

Preattentive Binding of Auditory and Visual Stimulus Features

István Winkler^{1,2}, István Czigler¹, Elyse Sussman³,
János Horváth¹, and László Balázs¹

Abstract

■ We investigated the role of attention in feature binding in the auditory and the visual modality. One auditory and one visual experiment used the mismatch negativity (MMN and vMMN, respectively) event-related potential to index the memory representations created from stimulus sequences, which were either task-relevant and, therefore, attended or task-irrelevant and ignored. In the latter case, the primary task was a continuous demanding within-modality task. The test sequences were composed of two frequently occurring stimuli, which differed from each other in two stimulus features (standard stimuli) and two infrequently occurring stimuli (deviants), which combined one feature from one standard stimulus with the other feature of the other standard stimulus. Deviant stimuli elicited MMN responses of similar parameters across the different attentional conditions. These results

suggest that the memory representations involved in the MMN deviance detection response encoded the frequently occurring feature combinations whether or not the test sequences were attended. A possible alternative to the memory-based interpretation of the visual results, the elicitation of the McCollough color-contingent aftereffect, was ruled out by the results of our third experiment. The current results are compared with those supporting the attentive feature integration theory. We conclude that (1) with comparable stimulus paradigms, similar results have been obtained in the two modalities, (2) there exist preattentive processes of feature binding, however, (3) conjoining features within rich arrays of objects under time pressure and/or long-term retention of the feature-conjoined memory representations may require attentive processes. ■

INTRODUCTION

A key question of modern theories of perception is the role of attention in processing sensory information. Despite 40 years of research dedicated to this issue, there is still no consensus about the extent to which the unattended sensory input is processed in the human brain. One long-standing debate focuses on whether the integration of sensory features to unitary code (also termed feature binding) occurs only for attended stimuli or if it also occurs for stimuli outside the focus of attention. Although this issue has been mainly considered in visual perception, it is also highly relevant in the auditory modality. Both the segregation of auditory sources and the identification of visual objects require the analysis of feature conjunctions.

The need for attentive processes in conjoining visual stimulus features has been first suggested in the now-classic “feature integration theory” (Treisman, 1993; Treisman & Gelade, 1980). According to this theory, stimulus features (such as color, shape, etc.) are pre-

attentively analyzed in parallel and linked to a master map of the visual space. There is abundant support for parallel preattentive feature analysis both from behavioral and neurophysiological research (Marr, 1982). The key suggestion of Treisman’s attentive feature integration theory is that integration between features is a serial process limited to one spatial location at a time. The role of spatial attention is to select the location for feature integration. This claim receives important support from evidence showing that illusory conjunctions can emerge outside the focus of attention (Treisman & Schmidt, 1992). That is, when subjects are prevented from exploring the whole display (e.g., by short presentation times), features appearing on separate objects within the display may be erroneously perceived as co-occurring on a single object. The possibility that features from different objects may be erroneously conjoined when the objects fall outside the focus of attention suggests that feature binding processes cannot operate normally without focused attention. Further, in contrast to the fast parallel search for individual features, searching for objects combining two features is a slow serial process (e.g., Treisman & Gelade, 1980). However, these and other evidence cited in support of attentive feature integration theories have also alternative explanations

¹Hungarian Academy of Sciences, ²Helsinki University, ³Albert Einstein College of Medicine, New York

(e.g., Duncan & Humphreys, 1989; Wolfe, Cave, & Franzel, 1989).

According to “object-based accounts” of visual perception (for a review, see Scholl, 2001), the locus of attentional effects is beyond the formation of perceptual units (objects). A strong argument for this claim is that feature detection is more efficient within a single object than between different objects (“same-object advantage”; e.g., Awh, Dhaliwal, Christensen, & Matsukura, 2001; Duncan, 1984). Furthermore, identity of similar visual objects is preserved after dislocation of the objects (Pylyshyn & Storm, 1988) and some forms of visual neglect appear to be object-based (for a review, see Rafal, 1996).

Much less research has been conducted regarding feature binding in the auditory modality. As with the visual system, there is evidence that different neuronal populations respond to the different auditory features (Giard et al., 1995; Pantev, Hoke, Lehnertz, & Lütkenhöner, 1989; Pantev, Hoke, Lehnertz, Lütkenhöner, Anogianakis, et al., 1988). However, there are scarce data concerning the role of attention in auditory feature binding.¹ The most relevant auditory studies on this topic have tested the emergence of illusory feature conjunctions. Emergence of illusory conjunctions between auditory features has been reported under conditions of fast dichotic stimulation (Cutting, 1976; Deutsch, 1975; Effron & Yund, 1974) and when two or more tones were presented simultaneously from different illusory sound sources (“sound sources” were lateralized by manipulating the interaural time difference of sounds presented through headphones; Hall, Pastore, Acker, & Huang, 2000). However, the conditions employed in these studies were not optimal for separating the test sounds in space. For example, Hall et al. (2000) presented, in separate conditions, arrays of the two or four sounds delivered simultaneously. Individual sounds were 2-sec long single-tone recordings of five different musical instruments. The sounds presented together in a single auditory array differed both in pitch and timbre (i.e., a different note was played by each instrument). In each trial, a cue sound (a single tone presented by one of the five instruments) followed the test array after a 500-msec long silent interval. Participants were asked to judge whether the combination of pitch and timbre in a cue sound was present in the preceding sound array. The pattern of errors indicated the presence of illusory conjunctions between pitch and timbre: Subjects relatively often indicated that the cue sound was present in the array when both the instruments playing the cue sound and the pitch of the cue sound appeared in the array in separate sounds. It is, however, possible that the segregation of the two or four test sounds was incomplete because of the single-cue approximation of spatial location, the possible spectral overlap between sounds, and the lack of context (i.e., the build-up time of auditory streams can exceed 2 sec; see Bregman,

1978). Therefore, the emergence of illusory conjunctions in Hall et al.’s study may not have reflected attentional capacity limitations, but possibly insufficient information to correctly assign features to the lateralized auditory sources.

Two studies (Thompson, Hall, & Pressing, 2001; Woods, Alain, & Ogawa, 1998) presented sounds sequentially, asking participants to judge whether sounds with a given combination of features appeared in the sequence. The results of these studies are contradictory. Whereas Woods et al.’s results indicated that auditory feature conjunctions were processed in parallel along with the analysis of the individual auditory features, Thompson et al.’s results suggested serial processing of auditory feature conjunctions by showing evidence for illusory conjunctions.

A critical factor common to these studies was that participants were asked to judge whether a sound with two specific feature levels has appeared in a given sequence. Participants may have looked for the presence of one feature, and then tried to decide whether the other target feature also matched the criterion. This procedure is prone to illusory conjunction type of errors, especially, when the memory trace of the target template can interfere with the memory of the test sounds (such as in Thompson et al., 2001; Hall et al., 2000). Using a measure that does not require participants to perform some task with the sounds would allow a better test of the question of whether attention is needed for feature integration. With this approach, we tested feature binding for both the auditory and visual modalities in separate experiments.

EXPERIMENT 1: MMN AND AUDITORY FEATURE CONJUNCTIONS

The Mismatch Negativity Event-related Brain Potential

For testing the formation of auditory feature conjunctions, we used the mismatch negativity (MMN) component of event-related brain potentials (ERPs; for recent reviews of MMN, see Picton, Alain, Otten, & Ritter, 2000; Näätänen & Winkler, 1999). ERPs have high temporal resolution and can provide information about the timing of cognitive processes associated with stimulus events. The MMN response is particularly useful for the present purpose because it does not require the experimental participant to respond to the sounds or to indicate his/her perception of them. When measuring MMN, participants are often instructed to disregard the auditory stimuli and to read a book, watch a movie, or play a video game (this is termed the “passive condition”). Although MMN is widely regarded as a reflection of preattentive processing (for recent supporting evidence, see Sussman, Winkler, & Wang, 2003),² it is important to note that the elicitation of MMN, in and of itself, does not constitute

grounds for arguing that all processes preceding MMN elicitation are also preattentive (cf. Sussman, Winkler, Huotilainen, Ritter, & Näätänen, 2002).

MMN is elicited when a sound is detected as violating some regularity of the preceding auditory sequence. Thus, it is elicited by infrequent changes in simple repetitive sound features (such as frequency, intensity, duration, location, etc.) and also by violations of structural and sequential regularities (for a review of the higher-level functions displayed by MMN, see Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). It has been established that MMN is based on a representation of the regularities extracted from the acoustic stimulation (Näätänen & Winkler, 1999). MMN is therefore not elicited by stimulus change per se or by the occurrence of rare stimuli not preceded by a sequence of regularly occurring stimuli. Thus, MMN can be used not only to study which auditory events were treated by the brain as being irregular, but also what auditory regularities were detected and represented in the auditory system.

In passive conditions, when participants performed a visual primary task and were instructed to ignore the sounds, MMN has been elicited by irregular conjunctions of auditory features (Takegata, Huotilainen, Rinne, Winkler, & Näätänen, 2001; Takegata, Paavilainen, Näätänen, & Winkler, 1999, 2001; Sussman, Gomes, Nousak, Ritter, & Vaughan, 1998; Gomes, Bernstein, Ritter, Vaughan, & Miller, 1997). For example, in Sussman, Gomes, et al.'s (1998) study, subjects were presented with sequences containing four types of tones. Three of the four types of tones appeared within the sequence with equal, $p = .3$ probability ("standards"). Each of the three standard tones differed from the other two in frequency and the location of origin (the sound source). The fourth tone (deviant), which was presented with .1 probability, had the frequency of one of the standard tones and the source location of another. The order of the tones within the sequence was randomized. Because the deviant introduced no feature that was not also present in one of the standard tones, it would only elicit MMN if the frequently occurring conjunctions between frequency and location have been detected and encoded in the memory representations involved in the MMN-generating process. MMN was elicited by the deviant tones suggesting the combinations of tone frequency and source location have been established both for the standard and the deviant tones. Although the results of this and similar other MMN studies are compatible with the notion of preattentive feature integration, they do not constitute a strong argument for this hypothesis. This is because these studies did not provide a strict control of attention. Using an uncontrolled primary visual task (such as reading) does not rule out the possibility that experimental participants could have attended the auditory stimuli and, as was already noted, the presence of MMN in and of itself does not prove that

the processes underlying it are preattentive. Furthermore, some theorists would argue that the attentional capacities for visual and auditory processing may be, to some extent, independent from each other and, therefore, performing a primary visual task would not prevent participants from also attending to sounds (see, e.g., Duncan, Martens, & Ward, 1997).

The present study used the MMN method to test the formation of feature conjunctions. However, in contrast to previous MMN studies, the present study was designed to provide a more stringent control of attention (a) by introducing a demanding primary *auditory* task and (b) by varying the demand on the participant's attention using different task instructions across conditions.

