability was observed on the amplitude of the MMN, $F(2,18) = 6.83$, $p = .011$, $\epsilon = 0.82$. Tukey post hoc tests indicated that the MMN was significantly larger in the 90-10 and 70-10-10-10 conditions than in the 70-30 condition. The MMN was marginally larger in the 70-10-10-10 condition than in the 90-10 condition (see Table 3), but the difference was not significant. There was no main effect of deviating feature, nor was there an interaction between the factors of feature and probability. Figure 5 shows the difference waveforms associated with the intensity deviant at all electrodes, illustrating both the amplitude effect and the consistency of the topography of the MMN across conditions.

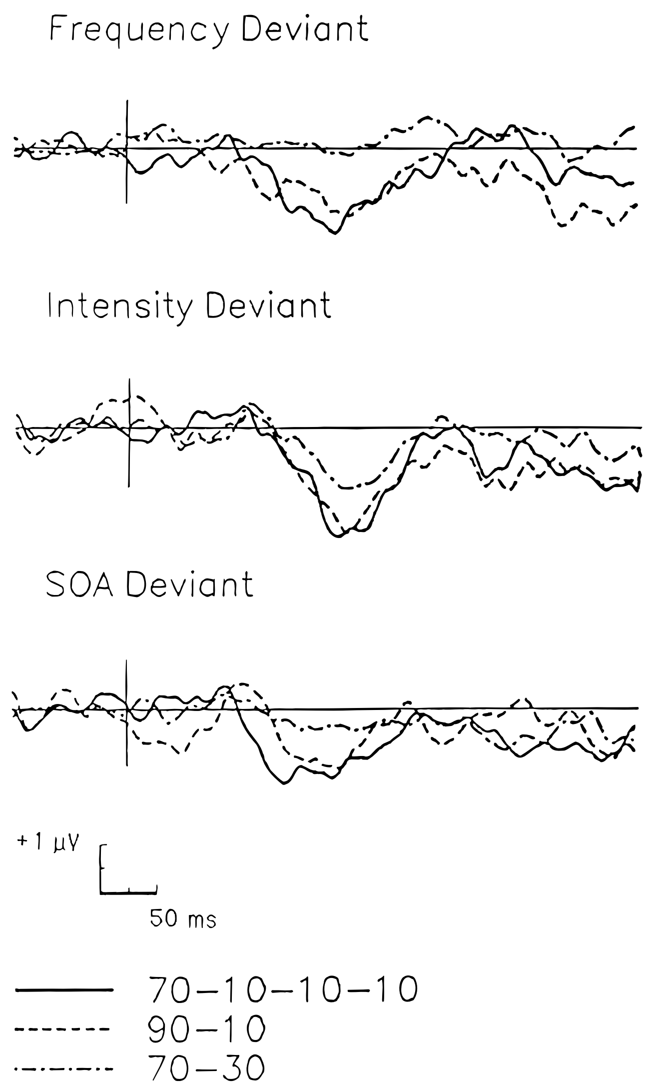

Figure 4. Effects of the probability manipulation demonstrated in the grand mean mismatch negativity (MMN) difference waveforms at Fz for the three types of deviants (frequency, intensity, and stimulus onset asynchrony [SOA]). The MMN was smallest in the 70-30 condition and slightly larger in amplitude in the 70-10-10-10 than the 90-10 condition.

Table 3. *Mean (SD) Amplitudes of the MMN (in μV) at Fz*

	Condition		
	70-10-10-10	70-30	90-10
Intensity deviant	−1.99 (0.93)	−1.07 (1.00)	−1.78 (1.28)
Frequency deviant	−1.44 (1.03)	−0.03 (0.76)	−1.14 (0.95)
SOA deviant	−1.32 (1.54)	−0.42 (0.56)	−1.09 (1.12)
M	−1.58	−0.51	−1.37

Note: MMN = mismatch negativity; SOA = stimulus onset asynchrony.

Discussion

The amplitude of the MMN was larger in the 90-10 than the 70-30 condition. Hence, when only a single deviant was in a given run, the amplitude of the MMN was inversely related to the probability of the deviant, and this finding was true for all of the three features that were manipulated. These results were expected based on the literature. The critical result was that the amplitude of the MMN in the 70-10-10-10 condition was more like that of the 90-10 condition than the 70-30 condition. Specifically, the MMN was larger in the 70-10-10-10 condition than in the 70-30 condition and even somewhat larger (although nonsignificantly so) than in the 90-10 condition. Thus, the amplitude of the MMN was related to the specific probability of the deviating feature rather than the global probability of deviance, and this observation was also true for all of the three features that were manipulated. The data therefore support the hypothesis that the MMN system was operating on the bases of features rather than gestalts. That is, the MMN system was prepared for individual features rather than for the standard with its combination of features.

