The Nature of Preattentive Storage in the Auditory System

Hilary Gomes
City College of New York and Albert Einstein College of Medicine

Walter Ritter
Albert Einstein College of Medicine and Lehman College

Herbert G. Vaughan, Jr.
Albert Einstein College of Medicine

Abstract

- Event-related potentials were recorded to tones that subjects ignored while reading a book of their choosing. In all conditions, 90% of the tones were 100 msec in duration and 10% of the tones were 170 msec in duration. In a control condition, a customary oddball paradigm was used in which all of the tones were identical except for the longer duration tones. In two conditions, the tones varied over a wide range of tonal frequencies from 700 to 2050 Hz in 10 steps of 150 Hz. In another condition, the tones varied over the same frequencies but also varied in intensity from about 60 to 87 dB in steps of 3 dB. Thus, there was no "standard" tone in the sense of a frequently presented tone that had identical stimulus features. A mismatch negativity (MMN) was elicited in all conditions. The data are discussed in terms of the storage of information in the memory upon which the MMN is based.

INTRODUCTION

In this study, electrophysiological evidence was sought concerning how information is stored in a transient auditory memory. The tool used in this endeavor was an event-related potential (ERP) component termed the mismatch negativity (MMN). The MMN is obtained when there is virtually any discriminable physical change in an otherwise repeating auditory stimulus. Examples of ways in which the MMN can be elicited include changes in the pitch, loudness, duration, or perceived location of an auditory stimulus (for a review, see Näätänen, 1992). On the basis of electrical recordings in monkeys (Javitt, Schroeder, Steinschneider, Arezzo, & Vaughan, 1992) and cats (Cseppe, Karmos, & Molnar, 1987), magnetic recordings in humans (Hari, Hämäläinen, Ilmoniemi, Kaukoranta, Reinkainen, Salminen, Alho, Näätänen, & Sams, 1984), and dipole analysis of scalp-recorded ERPs (Scherg, Vajsar, & Picton, 1989), the neural generator of the MMN has been determined to be located on the supratemporal plane in either primary or immediately adjacent auditory cortex. Consistent with this view, the MMN recorded from the scalp is largest in the frontocentral region and inverts in polarity at the mastoids (Alho, Paavilainen, Reinkainen, Sams, & Näätänen, 1986; Novak, Ritter, Vaughan, & Wiznitzer, 1990). In contrast to later ERP components related to cognitive processes, such as N2 and P3, the MMN is similar in amplitude for predictable as well as unpredictable changes (Scherg et al., 1989). Consequently, expectation on the part of the subject has little effect on the MMN.

The MMN is not generally affected by whether or not subjects attend to the stimuli (Näätänen, 1992; Ritter & Ruchkin, 1992). In selective attention tasks, the MMN elicited by deviant tones is equivalent in amplitude and latency for the relevant and the irrelevant stimuli (Näätänen, Gaillard, & Mantysalo, 1978). Novak, Ritter, and Vaughan (1992) found that in an intramodal selective attention task, the MMN was unaffected by the level of difficulty of discriminating between the attended and unattended tones. Alho, Woods, Algazi, and Näätänen (1992) found that in an intermodal selective attention task, when the subject attended to the visual stimuli the MMN elicited by irrelevant deviant tones was unaffected by the level of difficulty of discrimination between the visual stimuli. In both studies, allocating additional attentional resources to discriminate relevant from irrelevant tones (Novak et al., 1992) or a deviant from a standard visual stimulus (Alho et al., 1992) had no effect on the MMN elicited by deviant auditory stimuli. Also, when a single train of stimuli is presented, as in the "oddball" paradigm where one stimulus (termed the "standard") occurs frequently and another (termed the "deviant")
occurs infrequently, the MMN is not affected by whether subjects actively discriminate between the stimuli or ignore them while reading a book (e.g., Ritter et al., 1992). An exception to this rule is the suppression of the MMN for deviant tones among the irrelevant stimuli when strongly focused attention is required for the performance of an especially difficult selective attention task (Woldorff, Hackley, & Hillyard, 1991). However, when strongly focused attention is not required, the MMN is not usually affected by the direction of attention, as shown by the studies of selective attention mentioned above and by the presence of the MMN in sleep (Alho et al., 1990; Csepe et al., 1987; Campbell, Loewy, Bastien, & Bell, 1991) and even during anesthesia (Csepe, Molnar, Karmos, & Winkler, 1989). Taken together, the data suggest that the MMN reflects preattentive processing associated with comparison of the current stimulus with a memory of the physical characteristics of a preceding, repetitive stimulus. Näätänen (1992) has proposed that the memory in question is sensory memory. Novak et al. (1990) have suggested that the MMN reflects the outcome of a mismatch operation that detects a discrepancy between incoming and stored information. The present study focuses on the nature of the stored information.