Procedure for the Auditory Experiment

The Test Tone-Sequences

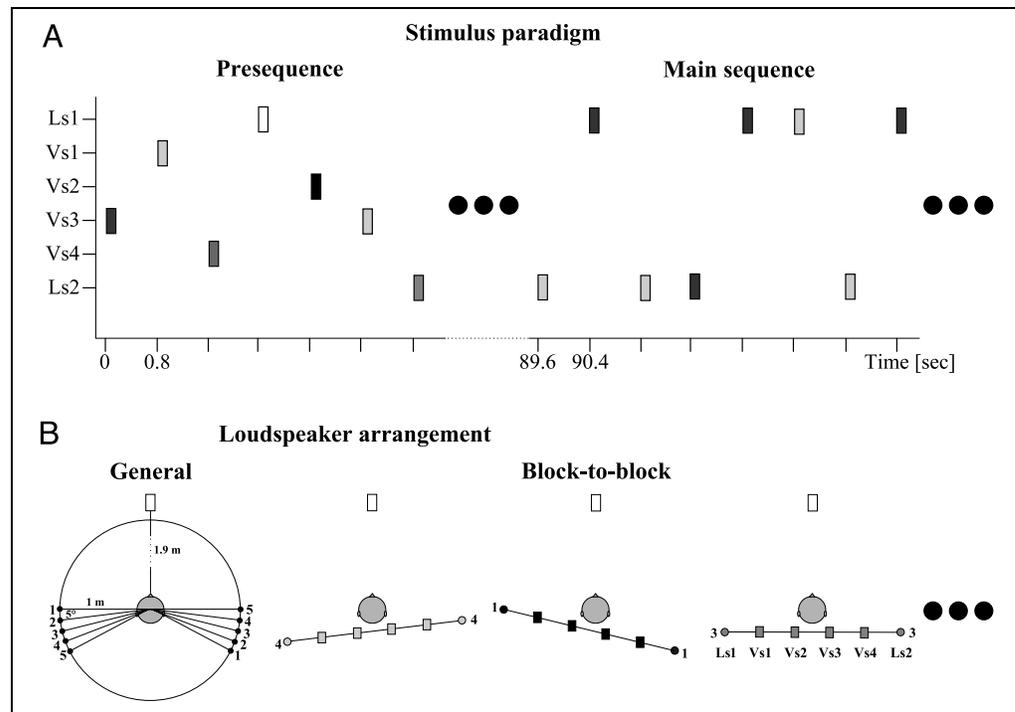
Four types of tones were presented in the main part of the test sequences in randomized order (see Figure 1, Panel A, right side: the "Main sequence"). Forty-five percent of the tones had one of four possible combinations of two frequencies \times two loudspeakers. Another 45% of the tones had the opposite combination of frequency and sound source (i.e., different frequency and different loudspeaker). These two types of tones were the "standards" in these "oddball" sequences. The remaining 10% of the tones (deviants) had, with equal probability, the other two possible combinations of frequency and loudspeaker.

Experimental Conditions

Four conditions were administered. In the "Attend Noise" (AN) and "No Tones" (NT) conditions, participants were instructed to detect slight intensity changes in a continuous stream of noise delivered by a loudspeaker placed in front of them. Participants were strongly motivated to do this task as well as they could by promising substantial bonus payments linked with their performance in the noise change detection task. In the AN condition, task-irrelevant tone sequences (the test sequences, see below) were delivered to participants by two loudspeakers placed behind their head. Participants were informed about the presence of these tones and were instructed to ignore them. In the NT condition, the tones were absent. This condition served as a control, providing information about the level of performance in the noise change detection task.

The remaining two conditions provided a comparison with previous MMN studies of feature conjunction that used the passive situation. The first of these two conditions was a commonly used passive condition, in which subjects were instructed to watch a movie presented on a TV screen in front of them and disregard sounds (both noise and tones). This condition is termed the "Attend

Figure 1. Experiment 1: Schematic illustration of the stimulus paradigm. (A) Each stimulus block consisted of two parts, the “presequence” (left) and the “main sequence” (right), which were presented without a break (uniformly 800-msec SOA throughout the stimulus block). The x-axis represents time (with the change from the presequence to the main sequence marked), the y-axis distinguishes the perceived sound source directions. The two extreme positions, Ls1 and Ls2 denote two physical loudspeakers, whereas the positions between Vs1 and Vs4 have been created by cross-fading the signal between the two loudspeakers (virtual sound sources). Tones are marked by small rectangle, the shade of gray filling (six different levels) representing tone frequency. Note that the four combinations of frequency and location in the main sequence (two frequent, the standards, and two infrequent ones, the deviants) do not appear in the presequence, although the two frequencies and the two locations do. (B) The five possible location-pairs of loudspeakers are shown in relation to the participant’s head (left side); loudspeakers forming a pair are marked with identical numbers. The loudspeaker used for the noise change detection task was set directly in front of the participant. The middle/right side of the figure illustrates how loudspeaker positions were changed from stimulus block to stimulus block. Circles at the ends of the lines correspond to the loudspeaker positions marked on the left side, as well as to Ls1 and Ls2 on Panel A. The four rectangles placed equidistantly on the lines depict the four virtual sound source locations created (Vs1 to Vs4). The shading of the sound sources refers to another feature of the design: also the set of tone frequencies (six individual frequencies in each set, see Panel A) changed from block to block; five sets of frequencies were used.



Video” (AV) condition. The other condition tested what happened if, despite instructions, subjects were to divide their attention between the primary task and the task-irrelevant sounds, as it is often assumed in arguments against the use of passive conditions. This condition was identical to the AN condition, except that we instructed participants to pay some attention also to the tone sequence. This condition is termed the “Attend Noise and Tones” (ANT) condition. Before starting the ANT condition (after all other conditions had been completed), participants were informed that after each stimulus block they would be asked a question about the tones they heard during the stimulus block. However, it was also made clear to them that correctly answering the questions was secondary to reaching good performance in the noise change detection task. Because asking questions about the tones was primarily used to direct some attention to the tones, most questions simply tested whether participants remembered which tones were presented in the preceding stimulus block, although one of the questions checked the emergence of illusory feature conjunctions. In summary, in this condition, subjects had a formal reason to allocate some attention to the test sounds, although the motivation to do so was

not too strong. This way the ANT condition modeled an assumed division of attention in passive conditions.

Predictions

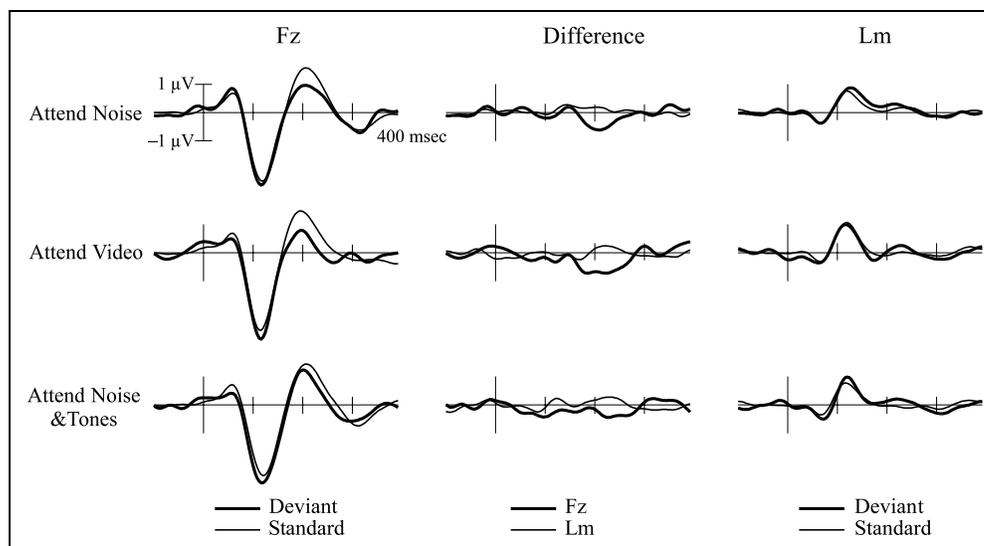
If auditory features were conjoined preattentively, MMNs of approximately equal parameters should be elicited in the AN, AV, and ANT conditions. If, on the other hand, the formation of auditory feature conjunctions requires focused attention, then no MMN should be elicited when subjects strongly focus their attention on the task-relevant sound channel and ignore the test sounds (AN condition), or, at least the MMN amplitude in this condition should be much lower than in the passive (AV) and the divided attention (ANT) conditions.

Results

Event-related Potentials

Figure 2 shows the ERP responses elicited by the standard and deviant tones (frequent and infrequent combinations of frequency and source location) at the Fz (frontal midline) and Lm (left mastoid) electrode

Figure 2. Experiment 1: Grand-average ($n = 12$) frontal (Fz, left side) and left mastoid (LM, right side) ERP responses to the standard (thin line) and deviant tones (thick line) in the AN, AV, and ANT conditions. The corresponding deviant-minus-standard difference curves are shown in the middle column (thick line: Fz, thin line: LM). MMNs of approximately equal size were elicited in all three experimental conditions.



locations in the three conditions presenting the tone sequences (AN, AV, ANT). Following the obligatory ERP responses (N1 and P2, see, e.g., Näätänen & Picton, 1987), the deviant-stimulus response was negatively displaced at frontal electrodes (positively at the mastoids) compared with the standard-stimulus response (see Figure 2). There was a significant main effect of stimulus type in the ANOVA analysis [Condition (AN vs. AV vs. ANT) \times Stimulus (Standard vs. Deviant) \times Electrode (F3 vs. Fz vs. F4)], showing a difference between the deviant- and standard-stimulus ERPs [$F(1,11) = 15.93, p < .01$; see Table 1 for mean amplitudes]. The only other significant effect obtained in the ANOVA was that for the main electrode factor [$F(2,22) = 5.00$; Greenhouse–Geisser $\epsilon = 0.67, p < .05$], representing the scalp distribution of the ERP responses in the measurement period. The latency and scalp distribution of the deviant-minus-standard difference waveforms (negative at the front, positive at the mastoids; was somewhat higher in amplitude at right than left fronto-central scalp locations) are compatible with the notion that MMN was elicited by the deviant tones. Neither the main condition factor nor any of the interactions showed significant effects. Because no significant effect of the attentional manipulations was found, we statistically tested with the city-block distance (cbd) method (Schröger, Rauh, & Schubö, 1993) whether the MMN amplitudes (deviant-minus-standard difference amplitudes) elicited in the AN and ANT conditions, in which the tones were actively ignored versus task-relevant, were significantly *similar* to each other. The MMN amplitudes elicited in the AN and ANT conditions were significantly similar (the city-block distance [cbd] was 5.7328, $p < .04$), whereas the comparison between the AN and AV conditions did not reach significance (cbd = 6.1959). Furthermore, no latency differences were found between the MMN responses elicited in the three con-

ditions [$F(2,22) = 1.032$]. Thus, we can conclude that the present attentional manipulations had no effect on the MMN response elicited by irregular conjunctions between frequency and spatial location.

Performance in the Noise Change Detection Task

The average reaction time (RT) in the NT (mean: 801.7 msec) condition was shorter than that in the AN (872.5 msec) and ANT (859.2 msec) conditions [one-way ANOVA: $F(2,22) = 8.86, p < .01$]. Post hoc comparisons revealed significant differences between NT versus AN and NT vs. ANT ($p < .01$ and $.05$, respectively). Hit rate (HR) was higher in the NT (0.77) than in the AN (0.68) and ANT (0.68) conditions [$F(2,22) = 9.23, p < .01$]. Again, post hoc comparisons confirmed the differences between NT versus AN, and NT versus ANT ($p < .01$, both). The number of false alarms (FA) were reasonably low compared with the number of hits, higher in the AN and ANT than in the NT condition (FA/HR were 0.19, 0.27, and 0.31 in the NT, AN, and ANT

Table 1. Grand-Average MMN Peak Latencies and Amplitudes in Experiment 1

	AN	AV	ANT
Latency	206.6 (7.0)	193.3 (7.1)	206.0 (7.2)
F3	-0.44 (0.14)	-0.63 (0.13)	-0.24 (0.18)
Fz	-0.59 (0.12)	-0.69 (0.16)	-0.24 (0.28)
F4	-0.57 (0.11)	-0.66 (0.17)	-0.30 (0.18)

Grand-average ($n = 12$) mean (196–220 msec from stimulus onset) deviant *minus* standard (MMN) amplitudes (in μV) measured from the three frontal scalp locations and peak latencies (in msec) measured from the 150–250 msec interval at Fz in the AN, AV, and ANT conditions. Standard error of mean (SEM) values are given in parentheses.

conditions, respectively—FA rates could not be calculated for the noise detection task because there were no foil trials).