The question of whether the MMN system operates on the bases of features or gestalts was also examined in a recent study by Nousak, Deacon, Ritter, and Vaughan (1996). Sams, Alho, and Näätänen (1984) found that if two identical deviants are presented in a row, the MMN elicited by the second deviant is reduced compared with the MMN elicited by the first deviant. The data could be interpreted to mean that the first deviant weakened the memory for the standard. To test this theory, Nousak et al. (1996) presented two deviants in a row each of which differed from the standard on a different feature. The MMN elicited by the second deviant was equally large as that in another condition in which it was the only deviant. The presence of the first deviant, therefore, had no effect on the amplitude of the MMN elicited by the second deviant. Consequently, the data suggested that a deviant may weaken the memory for the feature by which it differs from the standard (as in Sams et al., 1984), but not for the memory of other features of the standard. Because the strengths of the representations of the features by which the first and second deviants differed from the standard varied independently of one another, the feature values may be considered to be stored separately.

Our results are also consistent with the finding of Giard et al. (1995) that the neural source of the MMN varies as a

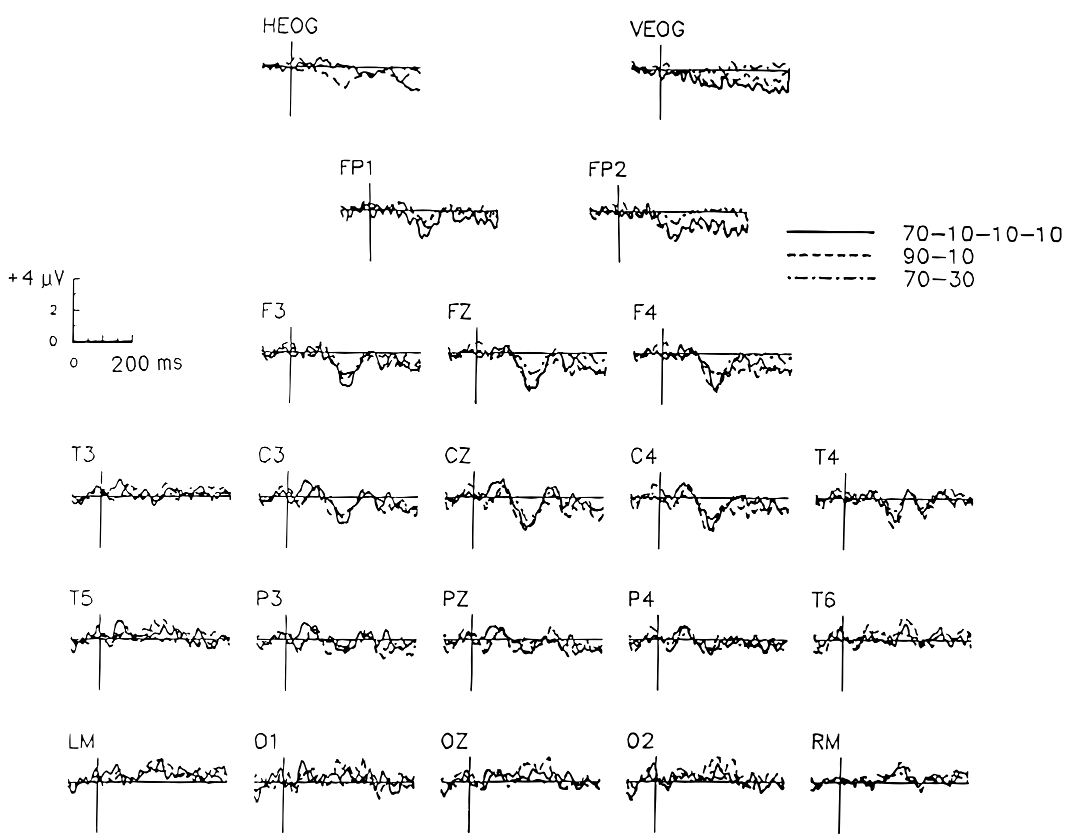

Figure 5. Grand mean difference waveforms at all recording sites and all probability conditions for the mismatch negativities associated with the intensity deviant.

function of the nature of the feature that deviates from the standard. Assuming that the different sources for the MMN implies different locations as to where the relevant memories are located, Giard et al. inferred separate storage for specific acoustic features.