There are at least two different ways in which information could be stored in the memory upon which the MMN depends. Stimuli could be stored as a unitary representation (i.e., each stimulus is represented by a composite that contains all of its features), or various features of stimuli could be stored independently of one another. Most previous studies of the MMN are ambiguous with regard to how information is stored. The most common manner in which the MMN has been studied is by presenting two tones in the oddball paradigm, in which one tone occurs frequently and is termed the standard and the other tone, which differs from the standard on one feature, occurs infrequently and is termed the deviant. With this arrangement it is not possible to determine whether the relevant information is stored as units or separately as features, since the deviant differs from the preceding standards both as a gestalt (the relationship of its features) and the single feature by which it differs from the standard. Winkler, Paavilainen, Alho, Reinkainen, Sams, and Näätänen (1990) have suggested that the relevant information is stored as units, each containing all of the physical features of specific stimuli, and Näätänen has concluded that "what is actually stored is sensory information on stimulus events rather than mere static stimulus properties" (Näätänen, 1992, p. 191). However, Nousak, Deacon, Ritter, and Vaughan (submitted) found that the strength of representations of individual features can vary independently of one another, suggesting that features may be stored separately. In an early study, Sams, Alho, and Näätänen (1984) found that if two identical deviants occur in a row, the amplitude of the MMN elicited by the second deviant is about half the size of the MMN elicited by the first deviant. In contrast, Nousak et al. (submitted) found that if two deviants occur in a row, each of which differs from the standard on a different feature, then the amplitude of the MMN elicited by the second deviant is not affected by the presence of the preceding deviant. Thus, whereas Sams et al. (1984) showed that the first of two successive deviants affects the representation of the feature by which it differs from the standard, Nousak et al. (submitted) showed that the first deviant does not affect the representation of other features of the standard. If the strength of the representations of features can vary independently of one another, then it should be possible to vary some features over wide ranges on every trial and, nevertheless, obtain a MMN for infrequent changes in a feature that is kept constant on most trials. In other words, whereas the strength of the representations of the features that constantly vary would be weak, the strength of the representations of features that are identical across most trials would be strong. In a sense, features that are constant across stimuli would take the place of the repeating standard stimulus in the ordinary oddball paradigm.

Winkler et al. (1990) delivered stimuli along analogous lines for a different reason. The MMN has been conceptualized as a change detector that serves the purpose of drawing attention to changes in unattended stimuli (Näätänen, Sams, & Alho, 1986). However, in everyday life repeating stimuli rarely impinge on the auditory system in a constant manner as they can in the laboratory, alterations being incessantly produced by movements of the organism and environmental perturbations. Consequently, Winkler et al. introduced small variations in the intensity of standards tones (or what they termed, "sub-standards") and, in one condition, presented deviant tones that were lower in intensity than all of the sub-standards and, in another condition, presented deviant tones that differed in tonal frequency from that of all of the sub-standards. In both conditions, the sub-standards varied in intensity across nine different levels. It was found that the MMN was elicited by the deviant tones in all conditions, rendering its hypothesized function more ecologically valid. An additional finding was that the MMN elicited by the deviants was smaller than in a control condition where the intensity of the standards was kept constant.

An interpretation offered by Winkler et al. (1990) for their results was that their sub-standards generated discretely stored representations (a second interpretation concerned with the possibility that a single representation was developed based on an average of the sub-standards was later discarded, see Winkler, Paavilainen, & Näätänen, 1992; Winkler, 1993). In support of the interpretation, it was found that the sub-standards themselves elicited small MMNs. It was thought that at least some of the sub-standards elicited MMNs because they differed from the traces of other sub-standards. The decrease in the size of the MMN elicited by the deviants...
in the experimental conditions was assumed to be due to weaker traces of the substandards because the probability of occurrence (and hence opportunity for reinforcement) of any given substandard was a fraction of the probability of the standards in the control condition. Notice that, on this interpretation, what is being reinforced in the experimental conditions are the representations of specific substandards with their particular intensities. If stimulus features were stored independently of one another, then it might be expected that the frequency deviant would elicit equivalent sized MMNs when intensity was varied because, so far as frequency was concerned, it was kept constant across the substandards. Considering frequency by itself, there was no difference with regard to its occurrence in the control and experimental conditions. Thus, it was inferred that the data “might indicate that stimuli keep their integrity in trace formation (their different features are stored as a unit)” (Winkler et al., 1990, p. 233). A corollary of the inference is that “a stimulus differing in any attribute from a trace cannot fully reinforce that trace, even with regard to the matching attributes” (p. 233). On the basis of this interpretation, the reason why the deviant tones elicited MMNs in the experimental conditions, despite the absence of the repeating standard used in the usual oddball paradigm, is that multiple representations were maintained in memory.