We checked whether the lower performance found in the AN and ANT compared with the NT condition was caused by the test tones masking the noise intensity changes. Table 2 shows the RT and HR results in the AN and ANT conditions as a function of the noise change to tone interval. Results show that in both conditions, the length of the interval between the noise change and the closest tone affected both RT and HR values in a manner compatible with the notion of auditory masking: RT was longer and HR lower when the interval between the noise change and the tone was <100 msec than when this interval was >200 msec. One-way ANOVA of the RTs as a function of the interval between noise change and tone (resolution as in Table 2) showed a significant effect [$F(7,77) = 8.26$ and 6.96 , $p < .001$, both for the AN and ANT conditions, respectively]. Planned comparisons between the <100 msec and >200 msec noise-change-to-tone intervals revealed that the effect was related to the temporal distance between the noise change and the closest test tone [$F(1,11) = 14.90$ and 12.33 , $p < .01$, both]. The same analysis of HRs displayed similar results [$F(7,77) = 2.79$ and 2.19 , $p < .01$ and $.05$ for the AN and ANT conditions, respectively; in planned comparisons $F(1,11) = 16.47$ and 4.99 , $p < .01$ and $.05$, respectively]. Outside the masking interval (>200 msec separation between the noise change and the closest test tone), there were no significant differences in RT [$F(2,22) = 0.98$] or HR [$F(2,22) = 0.78$] between NT and the other two conditions. This indicates that the significant main effect of condition on the noise change detection performance was due to masking caused by the test tones and not to redirection of attention.

Answers to the Questions Asked in the ANT Condition

The ratio of correct responses, except for Question 3 (which tested whether subjects remembered the frequency range of the tones in the preceding stimulus block), was not significantly different from chance (cor-

rect/incorrect: 6:6, 4:8, 7:5, 6:6, for Questions 1, 2, 4, and 5, respectively). The answers to Question 3 showed the presence of memory for the frequencies presented during the stimulus block (10:2, $p < .02$). The number of participants who correctly recognized tones presented in the first part of the stimulus block (the “presequence,” see Methods and the left side of Panel A of Figure 1; the related question is Question 4) was approximately equal to that of those correctly rejecting the “illusory conjunction” probe tone (Question 1). It should be noted that (a) the main purpose of asking these questions was to encourage participants to pay some attention to the tones and (b) technical limitations did not allow asking questions about location-related information. Therefore, with the exception of Question 1, the answers obtained are not relevant to the goal of this study (they are reported for the sake of completeness).

Discussion

Infrequent combinations of tone frequency and spatial location elicited the MMN ERP component in all three conditions. Despite the fact that the test tones were to be actively ignored in one condition and were task-relevant in another, MMNs of significantly similar amplitude were elicited by the infrequent feature combinations. These results demonstrate that the auditory system represented the frequent combinations of frequency and spatial location when participants performed a demanding auditory task (attending sounds presented in parallel with the test tones in the AN condition), as well as when the test tones were task-relevant (ANT condition) or when the primary task involved only visual stimuli (AV condition). It is also important to note that no indication of illusory conjunctions was observed in the answers to the questions following the stimulus blocks of the ANT condition. Therefore, the current results provide strong evidence that the formation of auditory feature conjunctions is independent of the direction of focused attention.

To maintain the attentive feature integration hypothesis in the face of the current MMN results, one should

Table 2. Grand-Average Performance Measures in Experiment 1

	–300 to 200	–200 to 100	–100 to 0	Overlap	0 to 100	100 to 200	200 to 300	300 to 400
AN								
RT	794.73 (35.72)	816.48 (53.07)	852.92 (41.72)	963.90 (49.25)	1038.35 (57.50)	901.29 (49.23)	843.26 (47.18)	806.28 (33.13)
HR	70.5 (4.6)	67.3 (5.7)	60.4 (2.2)	58.3 (3.3)	65.5 (5.1)	68.3 (4.7)	75.1 (5.4)	73.2 (4.1)
ANT								
RT	857.04 (39.25)	779.93 (41.11)	910.01 (57.32)	946.26 (27.46)	988.88 (39.00)	875.24 (39.12)	800.92 (35.20)	790.49 (48.90)
HR	75.2 (3.4)	65.2 (3.4)	61.7 (4.5)	64.6 (4.5)	67.5 (3.2)	65.6 (3.7)	73.4 (3.2)	72.1 (5.5)

Grand-average ($n = 12$) RT (in msec) and HR (in %) in the AN and ANT conditions presented as a function of the interval separating the noise change from the closest (in <0 ranges, preceding; in the >0 ranges, following) tone. The central column (“Overlap”) gives the RT and HR values for those noise change targets, which occurred during the presentation of a tone. SEM values are given in parentheses.

assume that participants attended the tones in all three conditions. The deterioration of the performance in the noise change detection task from the NT to the AN and ANT conditions seemingly supports this alternative. Decrease in the level of performance in detecting noise changes could mean that participants divided their attention between the noise and the test tones both in the AN and ANT conditions. However, once the masking effect of the test tones was removed, no significant condition effect was observed. Therefore, the present results do not lend support to the assumption that tones received approximately equal attention in all three experimental conditions.

The present results also confirm the dominant view concerning the utility of measuring the MMN potential in the passive condition for probing the outcome of preattentive sound analysis (Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995). The validity of this measure has often been challenged because of the possibility that attention would “leak” to the “to-be-ignored” channel. However, the MMNs measured in the passive AV condition, which were comparable with those obtained in previous similar studies (Takegata, Huotilainen, et al., 2001; Takegata, Paavilainen, et al., 1999, 2001; Sussman, Gomes, et al., 1998; Gomes et al., 1997), did not differ from the MMNs obtained when participants’ attention was strictly controlled (in the AN condition). Thus, our results substantiate the notion that MMN measured in the passive condition can be used to probe the outcome of preattentive processing (for corroborative evidence, see Sussman, Winkler, & Wang, 2003).

In summary, our results provide strong evidence that auditory features are conjoined irrespective of the direction of focused attention. Preattentive parallel processing of auditory feature conjunctions has two important advantages over attention-driven serial feature integration, it is faster and preserves limited capacities. Fast parallel processing of auditory information is needed in natural situations in which the auditory system must simultaneously keep track of several active sources. Maintaining representations of several sound sources in parallel allows instant switching from one to another as well as facilitating the ability to detect potentially important acoustic signals that fall outside the focus of attention.

EXPERIMENT 2: MMN AND VISUAL FEATURE CONJUNCTIONS

Mismatch Negativity in the Visual Modality

The aim of Experiment 2 was to investigate the possibility of preattentive formation of visual feature conjunctions with ERPs. We tested whether infrequent (deviant) conjunctions of two visual stimulus features elicit the visual mismatch negativity (vMMN) ERP component (Berti & Schröger, 2004; Stagg, Hindley, Tales, &

Butler, 2004; Heslenfeld, 2003; Czigler, Balázs, & Winkler, 2002; Tales, Newton, Troscianko, & Butler, 1999; Alho, Woods, Algazi, & Näätänen, 1992; Czigler & Csibra, 1992; Woods, Alho, & Algazi, 1992; for a recent review, see Pazo-Alvarez, Cadaveira, & Amenedo, 2003). The vMMN reflects a process detecting when the incoming visual stimulus mismatches the memory representation of the regular aspects of the preceding stimulus sequence (Czigler, Balázs, et al., 2002). Because the elicitation of this component does not require participants to actively detect deviant stimuli or, in general, to perform any task related to them, this method can be used to test whether visual stimulus features are conjoined for stimuli appearing in nonsegregated (background) areas of the visual field, outside the focus of attention. vMMN has been found to be insensitive to the difficulty of the concurrent primary visual task (Heslenfeld, 2003). Results of the abovementioned vMMN studies are compatible with the notion that preattentive mechanisms play an important role in implicit memory registration of irrelevant objects (e.g., DeSchepper & Treisman, 1996). The visual MMN can be regarded as the visual analogue of the auditory MMN component. The auditory MMN has been shown to be elicited by infrequent tones that differed from the preceding tone sequence only in their combination of two auditory features (see details in Experiment 1). If vMMN were also elicited by irregular (deviant) feature conjunctions, this result would support the view that also visual stimulus features are conjoined even for unattended stimuli.

Procedure

The Test Grating Sequences

Four types of grating patterns were presented in the test sequences in randomized order. Forty-five percent of the gratings had one of four possible combinations of two grating orientations \times two colors. Another 45% of the gratings had the opposite combination of orientation and color (i.e., different grating orientation and different color). These two types of grating were the “standards” in these “oddball” sequences. The remaining 10% of the gratings (deviants) had, with equal probability, the other two possible combinations of grating orientation and color. The paradigm was similar to the corresponding part of Experiment 1, which is illustrated on the right side of Panel A of Figure 1, the “Main sequence.” Replace loudspeakers with grating direction (the y -axis) and frequency with color (shading of the rectangles). Timing, however, was different (see Methods).

Experimental Conditions

In separate task conditions, two levels of attentional demand were administered. In the “Attend Fixation

Cross" (AFC) condition, participants were instructed to detect infrequent stimulus changes of the fixation cross presented in the center of the visual field. A white cross was displayed in the middle of a central stripe throughout the stimulus blocks. From time to time, the cross became wider or longer. In the AFC condition, these unpredictable changes required speeded reaction. The participants were also instructed to ignore the grating patterns appearing around the target stimulus. In similar task conditions, changes in the characteristics of the background stimuli (in our case, the grating patterns) have been shown to go unnoticed (Mack & Rock, 1998; Rensink, O'Regan, & Clark, 1997). In the "Attend Grating Patterns" (AGP) condition, one of the infrequent color/grating-direction combinations was designated as the target. Participants were requested to give speeded reaction to the emergence of this particular color/orientation conjunction.

Predictions

Elicitation of similar vMMN responses in the two conditions would support the notion that the formation of visual feature conjunctions does not require focused attention. In contrast, if attention was needed for the formation of feature conjunctions, no MMN or an MMN of lower amplitude should be elicited in the AFC condition (compared with that in the AGP condition). In the AGP condition, gratings with one or both target features were also expected to elicit the well-known attention-related visual ERP components, the anterior and posterior selection negativity, and the late positive components (see

e.g., Kenemans, Kok, & Smulders, 1993; Czigler & Csibra, 1992; Wyers, Mulder, Okita, & Mulder, 1989).

Results

The Attend Fixation Cross Condition

Figure 3 (left side) shows the ERP responses elicited by the standard and deviant stimuli. Over anterior brain areas (Fz, Cz), a positive peak (P1, ca. 100 msec peak latency from stimulus onset) was followed by a negative component (N1; 148 msec peak latency). At occipital scalp locations, the negative wave peaking at 98 msec from stimulus onset was identified as the pattern-specific component termed CII (Jeffreys & Axford, 1972). This negative ERP component was followed by a broad positive wave at Oz, whereas at the lateral occipital locations, two somewhat narrower positive peaks followed the CII. The right side of Figure 3 shows the deviant *minus* standard difference waveforms. The difference waveforms showed a negative wave peaking at 128 msec (at Oz) at the posterior electrode locations. This was followed by a positive wave, which peaked at 188 msec. The difference waveforms obtained from the anterior electrode locations showed a small positive difference wave in the 100–140 msec range followed by a similarly low-amplitude negative difference wave in the 150–200 msec interval.