Although the present data are in line with the interpretation that the comparator mechanism operates on the representations of features, this finding does not rule out that comparisons may also be performed on the basis of gestalt representations. Indeed, as was mentioned in the Introduction, Winkler et al. (1990) interpreted their data as indicating gestalt storage and Gomes et al. (1997) found data that are consistent with (although they do not prove) gestalt storage. If there is gestalt storage, we do not know why the present results did not indicate that the MMN system operated on the basis of gestalt information. Perhaps if the MMN can be elicited on the basis of both gestalt and feature information, one of the ways can inhibit the other. For example, a rule may apply such that whichever of the two ways elicits a MMN earlier terminates the other way of triggering a MMN. In support of this possibility, Czigler and Winkler (1996) have found that if a tone differs from a standard in both frequency and duration, a MMN is elicited by the frequency but not the duration feature. In this situation, the earliest possible detection of duration deviance was offset by 100 ms after stimulus onset. In Gomes et al., the latency of the MMN was significantly earlier for a control condition in which the deviant had values of intensity and frequency neither of which were present in the standards (hence the MMN could have been elicited by the difference in the individual features of the standards and deviant) than for the condition in which the deviant differed from the standard in its combination of feature values (and hence the MMN could only be elicited on the basis of a difference in the conjunction of the feature values of the standards and deviant). Although the result is in keeping with the proposed rule, it is not known whether Czigler and Winkler's finding for two features applies to two different bases for the MMN to operate upon, nor whether the period of time between when the two bases of operation might have occurred was great enough for inhibition to occur. It would be interesting to know what would happen if the relative difficulty of the discriminations were reversed such that a MMN based on a change in the conjunction of features could occur earlier.

Finally, the data reported here are potentially of consequence to clinical investigations using the MMN. One way to consider the results obtained is that the 70-10-10-10 condition contained three oddball paradigms, one for each of the three features that deviated from the standard. From a practical point of view, this set up means that if the MMN were used in a clinical setting to examine processing of three different features that could elicit it, a considerable saving in time could be achieved: similar results would be obtained if the three features of interest were combined in a 70-10-10-10 condition versus if three separate 90-10 conditions were conducted, one for each feature.

REFERENCES

Campbell, K., Loewy, D., Bastien, C., & Bell, I. (1991). *Mismatch negativity during natural sleep*. Paper presented at the 10th international conference on event-related potential of the brain, Eger, Hungary.

Czigler, I. & Winkler, I. (1996). Preattentive auditory change detection relies on unitary sensory memory representations. *NeuroReport*, *7*, 2413–2417.

Deacon, D., Pilotti, M., & Tinsley, C. (1995). Is the pre-attentive comparison of auditory stimuli effected on the basis of featural or gestalt representations? In G. Karmos, M. Molnar, V. Csépe, Czigler, I., & Desmedt, J. E. (Eds.), *Perspectives of event-related potentials research* (pp. 190–196). Amsterdam: Elsevier.

Csépe, V., Karmos, G., & Molnár, M. (1987). Evoked potential correlates of stimulus deviance during wakefulness and sleep in cat: Animal model of mismatch negativity. *Electroencephalography and Clinical Neurophysiology*, *66*, 571–578.

Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O. M., Pernier, J., & Näätänen, R. (1995). Separate representation of stimulus frequency, intensity and duration in auditory sensory memory: An event-related potential and dipole model analysis. *Journal of Cognitive Neuroscience*, *7*, 133–143.

Gomes, H., Bernstein, R., Ritter, W., Vaughan, H. G., Jr., & Miller, J. (1997). Storage of feature conjunctions in transient auditory memory. *Psychophysiology*, *34*, 712–716.

Gomes, H., Ritter, W., & Vaughan, H. G. Jr. (1995). The nature of pre-attentive storage in the auditory system. *Journal of Cognitive Neuroscience*, *7* , 81–94.

Javitt, D. C., Steinschneider, M., Schroeder, C. E., Vaughan, H. G., Jr., & Arezzo, J. (1994). Detection of stimulus deviance within primate auditory cortex: Intracortical mechanisms of mismatch negativity (MMN) generation. *Brain Research*, *662*, 192–200.

Kropotov, J. D., Näätänen, R., Sevostianov, A. V., Alho, K., Reinikainen, K., & Kropotova, O. V. (1995). Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. *Psychophysiology*, *32*, 418–422.

Levänen, S., Hari, R., McEvoy, L., & Sams, M. (1993). Responses of the human auditory cortex to changes in one vs. two stimulus features. *Experimental Brain Research*, *97*, 177–183.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.

Nousak, J. K., Deacon, D., Ritter, W., & Vaughan, H. G., Jr. (1996). Storage and comparison of information in transient auditory memory. *Cognitive Brain Research*, *4*, 305–317.

Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., & Vaughan, H. G., Jr. (1995). The mismatch negativity of event-related potentials as a probe of transient memory: A review. *Ear and Hearing*, *16*, 52–67.

Sams, M., Alho, K., & Näätänen, R. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology*, 21, 434–441.

Schröger, E. (1995). Processing of auditory deviants with changes in one versus two stimulus dimensions. *Psychophysiology*, *32*, 55–65.

Treisman, A. (1992). Perceiving and re-perceiving objects. *American Psychologist*, *47*, 862–875.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, *742*, 239–252.

Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen. (1990). The effect of small variation of the frequent auditory stimulus on the event-related brain potential to the infrequent stimulus. *Psychophysiology*, *27*, 228–235.

Zeki, S. (1993). *A vision of the brain*. London: Blackwell.

(Received May 22, 1996; Accepted November 14, 1997)