There are limits, however, to the number of discrete representations that can be held in the relevant memory that can support evocation of the MMN. One limitation pertains to the amount of time that the memory persists. Estimates vary, but the longest estimate reported to date is about 10 sec (Böttcher-Gandor & Ullsperger, 1992; Sams, Hari, Rif, & Knuttila, 1994). Another limitation is provided by the report of Cowan, Winkler, Teder, and Näätänen (1993) that at least two or three presentations of a standard tone are required before it can provide the basis for a MMN to a deviant tone. One interpretation they offered was that a single presentation of a stimulus does not produce a strong enough representation to provide the basis for elicitation of the MMN. Hence, combining these limits and holding other possible factors aside (such as the maximum number of separate representations the memory can hold), an approximation of the number of discrete representations that can be simultaneously maintained that are capable of providing a basis for MMNs would be a combined function of the rate at which stimuli are delivered and the amount of time between presentations of identical stimuli. In the Winkler et al. (1990) study, the substandards were presented in a random manner at a rate of one every 560 msec. Since on the average there was only about 5 sec between successive presentations of any given substandard, it is possible that individual traces for at least some of their nine substandards could have been maintained long enough to be strengthened by additional presentations.

An intent of our experiment was to present substandards in such a way that, on the basis of the limitations just described, it was not possible for discrete representations of them to be maintained long enough in memory to be reinforced (to be strong enough to provide a basis for MMN evocation). This was accomplished in Condition 4 by using 10 different values of tonal frequency along with 10 different values of intensity, combined randomly with minimal constraints. Successive substandards were different in both tonal frequency and intensity. The tones were presented at a rate of one per second and the mean number of substandards that were presented between two substandards of a given combination of intensity and frequency was 72. The prediction was that, on the hypothesis that the memory stores stimulus features separately, an MMN would be elicited by tones that differed from the substandards along a dimension that was common across all standards (stimulus duration). In all conditions, one out of each 10 successive stimuli was randomly replaced with a tone that was longer in duration. In other words, the substandards varied in tonal frequency and intensity whereas the deviants varied in duration. The stimuli that had the same (short) duration on 90% of the trials are termed the “standard duration tones” and the stimuli that had the longer duration on 10% of the trials are termed the “deviant duration tones.” Subjects ignored the stimuli and read books of their own choosing. The intent of the latter was to minimize attention to the stimuli.

RESULTS

Condition 1 was a control condition in which an ordinary oddball paradigm was employed. There was no variation among the standards. The standard duration tones were identical to one another and the deviant tones differed from the standards in duration. Figure 1 presents the grand mean ERPs at all recording sites averaged across subjects for the standard (thin lines) and deviant (thick lines) tones. In this and all subsequent conditions, the ERPs elicited by the standard duration tones contained a negative component that peaked around 100 msec, was largest at Cz, and inverted in polarity at the mastoids (N1), and a positive component that peaked around 150 msec and was also largest at Cz (P2).

As expected, the deviant tones elicited an MMN. As can be seen in Figure 1, the ERPs elicited by the deviant duration tones (thick lines) separated from those elicited by the standard duration tones in a negative direction around 145 msec, with the waveforms associated with the deviant duration tones becoming more negative at frontocephrental sites but more positive in the same latency region at the mastoids. This difference between the two kinds of tones represents the MMN. The MMN was largest at Fz, Cz, RC1, and LC1, next largest at the mastoids (where it inverted in polarity), small at Pz and Oz (where
Figure 1. Grand mean ERPs elicited in Condition 1 at all recording locations. The thin lines are the ERPs elicited by the standard duration tones and the thick lines are the ERPs elicited by the deviant duration tones. The lack of complete waveforms for some recording sites in this and in Figures 3-5 was due to equipment failure. In this and all subsequent figures, stimuli were presented at time zero.

In Condition 2, the stimuli differed in tonal frequency over a wide range from 700 to 2050 Hz in 10 steps of 150 Hz. The initial sequence of 10 tones was randomly selected without replacement from the 10 values of tonal frequency, but for each of the succeeding 10 tones

the identical sequence was presented. The purpose of the latter arrangement was to reduce the possibility that a memory for a given frequency would persist, since tones of the other nine frequencies intervened between the presentation of two tones of the same frequency. On a random basis, one out of each sequential set of 10 tones was longer in duration. The frequency of the longer duration deviant tones was randomly selected without replacement for each set of 100 trials among the 10 frequencies employed.

Figure 3 presents the grand mean ERPs for the standard duration tones averaged together (thin lines) and deviant duration tones (thick lines). The second panel of Figure 2 shows the grand mean difference waveforms at Fz and the mastoids. As can be seen, an MMN was elicited by the deviant duration tones, which had about the same latency as in Condition 1 and was somewhat smaller in amplitude. The amplitude of the MMN was again significantly different from zero at the six recording sites where it was measured (Table 1).