Figure 4 shows the mean amplitudes for the occipital deviant–standard difference waveforms in the 108–148 msec (occipital negativity) and 168–208 msec (occipital positivity) latency ranges. The ANOVA for the occipital negativity [factors: Stimulus type (standard vs.

Figure 3. Experiment 2: Grand-average ($n = 20$) ERP responses to standard and deviant stimuli (left side) and the respective difference waveforms (right side) in the Attend Fixation Cross condition.

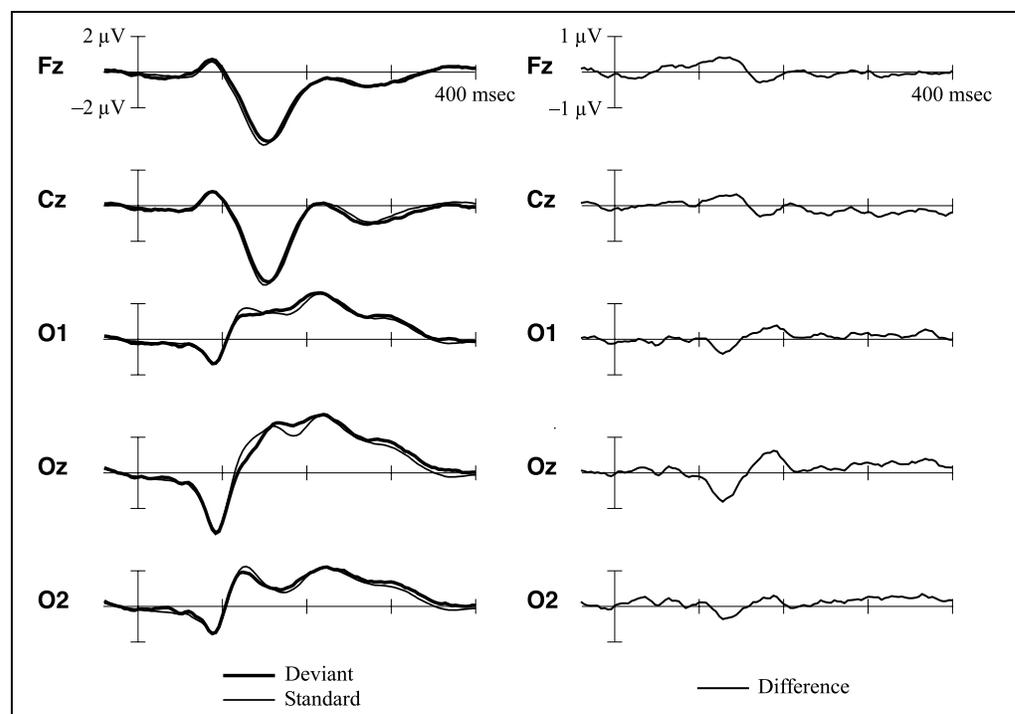
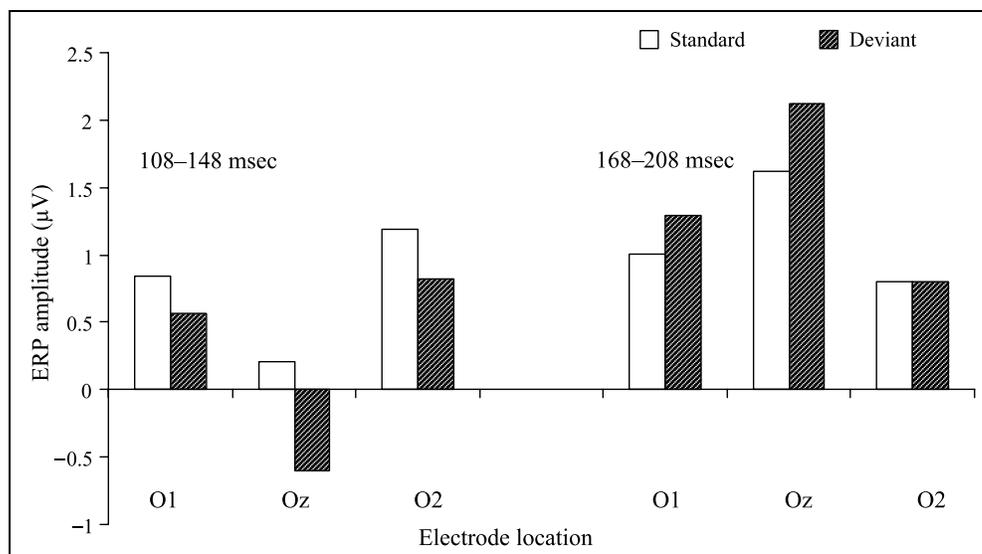


Figure 4. Experiment 2: Mean standard- and deviant-stimulus ERP amplitudes at three occipital electrode locations in the 108–148 msec (left side) and 168–208 msec (right side) latency range in the AFC condition.



deviant) \times Electrode location (O1 vs. Oz vs. O2)] showed that both of the main effects as well as their interaction were significant [$F(1,19) = 7.31, p < .05$ and $F(2,38) = 9.82, \epsilon = .933, p < .001$, for the stimulus type and electrode location factors, respectively; $F(2,38) = 10.38, p < .001$, for the interaction of the two factors]. Post hoc analysis showed that the deviant and standard responses significantly differed at all three electrode locations. The interaction resulted from the higher difference amplitude at Oz than at either one of the lateral electrodes locations (see Figure 4). The similar ANOVA test of the occipital positivity showed a significant effect of electrode location [$F(2,38) = 9.52, \epsilon = .986, p < .001$] and a significant interaction between stimulus type and electrode location [$F(2,38) = 4.41, p < .05$]. Post hoc analysis showed that the deviant and standard responses significantly differed only at Oz.

In contrast to the significant occipital effects of stimulus deviance, no significant difference was found between the ERPs to standard and deviant stimuli at the frontal and central (Fz and Cz) electrode locations. In the two-way ANOVA [factors: Stimulus type (standard vs. deviant) \times Electrode location (Fz vs. Cz)] of the mean amplitudes in the 108–148 msec latency range, only the electrode location main effect was significant [$F(1,18) = 14.59, p < .01$].

The detection of the fixation cross changes was fairly accurate. The mean HR was 96.3%, false alarm rate 3.3%. The mean reaction time (RT) was 455 msec.

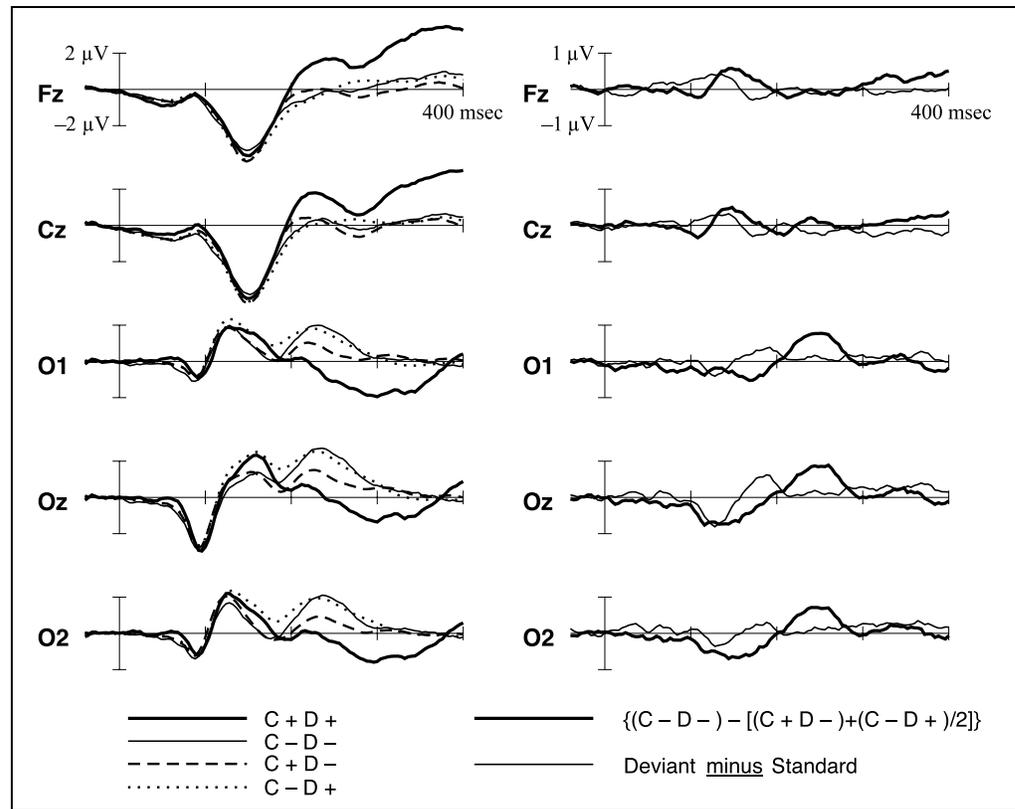
The Attend Grating Patterns Condition

Identification of the target grating was rather difficult: The average HR was only 67%, the false alarm rate 2%. On the basis of median HR, the group was divided into subgroups, one with higher and the other with lower performance rate. The average RT in the higher-

performance subgroup was faster than that in the lower-performance subgroup [537 vs. 594 msec, $t(18) = 2.56, p < .05$]. The ERP results, however, showed no significant difference between the two subgroups (see below).

Figure 5 (left side) shows the ERP responses elicited by the four stimulus classes (C+D+, C+D-, C-D+, and C-D-; where C stands for color, D for grating direction, the “+” sign denotes those stimuli, which had the target feature level in the given stimulus dimension, whereas the “-” sign denotes those stimuli, which differed from the designated target in the given stimulus dimension). In the 0–150 msec range, the responses were similar to each other and also to those obtained in the AFC condition. In later latency ranges, for the posterior electrode locations, attention-related components dominated the ERP responses, especially for the (C+D+) and (C+D-) stimuli. The right side of Figure 5 compares the deviant *minus* standard difference waves between the AFC and AGP conditions. For the AGP condition, only the nontarget deviant (C-D-) *minus* standard $\{[(C+D-)+(C-D+)]/2\}$ difference potentials were calculated, because the ERP response to C+D+ deviants is dominated by target-related ERP components. In the 100–140 msec poststimulus interval, at the posterior electrode locations, the deviant *minus* standard difference waveforms were quite similar between the two conditions. Therefore, the deviance-related negativity was assessed in the same 108–148 msec interval for the AGP condition, just as it was for the AFC condition. The ANOVA test [factors: Subgroup (above-median performance vs. below-median performance) \times Stimulus type (nontarget deviant vs. standard) \times Electrode location (O1 vs. Oz vs. O2)] showed significant effect of location [$F(2,34) = 10.88, \epsilon = .920, p < .001$] and a significant Location \times Stimulus interaction [$F(2,34) = 6.57, p < .01$]. The post hoc test showed that the ERP elicited by the nontarget deviant was

Figure 5. Experiment 2: Left side: Grand-average ($n = 19$) ERP responses elicited by the target C+D+, and three nontarget grating patterns C-D-, C+D-, and C-D+. Right side: Deviant *minus* standard difference waveforms for the AFC (thin line) and AGP conditions (thick line). In the AGP condition, deviant *minus* standard differences were calculated only for the nontarget deviant (C-D-) by subtracting from the nontarget deviant ERP the average of the ERPs elicited by the two frequent grating patterns $\{(C-D-) - [(C+D-) + (C-D+)/2]$.



significantly more negative than standard-stimulus response at Oz and O2, but not at O1. No significant differences were found between the two subgroups separated according to their rate of performance.