In Condition 3, the same tonal frequencies were used as in Condition 2, but the sequence in which they were
delivered was randomized across successive sets of 10 stimuli. The purpose of randomizing the sequences of the tonal frequencies was to rule out the possibility that the sequence had been memorized. Under such a circumstance, the deviant duration tone could stand out as different from the standard duration tone in the sense that it was not a member of the sequence.

Figure 4 depicts the grand mean ERPs elicited by the standard duration tones and the deviant duration tones and the bottom panel of Figure 2 presents the grand mean difference waveforms at Fz and the mastoids. As in the previous conditions, the deviant duration tones elicited an MMN of about the same latency. The amplitude of the MMN was again significantly different from zero at all recording sites at which it was measured (Table 1).

A two-way ANOVA was conducted using all six recording sites where the MMN was found to be significantly different from zero (factors of condition and electrode). For the analysis, microvolt values at the mastoids were changed from positive to negative scores. A significant main effect of condition was obtained \( [F(3/27) = 4.98, \quad \varepsilon = 0.935, p = 0.009] \). Tukey post hoc comparisons indicated that at several recording sites the MMN obtained in Condition 1 was significantly larger than one or two of the other three conditions, indicating that the significant main effect was due to a larger MMN when the standards did not vary compared to when they did vary. Also, in only one instance was there a significant difference among Conditions 2-4 (Condition 3 was larger than Conditions 2 and 4 at RC1). Given the large number of comparisons, the probability of a type I error was high. Hence, the significant main effect of condition suggests that the MMN was larger when the standards did not vary.

Table 1. MMN Measured from the Difference Waveforms (Deviant - (Sub)standard) as a Mean Amplitude (μV) during 205-255 msec

<table>
<thead>
<tr>
<th>Condition</th>
<th>Electrode</th>
<th>Amplitude</th>
<th>t(9)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fz</td>
<td>-2.71</td>
<td>-6.71</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>-2.58</td>
<td>-6.04</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>LC1</td>
<td>-2.03</td>
<td>-5.73</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>RC1</td>
<td>-3.68</td>
<td>-5.67</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>LM</td>
<td>1.51</td>
<td>5.43</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>RM</td>
<td>1.45</td>
<td>5.28</td>
<td>0.0002</td>
</tr>
<tr>
<td>2</td>
<td>Fz</td>
<td>-1.75</td>
<td>-4.86</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>-1.89</td>
<td>-4.78</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>LC1</td>
<td>-1.31</td>
<td>-5.79</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>RC1</td>
<td>-1.77</td>
<td>-3.85</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>LM</td>
<td>1.03</td>
<td>2.88</td>
<td>0.0091</td>
</tr>
<tr>
<td></td>
<td>RM</td>
<td>0.91</td>
<td>3.04</td>
<td>0.0070</td>
</tr>
<tr>
<td>3</td>
<td>Fz</td>
<td>-2.47</td>
<td>-5.54</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>-2.47</td>
<td>-5.12</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>LC1</td>
<td>-1.94</td>
<td>-5.26</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>RC1</td>
<td>-2.62</td>
<td>-5.09</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>LM</td>
<td>0.67</td>
<td>2.79</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td>RM</td>
<td>0.56</td>
<td>2.41</td>
<td>0.0197</td>
</tr>
<tr>
<td>4</td>
<td>Fz</td>
<td>-1.87</td>
<td>-6.07</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>-1.88</td>
<td>-5.39</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>LC1</td>
<td>-1.64</td>
<td>-4.95</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>RC1</td>
<td>-1.54</td>
<td>-4.64</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>LM</td>
<td>0.65</td>
<td>2.71</td>
<td>0.0120</td>
</tr>
<tr>
<td></td>
<td>RM</td>
<td>0.98</td>
<td>3.41</td>
<td>0.0039</td>
</tr>
</tbody>
</table>

Gomes et al. 85
of post hoc tests conducted, it was concluded that there was no significant difference in the amplitude of the MMN associated with the order with which the tonal frequencies were delivered (Conditions 2 and 3) or due to standard duration tones varying along one as opposed to two stimulus dimensions (Conditions 2 and 3 versus Condition 4).

The mean peak latency of the MMN at Fz was 216, 234, 221, and 229 msec for Conditions 1, 2, 3, and 4, respectively. A one-way ANOVA, used to determine whether variations in the tonal frequency and intensity of the standard duration tones would affect the timing of the MMN, yielded a nonsignificant result \[ F(3/27) = 1.49, \text{ } \epsilon = 0.763, \text{ } p = 0.25 \].