Target-related ERP responses were assessed by two-way ANOVA tests with factors of stimulus type [(C+D+) *minus* (C-D-) vs. (C+D-) *minus* (C-D-) vs. (C-D+) *minus* (C-D-)] and the electrode location (O1 vs. Oz vs. O2 or Fz vs. Cz). Target stimuli (C+D+) elicited larger posterior selection negativity (SN) than stimuli having only the relevant color (C+D-). Stimuli having the relevant direction but irrelevant color (C-D+) elicited no SN. This difference was reflected by the significant stimulus type main effect [$F(2,34) = 9.74, \epsilon = .968, p < .001$] of the ANOVA, and the significant results of the subsequent post hoc tests. SN was larger at Oz and O1 than at O2 [electrode main effect: $F(2,34) = 3.37, \epsilon = .959, p < .05$]. The hemispheric asymmetry was greater for the target than for the nontarget stimuli [$F(4,68) = 2.73, p < .05$ for the interaction between the stimulus type and electrode location factors]. The similar ANOVA of the anterior selection positivity (SP; the measurement electrodes were Fz and Cz) showed a stimulus type main effect [$F(2,34) = 13.58, p < .001$]. Results of the post hoc test indicated that this component emerged only for the target (C+D+) stimulus. The target stimulus also elicited the late positivity [stimulus type main effects: $F(2,34) = 10.81, \epsilon = .62, p < .001$ and $F(2,34) =$

$28.74, \epsilon = .52, p < .0001$ for the early (P3a) and late (P3b) part of this positivity, respectively].

Discussion

Infrequent combinations of color and grating pattern orientation elicited a posterior negative wave in the 120–160 msec latency range, which was followed by a posterior positivity. The negative component was identified as the vMMN described by Heslenfeld (2003) and Czigler, Balázs, et al. (2002). The elicitation of vMMN responses in the present study indicates that the visual system detected a difference between the rare (deviant) and the frequent (standard) grating patterns. Because our standard and deviant grating patterns differed only in the combination of the two test features, but not in any single feature, the current results suggest that memory representations were formed for the frequent feature conjunctions appearing in the stimulus sequences and the infrequent stimuli were found to mismatch these memory representations. One may argue that the emergence of the negativity could have been due to the refractoriness of the neuron population selectively tuned to the characteristics of the frequent stimuli (Kenemans, Jonh, & Verbaten, 2003). However, Czigler, Balázs, et al. ruled out this possibility. These authors showed that the elicitation of the vMMN component cannot be explained on the basis of differential refractoriness of stimulus-specific neuronal populations and,

therefore, one must assume that, similarly to the auditory modality, elicitation of the vMMN involves some memory representation of the regularities of the stimulus sequences. The current results lend further support to this interpretation in that the elicitation of vMMN in our paradigm cannot be attributed to neurons sensitive to individual visual features. It is not likely that there exist sufficiently large populations of neurons sensitive to a given combination of two features that differential refractoriness of such populations would produce a response detectable with scalp-recorded EEG.

Most importantly, the vMMN response was elicited with both of the attentional manipulations employed in the current study. This indicates that memory records representing visual feature conjunctions are formed independently of the direction of focused attention. In the AFC condition, the grating patterns were irrelevant for the participant's task, whereas in the AGP condition, grating patterns were to be attended and one of them was the target stimulus to which participants were required to respond. Performance levels in the conjunction discrimination task (AGP condition) showed that it was difficult to identify the target under the stimulus conditions of the present experiment, probably because both the stimulus durations and the interstimulus interval were short. However, the attention-related components (selection negativity, selection positivity, and late positivity) appeared to be similar to those recorded in other conjunction discrimination tasks (e.g., Kenemans, Kok, & Smulders, 1993; Czigler & Csibra, 1992). The emergence of the attention-related ERP components shows that a considerable number of correct responses were based on proper discrimination of the stimuli in participants achieving higher as well as those achieving lower performance levels.

Despite the low discriminability of the grating patterns (see behavioral responses in the AGP condition), deviant nontarget feature combinations elicited the vMMN in the AFC condition (in which the grating patterns were not attended) and this vMMN response was similar in amplitude and latency to the one obtained in the AGP condition (in which the grating patterns were attended). Further, the lack of significant differences between the vMMNs elicited by high- and low-performance participants (in the conjunction discrimination task) suggest that the formation of memory representations for feature conjunctions and the detection of deviation from these representations are independent of task performance. On the basis of these results we conclude that under the current stimulus conditions, both the formation of memory traces for feature conjunctions and the detection of irregular feature conjunctions occurred without the need for focused attention.

In the difference waveforms obtained in Experiment 2, the vMMN component was followed by a posterior positivity. In the auditory modality, emergence of a similar positive component is a frequent phenomenon,

and this component is usually identified as the auditory P3a. In the present study, the posterior distribution of the positive component argues against this interpretation. Therefore, further research is needed to clarify the relationship between the vMMN and the subsequent positive wave.

EXPERIMENT 3: MCCOLLOUGH EFFECT AND IMPLICIT MEMORY

Perceptual Aftereffect of Color/Orientation Conjunction

In Experiment 2, test stimuli were grating patterns with color and orientation as the critical features. Adaptation to colored gratings may give rise to a color-contingent aftereffect (McCullough, 1965). As an example, following simultaneous adaptation to a horizontal red/black and a vertical green/black grating, the white bars of a horizontal black/white pattern are perceived in the opponent color of red, whereas the white bars of a black/white vertical grating are perceived in the opponent color of green. The effect is long lasting, specific to the retinal area that was stimulated by the adaptation pattern (Stromeyer, 1978), and does not transfer to the non-stimulated eye (Murch, 1972). Although the latter characteristics suggest a low-level locus of the effect, fMRI results show the involvement of the fusiform and lingual gyri, that is, extrastriate structures (Humphrey, James, Gati, Menon, & Goodale, 1999). As Houck and Hoffman (1986, Experiment 1) pointed out, the McCollough effect is also elicited outside the focus of spatial attention. However, at the same time, the RT obtained in a visual search task (Experiments 1 and 2 of Houck and Hoffman) increased as a function of the number of distracter gratings, which showed that the search for a particular color/orientation conjunction requires spatial attention. On the basis of these results, Houck and Hoffman argued that the conjunction information underlying the McCollough effect is inaccessible by cognitive mechanisms. However, it is equally possible that, as in the case of various types of "fleeting memories" (Coltheart, 1999), accessibility is dependent on processing capacity (e.g., Chun & Potter, 1995). Because the rationale and stimuli of the present study were similar to that of Houck and Hoffman, it was necessary to test the emergence of McCollough effect in our design. Therefore, in addition to testing whether infrequent color/orientation conjunctions were preattentively detected, we also investigated the participants' explicit memory of the grating patterns.

Procedure

To investigate the emergence of the McCollough aftereffect, procedures of the AFC condition of Experiment 2

were repeated, except that brain electric activity was not measured. Some of the stimulus blocks were followed by the presentation of vertical and horizontal black/white gratings. Participants were instructed to identify the hue/saturation of the bright bars of the gratings by selecting the closest equivalent from a set of colored squares. At the end of the experiment, an incidental recognition test was administered. As the first step of this test, participants were asked to select patterns that they recognized as background stimuli of the change detection task. In the second step, participants were told to decide whether the selected patterns appeared frequently or infrequently during the change detection task.

Results

Change Detection

Performance in detecting changes of the fixation cross was similarly high as that seen in the AFC condition of the main experiment: The average HR was 98.8%, the false alarm rate 0.8%. The main RT was 414 msec.

McCollough Effect

Results on the emergence of McCollough effect were clear. The average rating on a ± 4 scale was $+0.17$ ($SD = .24$), which did not significantly differ from zero [$t(9) = 1.32$]. Out of the 60 judgments, 40 were zero themselves (i.e., no color aftereffect at all). Thus, it appears that the current stimulus conditions did not evoke a measurable perceptual color-contingent aftereffect.

Stimulus Recognition

Only one of the 10 participants selected gratings of nonpresented color (1 green and 1 yellow, both as “infrequently presented”). All participants selected the red/vertical grating as one that they remembered having seen during the experimental session, whereas the other three gratings that appeared during the experiment were selected by 9 participants. Thus, participants remembered very well which gratings appeared in the background, although they focused on the fixation cross. However, when having to decide whether the selected gratings appeared frequently or infrequently in the stimulus sequences, participants marked only 6 of the altogether 40 selected gratings (10 subjects \times 4 selections) as infrequent ($\chi^2 = 19.00$, $df = 1$, $p < .0001$). Further, even from these 6 gratings marked as “infrequent,” 3 gratings actually appeared frequently in the stimulus sequences. This result shows that participants did not remember which color/direction conjunctions appeared frequently and which appeared infrequently in the task sequences.

Discussion

No color-contingent (McCollough) perceptual aftereffect was obtained under the stimulus and task conditions identical to the AFC condition of Experiment 2. Being a long-lasting effect (Stromeyer, 1978), it is improbable that such effect was working throughout the stimulus blocks but dissipated by the time of the test. Accordingly, on the basis of the present results, the possibility that the McCollough effect and the elicitation of vMMN by infrequent feature conjunctions have a common origin can be ruled out.

Participants did not remember the probabilities with which the four color/orientation conjunctions appeared in the stimulus sequences. Thus, it is reasonable to assume that during the AFC condition of Experiment 2, which was identical both in stimulation and in the task to Experiment 3, participants did not notice the probabilities with which the four grating stimuli appeared within the stimulus sequences. This assumption receives further support from the low performance levels obtained in the AGP condition of Experiment 2, which suggest that it was difficult to discriminate the four grating patterns under the stimulus conditions used in Experiments 2 and 3. Even so, the elicitation of vMMN in Experiment 2 demonstrated that memory traces were formed that represented the frequent feature combinations. Similar dissociations have been found in the auditory modality between voluntary (top-down) and implicit (stimulus-driven) access to the memory representations involved in the MMN-generating process. An incidental recognition test showed that participants do not remember the rate with which tones occurred in a task-irrelevant sequence (Winkler, Sussman, et al., 2003), whereas the present Experiment 1 showed that participants remembered the pitch of the tones from a sequence they just heard. Furthermore, with a complex sequential auditory regularity (the rule of the tone sequence was: the higher the frequency the higher or, in separate stimulus blocks, lower the intensity), Paavilainen, Simola, Jaramillo, Näätänen, and Winkler (2001) found that, whereas most participants could not tell what was regular in the tone sequence, in a passive situation, deviants violating this regularity elicited the MMN response. The findings of the present study demonstrated the existence of implicit memory records storing conjunctions of visual features. These memory representations may be involved in storing characteristics of the visual background as well as in the processing of reappearing objects, the latter also reflected by results of negative priming experiments (DeSchepper & Treisman, 1996).

GENERAL DISCUSSION

The results of the current experiments demonstrated the existence of memory traces encoding the conjunc-

tion of two features under different task conditions (task-relevant vs. task-irrelevant) in the auditory as well as in the visual modality. Memory traces containing feature-conjoined stimulus information were indexed by the auditory and visual MMN ERPs, which are elicited when a stimulus deviates from the regularities detected from the preceding stimulus sequence. In the current experiments, deviant stimuli differed from the regular ones (standard stimuli) only in the specific combination of two stimulus features. That is, in the current stimulus sequences, deviants could elicit MMN only if (1) the memory representation of the regularities encoded those combinations of stimulus features that occurred regularly within the stimulus sequences; (2) the combination of stimulus features was also evaluated for the deviant stimuli; and (3) difference between the deviant and standard feature combinations was detected. Thus, the elicitation of the MMN responses in the current stimulus paradigms demonstrated that features were conjoined for all (standard as well as deviant) stimuli within the test sequences.

Three different attentional conditions were set up in the auditory experiment (one in which the tone sequences were partially attended and two in which they were ignored; the latter two differed in the stimulus modality and attentional demand of primary task) and two different attentional conditions in the visual experiment (in one the grating sequences were task-relevant, in the other they were to be ignored). No differences were found in the elicitation and component parameters of MMNs as a function of the attention in either modality. In fact, statistically significant similarity of the MMN amplitudes has been found between the two extreme attentional conditions of the auditory experiment. These results support our conclusion that in the current stimulus paradigm, features of the stimuli were conjoined irrespective of the direction of focused attention.

Our conclusion appears to contrast the main suggestion of the feature integration theory, which claims that conjoining stimulus features requires focal (spatial) attention. Therefore, in the following we examine the relationship between the current results and those that led to the formation of the feature integration theory. The feature integration theory is primarily based on results obtained in visual search tasks (for a review, see Wolfe, 1994). In visual search tasks, a relatively rich array of stimulus elements is usually presented simultaneously and subjects are required to find the location of a designated target or targets. The involvement of attentive processes in finding feature-conjunction targets (i.e., targets defined by the co-occurrence of two or more features) has been inferred from results showing an approximately linear relationship between the reaction times of detecting targets and the number of distracter (nontarget) items appearing together with the target(s). The stimulus paradigms used in the current experiments are markedly different from the basic

stimulus configuration used in visual search studies. In the current paradigms, only two objects were presented simultaneously (i.e., the continuous noise and a tone sequence in the auditory experiment and the fixation cross and a sequence of grating patterns in the visual experiments). Therefore, although in our task-irrelevant conditions the test sequences fell outside the focus of the participant's attention (serving as background to the task-relevant stimuli), one can argue that when only one object falls outside the focus of attention, its features may be conjoined "by default," that is, without the need for focused attention. Indeed, Treisman (1998) suggests: "Binding failures typically occur with high load displays when several objects must be processed under high time pressure. When there is only one unattended object, its features must belong together, so there should be no problem determining what goes with what" (p. 1305). This would explain the contrast between results obtained in paradigms presenting several stimuli concurrently (e.g., Treisman & Gelade, 1980 in the visual and Hall et al., 2000 in the auditory modality) and those obtained when only one stimulus was unattended at a time (see Experiment 2 for a visual example, the present Experiment 1, and Woods, Alain, et al., 1998, for auditory results).

However, if attention is only required for conjoining features under special circumstances (high load displays processed under high time pressure), then the mechanism of feature binding should "normally" work preattentively. Although our everyday environment is considerably richer in sensory events than the stimulation provided by most laboratory experiments, we do not often have to operate under high time pressure. It has also been suggested that once an object representation has been formed (this includes binding the features of this object), the formation of representation for subsequent similar stimulus instances does not require focused attention (Treisman, 1993; see, however, Wolfe, 1999). Finally, if we consider that segregating objects by certain cues (such as texture-based segregation in vision and pitch-based segregation in audition) is also assumed to be fully preattentive, the impression is that attentive feature binding mainly occurs in bodyguards trying to locate an assassin in the crowd before he can hit the target. This argument can be restated in a theoretical manner. Because we know that there are situations in which stimulus features are conjoined without the requirement of focused attention, there must exist preattentive processes of feature binding. What is then the role of attention in feature binding? Perhaps the attention effects attributed to feature binding are related to the specific tasks in which they were observed rather than generally to the process of feature binding. Some modern views on attention are compatible with this suggestion (Pashler, 1998).

Another difference between the present experiments and those that form the evidence base of the feature

integration theory is that in our study, the encoding of feature conjunctions in memory representations were inferred from the elicitation of the MMN an ERP response, the latter studies analyzed performance measures and/or assessed the participants' experiences. One should therefore discuss the relationship between the memory traces indexed by MMN and the memory traces underlying performance measures and the participants' experiences (e.g., of illusory conjunctions).

In the auditory modality, several studies demonstrated that the memory traces involved in MMN generation and behavioral indexes of perception correspond to each other (MMN parameters compared with performance measures as well as with individual perceptual abilities; for a review see, Näätänen & Winkler, 1999). However, as was already noted in the discussion of Experiment 3, in some cases, performance-based measures of memory were dissociated from the MMN measure of memory (Winkler, Sussman, et al., 2003; Paavilainen et al., 2001), although, because in these studies, behavioral and MMN measurements were separated in time, one may argue that access to the memory traces have been lost during the time between the two measures (e.g., due to interference). Some studies found dissociation between perception and preattentive change detection reflected by MMN. Sussman, Winkler, Kreuzer, et al. (2002) found that two successive deviant tones delivered within 200 msec (the assumed temporal window of integration) elicited only a single MMN response even though subjects indicated (on-line) that they heard two separate deviant tones. Furthermore, Berti, Schröger, Cowan, and Winkler (2000) found no significant MMN for deviants separated from the preceding standard by a long (>10 sec) silent interval, although participants could discriminate the deviant tone from the standard tones presented in the preceding short train. However, both of these effects can also be explained by characteristics of the MMN-generating process without assuming that the corresponding information was not encoded in the memory traces that underlie the MMN-generating process.

Much less is known yet about the memory traces indexed by the visual MMN. However, a considerable amount of research shows the existence of "fleeting memories" in vision (Coltheart, 1999). Implicit memory traces, including ones with semantic properties, may be established for stimuli that appear outside the focus of attention, or for which the temporal constraints of the situation prevent the formation of a retrievable memory trace. Negative priming (DeSchepper & Treisman, 1996) and attentional blink studies (Vogel et al., 1998) showed that such memory traces may still influence the processing of other stimuli even when participants cannot retrieve these memory traces. It is possible that the memory representations involved in MMN generation are "fleeting" implicit memories. Implicit memories can be formed preattentively but attention may be needed for their consolidation (Coltheart, 1980) and/or

to transfer these implicit memories to further stages of information processing (Potter, Stiefbold, & Moryadas, 1998). Thus, according to this view, stimulus features are conjoined preattentively, but preserving the conjoined representations may require attentive processing.

In summary, the current as well as previous studies showed that, with compatible stimulus paradigms, comparable results are obtained in the auditory and visual modalities with respect to the role of attention in binding stimulus features. Although the current results demonstrated the existence of preattentive feature-conjunction processes, conjoining features within rich arrays of objects under time pressure and/or long-term storage of the feature-conjoined memory representations may require attentive processes.

METHODS

Experiment 1

Participants

Fourteen healthy volunteers (5 men; 18 to 31 years of age, average age 20.4 years) with normal hearing (checked with audiometry before the experiment) were paid for their participation. Written informed consent was obtained from all participants after the procedures of the study were explained to them. Two participants' data were rejected due to extensive electrical artifacts.

Stimuli and Procedure

Tone sequences. Tones of 78 dB intensity (SPL, measured at the participant's head) and 100 msec duration (including 5 msec rise and 5 msec fall times) were presented with a constant, 800-msec long onset-to-onset interval via two loudspeakers positioned behind the participant (Figure 1, Panel B, left side). Both loudspeakers were positioned on an arc of 1-m radius centered on the participant behind the participant's head.

Five equidistant loudspeaker location pairs (160° apart) were used in the experiment. On each side of the participant's head, neighboring loudspeaker positions were separated by 5°. The leftmost location on the left side of the subject was paired with the leftmost location on the right, the remaining four pairs constructed by progressing from left to right in parallel on the two sides (Figure 1, Panel B, left side). Carryover of location information from one block to the next was prevented by changing the loudspeaker locations between stimulus blocks (Figure 1, Panel B, right side). Loudspeaker location pairs were selected equiprobably in a pseudorandomized order that excluded using the same loudspeaker location pair in consecutive stimulus blocks. In addition to the two real loudspeakers, four "virtual" sound source locations were created through linear cross-fading of the outputs of the two loud-

speakers, using the four proportionally equidistant points of the cross-fading function. The direction of the four virtual sound sources was perceived as being between the two actual loudspeakers (Figure 1, Panel B, right side). Informal testing showed that the directions of these virtual sound sources could be easily distinguished both from each other and from the two real loudspeakers.

Five sets of tone frequencies were used, only one set in each stimulus block. The base frequencies for these frequency sets were 307.63, 390.96, 545.45, 762.99, and 969.70 Hz. From each of these base frequencies, six frequency values were calculated with proportionally equal, 10% cascading increments starting from the base level. Again, to prevent carryover between stimulus blocks, the frequency set was changed between blocks. Frequency sets were selected equiprobably and independently of the loudspeaker location pairs in a pseudorandomized order that excluded using the same frequency set in consecutive stimulus blocks. In summary, consecutive tone sequences were composed of a different set of six frequencies and were presented from a different pair of loudspeaker locations (which also made the directions of the four virtual sound sources to differ between consecutive sequences).

Tone sequences were made up of two parts that had somewhat different characteristics and presented different tones. The two parts were delivered in a continuous manner (i.e., without breaking the uniform rhythm of tone delivery). Thus, if attention was focused on the tones at the beginning of the tone sequence (even when participants were instructed to ignore them), no indication of the specific tones used for testing the formation of feature conjunctions could be gleaned from the tones appearing in the beginning of the stimulus blocks. This precaution was taken because one version of the attentive feature conjunction theory suggests that, although building an object file requires focused attention, once such a file has been created, further instances of the same stimulus can be identified without need for focused attention (Treisman, 1993).

During the first, approximately 1.5 min of each stimulus block, 112 tones were delivered (the “presequence,” Figure 1, Panel A, left side). This tone sequence was composed of 32 of the 36 (6×6) possible combinations of the six different frequencies of the frequency set and the six different sound source locations (the 2 loudspeakers plus the 4 virtual sound sources created by cross-fading the loudspeaker intensities). The tones were presented in a random order, half of them 3 times, the other half 4 times within the presequence. Four out of the 36 possible frequency–location combinations did not appear within the presequence. These four were the possible combinations between the two real loudspeakers and two frequencies, the second and the fifth frequency values of the frequency set selected for the stimulus block. Note that

although these four frequency–location combinations did not appear within the presequence, on separate tones, both the frequency values and the source locations appeared within the presequence.

In the second part of the stimulus blocks (the “Main sequence,” Figure 1, Panel A, right side), 400 tones were delivered. The main sequence was composed of four types of tones, none of which appeared in the presequence. These four tones covered all possible combinations of the Two loudspeakers \times Two frequencies (the second and the fifth frequency values of the frequency set selected for the stimulus block). Forty-five percent of the tones were delivered by one of the two loudspeakers and had one of the two frequencies. Another 45% of the tones were delivered by the other loudspeaker and had the other frequency. These were the standard tones of the oddball type of main sequence. The remaining 10% of the tones had, with equal probability, the two remaining combinations of the two loudspeakers and the two frequencies (deviant tones). Sequences were separately pseudorandomized with the possibility of delivering two deviants in a row excluded.