The design of Condition 4, where successive standard duration tones repeatedly changed in both tonal frequency and intensity, ruled out the possibility that memories could have been made of the standard duration tones (one interpretation offered by Winkler et al., 1990). One remaining concern, however, was the possibility that memories were established of the averaged values of the stimulus dimensions along which the standard duration tones varied (another interpretation offered by Winkler et al., 1990). To examine the possibility that in Conditions 3 and 4 (where the tonal frequency of the standard duration tones were presented in comparable ways) the subjects had developed a memory based on the averaged frequency of the 10 standard duration tones, the ERPs elicited by the standard duration tones of the middle two frequencies were averaged and separately subtracted from the ERPs elicited by each of the standard duration tones of lower and higher frequency. Had the frequencies of the 10 standard duration tones been averaged together by the subjects, then the subtractions should have yielded MMNs.

Figure 6 shows the grand mean ERPs of Condition 3 at Fz, Cz, and the two mastoids for standard duration tones of selected tonal frequencies overlapped with the ERPs of the standard duration tones with the middle two frequencies averaged together. The thick lines depict the ERPs of the tones with the highest frequency (top row: 2050 Hz) and the four lowest frequencies (bottom four rows). The ERPs elicited by these tonal frequencies were selected because they were the only ones to show any observable difference between the ERPs of the standard

Figure 3. Grand mean ERPs elicited in Condition 2. Thick and thin lines as in Figure 2.
duration tones elicited by the middle two frequencies and the other frequencies. In all rows, the thin lines represent the averaged ERPs for the standard duration tones with the middle two frequencies. As can be seen, the only substantial differences between the thick and thin waveforms were in the latency region of P2, where the thick lines were more positive than the thin lines (the opposite of what would occur were these differences due to MMNs). In the bottom row, N1 and P2 were larger at Fz and Cz for the tones with the lowest frequency (thick lines) than for the tones with the middle two frequencies (thin lines). At both mastoids, N1 inverted in polarity and was larger for the lower frequency than the middle two frequency tones. A similar pattern of results can be seen in the third row, except that the differences between N1 were diminished. In the second and third rows, the pattern of results just described was even smaller, except that the difference in the amplitude of P2 persisted. The thick lines in the top row, which display the ERPs elicited by the tones with the highest frequency, exhibit a much smaller pattern of N1 and P2 results than was found for the same comparison for the tones with the lowest frequency (bottom row). The ERPs of the remaining standard duration tones are not shown because N1 and P2 differences were even smaller or unobservable for similar comparisons. The same overall pattern of results was obtained for the tonal frequencies of the standard duration tones in Condition 4.

Sustained Potentials

Long duration tones elicit a sustained potential that has a topography generally similar to that of the MMN (e.g., it is negative at the midline and inverts in polarity at the mastoid) (Picton, Woods, & Proulx, 1978). It is possible that the longer deviant duration tones were associated with a more extended sustained potential than the shorter standard duration tones. Consequently, the subtraction used to delineate the MMN could have produced difference waveforms that had the appearance of the MMN but instead were due to the sustained potential. To examine this possibility, six subjects were presented with complementary runs, one set where the standard duration tones lasted for 100 msec and the deviant duration tones 170 msec, and another set where the standard duration tones lasted for 170 msec and the deviant...
duration tones 100 msec. This was done for Condition 1, where the standards did not vary in any way, and Condition 4, where tones varied on every trial in frequency and intensity. The MMN was delineated by subtracting the ERPs associated with tones of a given duration when they were standards from when they were deviants, separately for each condition. In this way, the amount of sustained potential would be equivalent and cancel in the subtraction. Since the subjects were reading throughout, it was deemed appropriate to subtract ERPs across runs of a given condition because systematic changes of psychological state across the counterbalanced runs were unlikely to affect the ERPs.

For Condition 1, comparable difference waveforms were obtained both for the 100- and 170-msec duration tones as in the main experiment. Figure 7 shows the ERPs obtained at Fz, Cz, LM, and RM for Condition 4 in the run where the standard duration tones lasted for 100 msec (thin lines) and the run where the deviant duration tones lasted 100 msec (thick lines). In the latency range of the MMN, the ERPs elicited by the deviant duration tones were more negative than those elicited by the standard duration tones at the midline, and the reverse was true at the mastoids. The difference waveforms shown at the bottom of the figure for Fz (thick line) and the mastoids (thin lines) are similar to those depicted in Figure 2 in morphology, polarity, and latency. Similar results were obtained for the 170-msec tones. These data cannot be accounted for by differences in the sustained potential. As in the main experiment, the amplitude of the MMN was larger in Condition 1, where the standards did not vary, compared to Condition 4, where the tones varied in frequency and intensity on every trial.