The noise task. Continuous band-filtered (100–2100 Hz) white noise was presented from a loudspeaker placed directly in front of the participant at a 1.9-m distance. The participant’s task was to detect slight changes in the intensity (increase and decrease, equiprobably) of the continuously presented noise. The base noise intensity was 61 dB (SPL, measured at the participant’s head), intensity transitions were 5-msec long linear ramps. On average, intensity changes occurred once every 15.25 sec (even distribution between 0.5 and 30 sec). Participants were required to press a response key as fast and accurately as possible when they detected the intensity changes. Responses falling between 200 and 2000 msec from the intensity change onset were considered correct.

Training was given to participants in which the amount of intensity change was determined separately for each participant (between 1 and 3 dB). Noise blocks of 120 sec were administered. Starting from 3 dB, the amount of intensity change was decreased as long as the participant still performed above 80% and increased when detection performance did not reach 80% in two consecutive blocks of the same intensity change. The training period ended when a stable above-80% level of performance was established, typically after 10–12 training blocks. Participants received feedback of their task performance after each stimulus block of the main experiment. They were informed before starting the main experiment that they would receive a bonus payment for correct performance. The amount of bonus gradually increased, starting at 75% base performance. The base level was set to 75% because we anticipated a drop in performance level when the task was to be performed in the presence of the test tones. The bonus could substantially increase the participant’s fee.

Movie. Participants selected a subtitled movie before the test session. The movie was presented to them without sound on a 51-cm TV screen placed directly in front of them at a 2-m distance. The movie was shown during all four conditions.

Questions. The following questions were asked, in random order, one after each of the ANT stimulus blocks: (1) A tone with a combination of frequency and location, which did not appear in the stimulus block was presented to the participant. The participant was then asked whether this sound appeared during the stimulus block. The stimulus block, after which this question was asked, was slightly different from all other blocks. The first part of the sequence used only 26 of the 36 possible frequency–location combinations, leaving 6 combinations, which did not appear at all during the stimulus block, because 4 combinations were used in the second part of the block. One of these six tones was delivered to the participant. (2) All four tones that appeared in the second part of the stimulus block were presented. The participant was asked to tell which of these appeared frequently and which less frequently. (3) A tone was presented, which had one of the standard frequencies of the stimulus block preceding the just finished one (the loudspeaker location remained as it was in the just finished stimulus block). The participant was asked whether this sound appeared in the just finished stimulus block. (4) A tone was presented that appeared during the first part of the stimulus block. The participant was asked whether this sound appeared during the stimulus block. (5) One standard and one deviant tone of the stimulus block were presented. The participant was asked which of them appeared more frequently within the stimulus block.

Order of the conditions (see description at Experiment 1). The AN, AV, and NT conditions were administered in a counterbalanced order. The ANT condition was always presented at the end of the experiment, so as not to draw the participants' attention to the tones during the other conditions. Each condition received five stimulus blocks (20 stimulus blocks, altogether).

Electroencephalogram Recording and Data Analysis

Electroencephalogram (EEG) was recorded with Ag/AgCl electrodes from Fz, Cz, Pz, F3, C3, P3, F4, C4, P4 (10–20 system) and the left and right mastoids (Lm and Rm, respectively). The reference electrode was placed on the tip of nose. Horizontal electrooculogram (EOG) was monitored between electrodes placed lateral to the outer canthi of the two eyes, vertical EOG between an electrode placed above and another below the right eye. EEG and EOG signals were amplified, filtered (40 Hz, low pass), and digitized with 250 Hz sampling rate (NeuroScan Synamps amplifier, Scan software) then filtered off-line (1.5–30 Hz). Epochs of 90 msec prestimulus and 490 msec poststimulus period were averaged

separately for the standard and deviant tones delivered in the second part of the sequences. Epochs with an EEG or EOG change exceeding 100 μ V were excluded from averaging.

For each condition, the ERPs elicited by the different standard and deviant tones were separately collapsed. (Note that sound source location and frequency varied across the stimulus blocks. Therefore, the standard- and deviant-stimulus averages include responses to 10 different tones, each.) The elicitation of MMN was assessed by comparing the responses elicited by standard and deviant tones within the expected MMN time interval (196–220 msec in all current experimental conditions). This window included the range within which the deviant-minus-standard difference was close to maximal in the three EEG conditions, thus maximizing the signal-to-noise ratio while keeping the measurement interval constant across conditions. Statistical testing was conducted with a repeated-measures ANOVA [Condition (AN vs. AV vs. ANT) \times Stimulus (standard vs. deviant) \times Electrode (F3 vs. Fz vs. F4)]. Where applicable, Greenhouse–Geisser correction of the degree of freedom values was calculated. Similarity of the MMN amplitudes (averaged over the three frontal channels—F3, Fz, and F4) between the AN and ANT and between the AN and AV conditions was tested with the city-block distance method (Widman & Schröger, 1999; Schröger, et al., 1993; Edgington, 1980), which estimates the probability that two measurements are taken from the same distribution.

One-way ANOVA tests were used to test HR and RT effects in the primary task. Effects were further specified using planned and Sheffé-type post hoc comparisons.

Significant deviation from the chance level for correct versus incorrect answers to the questions asked following the ANT blocks was tested by comparing these results with the binomial distribution.

Experiment 2

Participants

Twenty adults (11 women and 9 men), aged 18–26 years, served as paid volunteers in the main (ERP) experiment. All participants had normal or corrected-to-normal visual acuity and correct color vision. Due to high artifact rate, one participant's data were discarded from the AGP condition of the experiment. Written informed consent was obtained from the participants after the procedures of the study were explained to them.

Stimuli and Procedure

Red/black and blue/black, horizontal and vertical square-wave gratings (2.3 cycle/deg, measured from the 1-m viewing distance) were presented over the dark background of an SVGA monitor. The pattern covered a field

of 12° horizontally and vertically, except for a central dark stripe (12° horizontally and 0.75° vertically). Gratings were presented for 17 msec with a variable ISI of 350–450 msec (rectangular distribution).

In the AFC condition, a white cross was displayed in the middle of the central dark stripe throughout the stimulus blocks. From time to time, the cross was made wider or longer. These changes occurred unpredictably (mean frequency: 1.33/min), and speeded button-press responses were required for them. Participants were instructed to ignore the grating patterns, whose only function was, according to the instruction, “to produce a more vivid display.” In the AGP condition, the white cross served as the fixation point.

In both conditions, the grating patterns were presented in blocks of 500 stimuli, 12 test blocks preceded by two practice blocks per condition. For half of the participants, the probability of the vertical red (VR) and horizontal blue (HB) gratings was .45 each (“standard stimuli”) and the probability of the horizontal red (HR) and vertical blue (VB) grating were .05 each (“deviant stimuli”). For the other half of the participants, these probabilities were reversed. Each subgroup of participants, separately, was further subdivided into two halves according to the two possible target grating patterns (AGP condition only).

ERP Recording

Electroencephalogram was recorded from Fz, Cz, Pz, T3, T4, T5, T6 O1, Oz, O2 points, all referred to the nose-tip. Fpz served as the ground electrode. Vertical and horizontal EOGs were recorded between above and below the right eye and between two electrodes placed laterally to the left and right outer canthi, respectively. Signals were amplified, filtered (1–30 Hz, 6 dB/octave), and digitized (250 Hz). ERPs (100 msec prestimulus and 650 msec poststimulus period) were averaged separately for the 4 different grating patterns (HR, VR, HB, and VB). Trials with electric changes exceeding 60 μ V on any recording channel were rejected from averaging. ERPs elicited by stimuli following a change of the fixation cross within 600 msec were also discarded (AFC condition, only). In the AGP condition, ERPs were also rejected for error trials (for both omission and commission errors).

Analysis of the Responses Related to Conjunction Deviance

In the AFC condition, responses specific to the infrequent nontarget conjunctions of color and direction (vMMN) were calculated as difference waveforms between the ERPs elicited by the deviant and standard stimuli (collapsed across the different stimuli within subjects).³ For the AGP condition, separate averages

were calculated for the target deviant (color and direction relevant: C+D+), the nontarget deviant (color and direction irrelevant: C–D–) and the two nontarget standard stimuli (color relevant, direction irrelevant: C+D– and color irrelevant, direction relevant: C–D+). Deviance-related ERP responses (the vMMN) were identified from the $(C-D-) - [(C-D+) + (C+D-)]/2$ difference. The target (C+D+) responses were not used in determining the vMMN because the target elicited large task-related ERP components.⁴

Based on the previously established properties of vMMN (Heslenfeld, 2003; Czigler, Balázs, et al., 2002), we identified the negative peak at Oz and the positive peak at Fz in the 80–160 msec poststimulus interval from the grand-average ERP difference waveforms. Additionally, a subsequent occipital positive wave was found in the 160–220 msec interval. For each participant, the mean voltage was calculated in 40-msec wide intervals centered on the latency of the peak in the corresponding grand-average difference waveform, separately for the standard and deviant responses (only nontarget deviants in the AGP condition). Measurements were taken at O1, Oz, and O2 for the posterior deviant-related components, and at Fz and Cz for the anterior component.

Analysis of the Attention-related ERP Responses

In the AGP condition, attention-related ERP components were identified in the $(C+D+) - (C-D-)$, $(C+D-) - (C-D-)$, and $(C-D+) - (C-D-)$ difference waveforms. The posterior selection negativity was measured in the 200–260 msec latency range at O1, Oz, and O2, and the anterior selection positivity in the 220–260 msec range at Fz and Cz. The early (P3a) and late (P3b) parts of the P3 complex were measured at Fz, Cz, and Pz in the 300–400 and 500–600 msec latency ranges, respectively.

Statistical Testing

Measurements in the main experiment were entered into ANOVA tests with factors of stimulus (standard vs. deviant) and electrode location. Where appropriate, degree of freedom values were adjusted with the Greenhouse–Geisser correction procedure (ϵ correction factors given). Tukey HSD tests were used for post hoc comparisons.

Experiment 3

Participants

Ten paid volunteers (6 women and 4 men, aged 21–25 years) participated in Experiment 3. None of them took part in Experiment 1 or 2. All participants had normal or corrected-to-normal visual acuity and correct color vision. Written informed consent was obtained

from the participants after the procedures of the study were explained to them.

Assessment of the McCollough Effect

Eight stimulus blocks, identical to those presented in the AFC condition of Experiment 2, were delivered. For half of the participants, the red/horizontal and blue/vertical gratings served as the standard stimuli, whereas the blue/horizontal and red/vertical gratings served as deviants. In the other half of the participants, the roles of the different grating stimuli were reversed. In order to assess the color-contingent aftereffect, two probe-test screens (one horizontal and one vertical) were presented simultaneously for 700 msec duration after Blocks 1, 3, and 8. The delay of the probe screens from the end of the stimulus block was 1 min. Probes were black/white gratings (one horizontal and one vertical), which were in all other respect identical to the color gratings. Probes were immediately followed by the presentation of a test screen. Test screens presented 9 colored squares (8 from the stimulus color to color opponent,⁵ and a white square) arranged in random order on a 3 × 3 matrix. Colors presented on the two screens corresponded to the horizontal and vertical probes, respectively. Participants were instructed to select the square that was “most similar to the appearance of the bright bars” in the preceding probe screen (i.e., the white bars). Responses were scored on a scale from +4 to -4. Positive values were assigned to the opposite colors with respect to the ones appearing in the stimulus blocks; white was coded as 0.