**DISCUSSION**

**The Standard Duration Tones**

In Condition 3, where the standard duration tones varied in frequency, the ERPs elicited by the standard duration tones with the middle two tonal frequencies were averaged and subtracted from the ERPs elicited by each of the standard duration tones of higher and lower frequency to determine whether the standard duration tones elicited MMNs (see Fig. 6). The pattern of results obtained was that the ERPs of the standard duration tones with the extreme frequencies had larger N1 and
P2 components than those of the standard duration tones with the middle two frequencies, and that this effect was greater for the lower tonal frequencies. For the higher frequency tones the N1 effect was confined to the mastoids. Since N1 and the MMN have almost identical scalp distributions, it is possible that the greater negativity in the N1 latency region for the standard duration tones with the extreme tonal frequencies could be attributed to the presence of an MMN. This would be especially true for the standard duration tones of the extreme tonal frequencies because the latency of the MMN is inversely related to the physical separation between standards and deviants (Näätänen & Gaillard, 1983). The point is that shorter latency MMNs can temporally overlap the N1 (e.g., Scherg et al., 1989). However, as can be seen in Figure 6, the standard duration tones with the extreme frequencies also had larger P2 components than the standard duration tones with the two middle frequencies. To our knowledge, ERPs associated with deviant stimuli that elicit the MMN do not display larger P2 components than the standard tones. Moreover, a straightforward interpretation of the present results is that the enhancement of the N1 and P2 components of the standard duration tones with the extreme frequencies is due to frequency-specific refractoriness (Butler, 1968). Since the standard duration tones with the extreme frequencies were, on the average, preceded by stimuli with more distant frequencies than was so for the standard duration tones of the two middle frequencies, the effects of refractoriness would be less for the former than the latter. Although studies of frequency-specific refractory effects have been essentially confined to measurements of N1, Figures 2 and 3 of Näätänen et al. (1988) clearly show that similar effects are obtained for P2. In addition, the greater effects of N1 enhancement for the standard duration tones with the lower as contrasted to the higher frequencies can be attributed to the general finding (reviewed in Jacobson, Lombard, Gibbens, Ahmad, & Newman, 1992) that N1 amplitude is inversely related to tonal frequency.

Figure 6. Grand mean ERPs elicited by the standard duration tones in Condition 3. In all rows, the thin lines are the ERPs elicited by the tones of the two middle frequencies, averaged together. In the top row, the thick lines are the ERPs elicited by the highest frequency tones (2050 Hz). In the bottom four rows, the thick lines are the ERPs elicited by the tones with the lowest (700 Hz), second from lowest (850 Hz), third from lowest (1000 Hz), and fourth from lowest frequency (1150 Hz).
In contrast to Winkler et al. (1990), therefore, our standard duration tones (what they called substandards) did not elicit MMNs. This indicates that representations of standard duration tones, with their specific values of frequency and intensity, were not reinforced such that they could provide the basis for MMNs elicited by standard duration tones with other frequency and/or intensity values, as in Winkler et al. (1990).

The Deviant Duration Tones

MMNs were obtained for the deviant duration tones in all conditions despite the extensive variations in the frequency and intensity of the standard duration tones. The circumstance that the standard duration tones did not elicit MMNs supports the conclusion of Winkler et al. (1992) that evocation of the MMN by deviant tones when the standard tones are varied (become substandards) is not due to an averaging process that produces a single representation of the substandards. The alternative interpretation offered by Winkler et al. (1990) for their data, that multiple representations of substandards could account for the evocation of the MMN by deviant tones, also does not apply to the present results. In Conditions 2 and 3 an average of 10 sec elapsed between two presentations of a standard duration tone of the same frequency. While it is conceivable that multiple representations could have been reinforced anyway, had they been, the standard duration tones should have elicited MMNs.

Examination of the standard duration tones in Condition 4, where both frequency and intensity varied, yielded the same results for frequency as in Condition 3. But in addition to the lack of evidence for representations based on averages, a given combination of tonal frequency and intensity for the standard duration tones occurred, on the average, once in 72 trials. Thus, if it is assumed that two or three repetitions of a given stimulus event (defined by its specific combination of values of various stimulus features) is necessary for its representation to provide the basis of an MMN (Cowan et al., 1993), there would have been no representations in sensory memory of the identical repeated stimuli required to elicit an MMN. The data of Conditions 2, 3, and 4, therefore, are consistent with the hypothesis that stimulus features are stored independently of one another in the memory upon which the MMN depends, and that their representations are capable of evoking the MMN. In support of that conclusion, Giard, Lavikainen, Reinkainen, Perrin, and Näätänen (1991) reported differ-

![Figure 7. Grand mean ERPs elicited at Fz, Cz, LM, and RM in Condition 4 for the 100-msec duration tones. Thin lines are taken from the run where these tones were the standard duration; thick lines are taken from the run where they were the deviant duration. MMN, deviant minus standard ERPs obtained at Fz (thick line) and the mastoids (thin lines).](image-url)
ent topographic distributions for MMNs elicited by deviant tones that differed from standard tones along three stimulus dimensions. The hypothesis guiding their study was that the “memory codes the different physical attributes (frequency, intensity, duration) of an acoustic stimulus in separate locations of the auditory cortex” and that “the neural generators for the frequency-, intensity- and duration-MMNs are located in separate cortical areas.” There is also ecological validity in the notion that the system underlying the MMN would track infrequent changes of particular stimulus dimensions despite wide variability in other stimulus dimensions. If a complex series of sounds come from a general direction (say the sounds of family members or young, which could vary extensively in pitch, loudness, etc.), detection of a sound that comes from a different direction would be useful.

Näätänen (1992) has suggested that the memory upon which the MMN depends is sensory memory. On the basis of an extensive review of the literature, Cowan (1984) divided auditory sensory memory into two phases, a literal store that lasts about 200–300 msec followed by a longer store that might last as long as 20 sec. It is important to recognize that on the assumption that the memory upon which the MMN depends is sensory memory, then in the present experiment (and most other studies that examined the MMN) the relevant memory was maintained in the longer phase of sensory memory, since there was 1 sec between tones. Thus, the interpretation that a feature (stimulus duration) was stored independently of its carrier in the present study does not apply to literal memory, since transformations of information could occur between what is stored in literal memory and the longer phase of sensory memory.

There is no implication in the present paper that stimulus features are stored in the relevant memory only independently of one another. We find the inferences drawn by Winkler et al. (1990) from their data that gestalt representations of stimuli are stored in the memory convincing. Recently, Gomes, Bernstein, Ritter, Vaughan, and Miller (in preparation) obtained data that directly support gestalt storage. If gestalt representations of stimuli are stored in the memory, then the data of the present study need not necessarily be interpreted in terms of completely independent storage of features. Though the gestalt representations of the standard duration stimuli may not have been reinforced through repetition, the individual representations of several of the most recent standard duration stimuli are nevertheless likely to have been present at the time when deviant duration stimuli were delivered. Winkler et al. (1992) have shown that at least two representations can be simultaneously present in the memory upon which the MMN is based. There is an implication in Gomes et al. (submitted) that at least three representations can be simultaneously stored. An alternative hypothesis to independent storage of features is that the comparator mechanism that underlies the
MMN is capable of analyzing features across gestalt representations extant in the memory. The hypothesis is schematically presented in Figure 8. The squares stand for gestalt representations of particular stimuli encountered in the recent past (time moves from the lower left to the upper right). The round circles at the corners of each square stand for the representation of an arbitrary number of feature dimensions (e.g., frequency might be represented in the upper left circles and loudness in the upper right circles). The bars that connect the circles of each square provide for the gestalt representation of each stimulus (i.e., the relationship between the features of a given stimulus). The bars that connect the circles in the same location from square to square provide the basis for an analysis of particular features across the gestalt representations. The thickness of the bars are intended to depict the strength or accessability of a given feature or gestalt. The hypothesis entails the assumption that the strength of gestalt representations can vary independently of the strength of representations for specific features. To illustrate, if a tone occurs that has identical stimulus parameters as one or more of the stimuli represented in the memory, then both that gestalt representation and the representation of similar features in the memory are strengthened. However, if a tone does not have identical stimulus parameters as any of the representations in the memory, then only the representations of similar features in the memory are strengthened. In the present study, this would mean that the standard duration tones could not have strengthened any gestalt representations in the memory, but they could strengthen the representation for specific features (in this case, duration).

Whatever conceptualization is used to explain how the memory works, it appears clear that the MMN not only responds to changes in "stimulus events" (i.e., gestalts) but also to changes in stimulus features independent of the nature of the gestalt characteristics of the carrier stimuli. Thus, there appear to be at least two aspects to the manner in which stimuli are analyzed by the comparator mechanism underlying the MMN. When an ordinary oddball paradigm is used, as in Condition 1, the MMN could be triggered both on the basis of the difference between the gestalt of the deviant and that of the standards as well as on the basis of the difference between the feature by which the deviant differed from the standards. In Conditions 2-4, the MMN could be triggered only on the basis of the difference between the feature by which the deviant duration tones differed from the standard duration tones. It could be that the reason why the amplitude of the MMN was larger in Condition 1, where an ordinary oddball paradigm was used, than in Conditions 2-4, where tones varied widely from trial to trial,1 was because in Condition 1 it was triggered on the basis of both gestalt and feature analyses whereas in Conditions 2-4 the MMN could be triggered only on the basis of a feature analysis.

METHOD

Subjects

Six female and four male subjects between 18 and 40 years of age were paid for their participation in the experiment. Seven of the subjects had participated in previous ERP experiments.

Experimental Procedure

The subjects sat in a comfortable chair and ignored the stimuli, reading a book of their own choice. The stimuli were pure tones presented binaurally with insert earphones at a rate of one per second. The tones were either 100-msec standard duration tones or 170-msec deviant duration tones. There were three runs of 500 trials in each condition, a run consisting of 450 standards and 50 deviants. The runs lasted approximately 7 min with 1-2 min between runs. The order of the conditions was counterbalanced for each subject (e.g., 1, 2, 3, 4, 4, 3, 2, 1), with the starting condition varied across subjects.