Recognition Test

At the end of the experiment, participants completed an incidental recognition test. Eight colored gratings (the 4 gratings that appeared in the stimulus blocks and 4 gratings with the color opposites of the colors used in the stimulus blocks) were presented in two rows in a randomized order. First, participants were instructed to select the gratings they recognized as having seen during the stimulus blocks. After that, participants were asked to indicate for each selected grating whether it appeared frequently or infrequently during the stimulus blocks.

Acknowledgments

This study was supported by the Hungarian Science Research Fund (OTKA T034112, and T-030739), and the National Institutes for Health (R01 DC04263). We thank Kinga Gyimesi, Lívía Pató, and Teréz Balázs for their assistance. Authors István Czizler and István Winkler contributed equally to the research reported.

Reprint requests should be sent to István Winkler, Institute for Psychology, Hungarian Academy of Sciences, H-1394 Budapest,

P.O. Box 398, Szondi u. 83-85, or via e-mail: iwinkler@cogpsyphy.hu.

Notes

1. A rhythmic electric fluctuation of approximately 40 Hz in frequency (the gamma-band potential) has been proposed to be associated with feature binding (Engel, König, Kreiter, Schillen, & Singer, 1992). It should be noted that the auditory-induced gamma band response is sensitive to attention, although not to specific levels of stimulus features (Tiitinen, Sinkkonen, May, & Näätänen, 1994).
2. Sussman, Winkler, and Wang (2003) showed that the elicitation of the MMN is only indirectly affected by selective attention. When similar deviations occur in two auditory channels, they compete for access to MMN generation. This competition can be biased by instructing subjects to focus on one channel and ignore the other channel. This explains the original contradictory findings about the attentional sensitivity of MMN (Näätänen, Paavilainen, Tiitinen, Jiang, & Alho, 1993; Woldorff, Hackley, & Hillyard, 1991). The current paradigm avoided this situation by presenting different types of deviations in the to-be-attended and to-be-ignored channel (see Experiment 1).
3. Responses obtained in the two subgroups with reversed standard-deviant stimuli were collapsed after finding no significant differences between them.
4. Responses obtained in the four subgroups (Standard-Deviant × Target) were collapsed after finding no significant differences between them.
5. Color opposites were empirically assessed in a separate negative afterimage matching procedure.

REFERENCES

- Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention: II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, *82*, 356–368.
- Awh, E., Dhaliwal, H., Christensen, S., & Matsukura, M. (2001). Evidence for two components of object-based selection. *Psychological Science*, *12*, 329–334.
- Berti, S., & Schröger, E. (2004). Distraction effects in vision: Behavioral and event-related potential indices. *NeuroReport*, *15*, 665–669.
- Berti, S., Schröger, E., Cowan, N., & Winkler, I. (2000). Attention and auditory sensory memory. In: C. Esceara, M. Tervaniemi, & E. Barcelona Yago (Eds.), *Abstract book of the second international congress on mismatch negativity and its clinical applications* (p. 67). Les Franqueses del Vallès: Grafiques Alle.
- Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 380–387.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109–127.
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception and Psychophysics*, *27*, 183–228.
- Coltheart, V. (1999). Introduction: Perceiving and remembering brief visual stimuli. In: V. Coltheart (Ed.), *Fleeting memories* (pp. 1–12). Cambridge: MIT Press.
- Cutting, J. (1976). Auditory and linguistic processes in

- speech perception: Inferences from six fusions in dichotic listening. *Psychological Review*, 83, 114–140.
- Czigler, I., Balázs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, 39, 869–873.
- Czigler, I., & Csibra, G. (1992). Event-related potentials and the identification of deviant stimuli. *Psychophysiology*, 33, 195–206.
- DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: Implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 27–42.
- Deutsch, D. (1975). Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 57, 1156–1160.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501–517.
- Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, 387, 808–810.
- Edgington, E. S. (1980). *Randomization tests*. New York: Dekker.
- Effron, R., & Yund, E. W. (1974). Dichotic competition of simultaneous tone bursts of different frequency: I. Dissociation of pitch from lateralization and loudness. *Neuropsychologia*, 12, 149–156.
- Engel, A. K., König, P., Kreiter, A. K., Schillen, T. B., & Singer, W. (1992). Temporal coding in the visual cortex: New vistas on integration in the nervous system. *Trends in Neurosciences*, 15, 218–226.
- Giard, M. H., Lavikainen, J., Reinikainen, K., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model study. *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.
- Hall, M. D., Pastore, R. E., Acker, B. E., & Huang, W. (2000). Evidence for auditory feature integration with spatially distributed items. *Perception and Psychophysics*, 62, 1243–1257.
- Heslenfeld, D. J. (2003). Visual mismatch negativity. In: J. Polich (Ed.), *Detection of change: Event-related potential and fMRI findings* (pp. 41–60). Boston: Kluwer Academic Publishing.
- Houck, M. R., & Hoffman, J. E. (1986). Conjunction of color and form without attention: Evidence from an orientation-contingent color aftereffect. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 186–199.
- Humphrey, G. K., James, T. W., Gati, J. S., Menon, R. S., & Goodale, A. (1999). Perception of the McCollough effect correlates with activity in extrastriate cortex: A functional magnetic resonance imaging study. *Psychological Science*, 10, 444–448.
- Jeffreys, D. A., & Axford, J. G. (1972). Source location of pattern-specific components of human visual evoked potentials: II. Components of extrastriate cortical origin. *Experimental Brain Research*, 16, 22–40.
- Kenemans, J. L., Jonh, T. G., & Verbaten, M. N. (2003). Detection of visual change: Mismatch or rareness? *NeuroReport*, 14, 1239–1242.
- Kenemans, J. L., Kok, A., & Smulders, F. T. Y. (1993). Event-related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirement. *Electroencephalography and Clinical Neurophysiology*, 88, 51–63.
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge: MIT Press.
- Marr, D. (1982). *Vision*. New York: Freeman.
- McCollough, C. (1965). Color adaptation of edge-detectors in the human visual system. *Science*, 149, 1115–1116.
- Murch, G. M. (1972). Binocular relationships in a size and orientation specific aftereffect. *Journal of Experimental Psychology*, 93, 30–34.
- Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*, 30, 436–450.
- Näätänen, R., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24, 375–425.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). “Primitive intelligence” in the auditory cortex. *Trends in Neurosciences*, 24, 283–288.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125, 826–859.
- Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology*, 38, 359–365.
- Pantev, C., Hoke, M., Lehnertz, K., & Lütkenhöner, B. (1989). Neuromagnetic evidence of an amplitopic organization of the human auditory cortex. *Electroencephalography and Clinical Neurophysiology*, 72, 225–231.
- Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., & Wittkowski, W. (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalography and Clinical Neurophysiology*, 69, 160–170.
- Pashler, H. E. (1998). *The psychology of attention*. Cambridge: MIT Press.
- Pazo-Alvarez, A., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. *Biological Psychology*, 63, 199–236.
- Picton, D. W., Alain, C., Otten, L., & Ritter, W. (2000). Mismatch negativity: Different water in the same river. *Audiology and Neuro-Otology*, 5, 111–139.
- Potter, M. C., Stiefbold, D., & Moryadas, A. (1998). Word selection in reading sentences: Preceding versus following contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 68–100.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179–197.
- Rafal, R. (1996). Visual attention: Converging operations from neurology and psychology. In: A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 139–192). Washington, DC: American Psychological Association.
- Rensink, R. A., O’Regan, J. K., & Clark, J. J. (1997). To see and not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373.
- 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 auditory memory: A review. *Ear and Hearing*, 16, 52–67.

- Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*, 1–46.
- Schröger, E., Rauh, R., & Schubö, W. (1993). Probability distribution of Minkowski distance between discrete random variables. *Educational and Psychological Measurement*, *53*, 379–398.
- Stagg, C., Hindley, P., Tales, A., & Butler, S. (2004). Visual mismatch negativity: The detection of stimulus change. *NeuroReport*, *15*, 659–663.
- Stromeyer, C. F. (1978). Form-color aftereffects in human vision. In R. Held, H. W. Leibowitz, & H.-L. Teuber (Eds.), *Perception: Handbook of sensory physiology* (vol. 8, pp. 97–142). New York: Springer-Verlag.
- Sussman, E., Gomes, H., Noursak, J. M., Ritter, W., & Vaughan, H. G., Jr. (1998). Feature conjunctions and auditory sensory memory. *Brain Research*, *793*, 95–102.
- Sussman, E., Winkler, I., Huottilainen, M., Ritter, W., & Näätänen, R. (2002). Top-down effects on stimulus-driven auditory organization. *Cognitive Brain Research*, *13*, 393–405.
- Sussman, E., Winkler, I., Kreuzer, J., Saher, M., Näätänen, R., & Ritter, W. (2002). Temporal integration: intentional sound discrimination does not modulate stimulus-driven processes in auditory event synthesis. *Clinical Neurophysiology*, *113*, 1909–1920.
- Sussman, E., Winkler, I., & Wang, W. J. (2003). MMN and attention: Competition for deviance detection. *Psychophysiology*, *40*, 430–435.
- Takegata, R., Huottilainen, M., Rinne, T., Winkler, I., & Näätänen, R. (2001). Changes in acoustic features and their conjunctions are processed by separate neuronal populations. *NeuroReport*, *12*, 525–529.
- Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letters*, *266*, 109–112.
- Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (2001). Preattentive processing of spectral, temporal, and structural characteristics of acoustic regularities: A mismatch negativity study. *Psychophysiology*, *38*, 92–98.
- Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch negativity in the visual modality. *NeuroReport*, *10*, 3363–3367.
- Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 128–140.
- Tiitinen, H., Sinkkonen, J., May, P., & Näätänen, R. (1994). The auditory transient 40-Hz response is insensitive to changes in stimulus features. *NeuroReport*, *6*, 190–192.
- Treisman, A. (1993). The perception of features and objects. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control. A tribute to Donald Broadbent* (pp. 5–35). Oxford: Clarendon Press.
- Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*, 1295–1306.
- Treisman, A. M., & Gelade, G. A. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Treisman, A. M., & Schmidt, H. (1992). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107–141.
- Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1656–1674.
- Widmann, A., & Schröger, E. (1999). Bootstrapping the distribution of the city-block distance between two repeated measures. Available at: <http://www.uni-leipzig.de/~biopsych/widmann/minkowski.html>.
- Winkler, I., Sussman, E., Tervaniemi, M., Ritter, W., Horváth, J., & Näätänen, R. (2003). Pre-attentive auditory context effects. *Cognitive, Affective, and Behavioral Neuroscience*, *3*, 57–77.
- Woldorff, M. G., Hackley, S. A., & Hillyard, S. A. (1991). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, *28*, 30–42.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*, 202–238.
- Wolfe, J. M. (1999). Inattentive amnesia. In: V. Coltheart (Ed.), *Fleeting memories* (pp. 71–94). Cambridge: MIT Press.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433.
- Woods, D. L., Alain, C., & Ogawa, K. H. (1998). Conjoining auditory and visual features during high-rate serial presentation: Processing and conjoining two features can be faster than processing one. *Perception and Psychophysics*, *60*, 239–249.
- Woods, D. L., Alho, K., & Algazi, A. (1992). Intermodal selective attention: I. Effects of event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography and Clinical Neurophysiology*, *82*, 341–355.
- Wyers, A. A., Mulder, G., Okita, T., & Mulder, L. J. M. (1989). Event-related potentials during memory search and selective attention to letter size and conjunction of letter size and color. *Psychophysiology*, *26*, 529–547.