Condition 1

This was an ordinary oddball paradigm with the deviant tone being longer in duration than the standard tones. All of the stimuli were 1000 Hz and 74 dB peak equivalent SPL. There was one deviant for each successive group of 10 trials, with the position of the deviant of each group selected randomly. Thus, the probability of the deviant was 10%.

Condition 2

Each successive group of 10 trials contained stimuli of 10 different tonal frequencies ranging from 700 to 2050 Hz in intervals of 150 Hz, with an average peak equivalent SPL of 75 dB. The sequence of frequencies for the first group of 10 trials was randomly selected, however, all subsequent groups of 10 stimuli had the identical sequence of frequencies. The frequency of the longer duration deviant tone for each group of 10 stimuli was randomly selected without replacement for each set of 100 trials. Thus, each frequency occurred on 10% of the trials and each frequency was assigned to 10% of the deviant tones.

Condition 3

The same stimuli were used as in Condition 2, except that the ordering of tonal frequencies of each sequence of 10 trials was random up to a sequence of 100 trials, after which the same sequence of 100 trials was repeated to the end of the run.
**Condition 4**

In addition to the stimulus values employed in Condition 3, the intensity of the tones ranged from about 60 to 87 dB peak equivalent SPL, with intervals of approximately 3 dB. Thus, there were 10 values for intensity and 10 for tonal frequency. The difference in intensity between any two successive tones was never less than 6 dB. Each intensity and frequency occurred once in every 10 stimuli on a random basis, up to a sequence of 100 trials, after which the same sequence was repeated to the end of a run. A given combination of values for the two stimulus dimensions for the standard duration tones occurred, on the average, once in 72 trials with a range from 8 to 100 trials. For each sequence of 10 trials in each run, a deviant duration tone occurred in one of the 10 positions, randomly selected without replacement.

**ERP Recording**

The electrical activity (filtered using bandpasses of 0.1–40 Hz) was recorded along the midline at Fz, Cz, Pz, and Oz. In addition, lateral electrodes were placed along a coronal chain from Fz to each mastoid consisting of electrodes one-third of the distance (LC1 and RC1 for the left and right coronal chains, respectively), two-thirds of the distance (LC2 and RC2), and the right (RM) and left (LM) mastoids. All recordings were referenced to the nose. Ocular potentials were monitored with bipolar electrodes above and below the left eye and at the right and left outer canthi, respectively. Trials on which electrical activity exceeded ±100 µV at all but the horizontal EOG recordings were automatically rejected. All recordings were subsequently visually assessed for residual artifact. Recordings began 100 msec prior to stimulus presentation and extended poststimulus for 500 msec.

**Data Analysis**

For each subject, the three runs of each condition were combined. Measurements of the MMN were obtained at Fz, Cz, RC1, LC1, RM, and LM in difference waveforms obtained for each subject in all conditions by subtracting the ERPs elicited by the 100-msec tones from the ERPs elicited by the 170-msec tones. Grand mean ERPs, averaged across subjects for a given condition, were obtained for purposes of display and selection of a latency window for amplitude and latency measurements. Across conditions, the peak latency of the MMN in the grand mean difference waveforms was about 230 msec. For the individual subjects, the amplitude of the MMN in each condition was measured as the mean voltage across a latency window from 25 msec before to 25 msec after its peak latency (205–255). The peak latency of the MMN was defined as the most negative voltage point between 170 and 270 msec poststimulus at Fz. If there was more than one candidate peak, a point half way between the peaks was chosen.

To establish the presence of the MMN at the six recording sites where its amplitude was measured, t tests for correlated data were conducted for each condition to determine whether the mean amplitude across subjects was significantly different from zero. Since it was predicted that the MMN would be obtained in all conditions, a one-tailed test was employed.

One-way ANOVAs for repeated measures were conducted to compare the amplitude and latency of the MMN at Fz across the four conditions. The Geisser-Greenhouse correction was used in reporting p values based on ANOVAs. Where appropriate, Tukey post-hoc tests were applied. An alpha level of 0.05 was used for all statistical tests.

**Acknowledgments**

This research was supported by Grants NS30029, MH06823, and HD01799 from the USPHS. We thank Nelson Cowan, Diana Deacon, Terry Picton, Daniel Ruchkin, and various members of the ERP laboratory at the Albert Einstein College of Medicine for helpful comments.

Reprint requests should be sent to Dr. Hilary Gomes, Department of Neuroscience, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461.

**Note**

1. After this paper was accepted for publication, an experiment came to our attention that used similar methods and obtained comparable results (Huotilainen et al., 1993).

**REFERENCES**


*Gomes et al.* 93


