Research report

Modulation of human auditory information processing by emotional visual stimuli

Veikko Surakka a, *, Mirja Tenhunen-Eskelinen b, Jari K. Hietanen a, Mikko Sams a,c

a University of Tampere, Department of Psychology, P.O. Box 607, FIN-33101 Tampere, Finland
b Tampere University Hospital, Department of Neurophysiology, P.O. Box 2000, FIN-33521 Tampere, Finland
c Laboratory of Computational Engineering, HUT, PO Box 9400, FIN-02015 Hut, Finland

Accepted 12 May 1998

Abstract

Auditory event-related potentials mismatch negativity (MMN) and N100 were recorded from seven subjects while they read text and watched emotionally negative, neutral, and positive pictures varying in valence and arousal. The MMN reflects automatic detection of change in auditory stimulus stream. Functionally different N100 is triggered by onset of various auditory stimuli. The N100 was stable during all visual conditions. The MMN was very similar during text reading, and neutral and negative slide viewing, but was significantly attenuated during viewing of positively valenced slides. We suggest that visual emotional information of high positive valence and low arousal is a signal of nonthreatening and nonappetitive environment. This kind of environment probably reduces the need for auditory change detection.

Keywords: Emotion; Mismatch negativity; Arousal; Attention; Amygdala; Evoked potentials

1. Introduction

Emotional responses have evolved because they are vital for survival and adaptation in environment [9,13]. It has been suggested that the processing of emotionally significant information and reacting to it can be automatic and precede conscious perception [20]. LeDoux has suggested that there are two basic modes of information processing: cognitive and affective [12]. There is also evidence that automatic generation of an affect precedes and molds cognitive processing [16,25].

Event-related potentials (ERPs) offer one way to study the effects of emotions on early sensory processing in humans. Previous studies have dealt with P3, P4, and late slow waves [8,23]. There are two well established measures of relatively early auditory processing. N100 deflection peaks at about 100 ms from the stimulus onset. The neurons underlying N100 are suggested to be capable of detecting the stimulus on-/offset and act as an internal attention trigger [17]. On the other hand, functionally different mismatch negativity (MMN), peaking at 100–300 ms from the stimulus onset reflects a deviance in any constant feature of repetitive auditory stimuli [14,17–19]. The MMN is automatic in a sense that it is elicited by changes in unattended stimuli [17,19].

The present aim was to study if visual emotional information affects basic auditory information processing as measured with the N100 and the MMN. ERPs to standard and deviant tones were recorded while the subjects read a text and watched visual stimuli differing in valence and arousal. We hypothesized that if emotional stimuli affect early auditory processing, it should be reflected in the amplitudes of the N100 and/or the MMN.

2. Subjects, materials, and methods

Eleven volunteer psychology students participated in the study. The data from four subjects had to be discarded because of excessive artifacts or lack of the MMN in the reading phase. Therefore, data from seven subjects (three males and four females) were used in the final analyses.

The auditory stimulus sequence consisted of frequent standard (1000 Hz) and infrequent deviant (1050 Hz) tones presented to the left ear (via eartube) at the probabilities of 0.85 and 0.15, respectively. At least 2 standard sounds
were presented between consecutive deviants. The constant stimulus onset asynchrony was 700 ms. The rise/fall time of the 100-ms, 40-dB stimuli was 30 ms (Hanning window). Auditory stimulus presentation was controlled by the Neuro Scan’s Stim™ program running on a Pentium PC.

Ninety digitized color pictures, thirty per each category, were chosen from the International Affective Picture System, IAPS [5]. The categories were negative, neutral, and positive (e.g., mutilations, mushrooms, and pleasant sceneries). The stimuli were presented using a Power Macintosh 7100/80 with 20-in., 1024×768-pixel, 75-Hz display. The stimuli were presented for 20 s separated by equally long blank screen. The stimulus presentation was fully randomized and controlled by PsyScope program [6].

The laboratory was introduced to the subjects and they were told that the purpose of the study was to investigate auditory processing during different kinds of visual information. Two conditions were run in the following order. First, in the reading condition the EEG was recorded for 15 minutes while the subject was reading a book. This was followed by the slide viewing condition during which the EEG was recorded for about an hour including two resting periods. Subjects were instructed not to move and, in addition, to fixate at the computer screen in the slide viewing condition. The viewing distance was 2.5 meters. The lights were dimmed and the slide sequence was started when the subject indicated being ready.

After the EEG measurement, the subjects rated their experiences evoked by the stimuli on two dimensions, valence and arousal (see Ref. [5]). The pictures were presented randomly. After both ratings were given the screen was replaced by the next picture. The ratings were made on nine point scales. The valence scale varied from unhappy (the lower end) to happy (the upper end). The arousal scale varied from calm to highly aroused. The responses were given using the keyboard and automatically recorded by the PsyScope.

Continuous EEG was recorded from 19 scalp sites (see Fig. 1) according to the international 10-20 system using Neuro Scan SynAmp preamplifier and Neuro Scan™ 3.0 digital EEG recording software on a 120 Hz Pentium PC. The ground electrode was placed half-way between Fpz and Fz. Electrodes attached below the right and above the left eye monitored eye movements. Linked ears were used as a reference. The electrode impedances were below 5 kΩ.

The EEG was amplified with a passband of 0.1–70 Hz (3 dB points) and later digitally filtered to 1–30 Hz. The 700-ms analysis period (sampling rate 1000 Hz) included a 100-ms prestimulus baseline. When there was activity exceeding ±65 µV in eye movement or Fpz channel the data from all channels were discarded from averaging.

Averaged responses to standards and deviants (minimum of 70 per category) were obtained during the reading condition and during the presentation of different emotional stimulus categories. For the quantification of the

![Fig. 1. Electrode sites and grand averaged auditory MMN difference waves at midline electrodes measured during different visual conditions.](image)
MMN, responses to standards were subtracted from those to the deviants. The MMNs in the reading condition were visually inspected to manually mark the approximate peak amplitude of the MMN. Then, the Neuro Scan’s Edit program was used to automatically find the latency at which the amplitude was at highest (20 ms integration window). This latency was used for measuring the MMN amplitude in different slide viewing conditions. The N100 peak amplitudes to the standard auditory stimuli were defined in a similar way.

The EEG frequency distribution during various slide categories was analyzed with the fast Fourier transform (FFT) analysis from 16-s period following the onset of a slide. The power spectrum measures were obtained for delta (1.5–3.5 Hz), theta (3.5–7.5 Hz), alpha (7.5–12.5 Hz), and beta (12.5–19 Hz) frequency bands and the relative power of each band was calculated.

The data were analyzed with repeated measures analyses of variance. All the reported significance levels are Greenhouse–Geisser corrected. Tukey’s HSD method was used for post-hoc tests.

3. Results

The mean ratings of experiences evoked by the positive, neutral, and negative slides were respectively for valence 7.2 ± 0.3 (mean ± S.E.M.), 5.1 ± 0.0, and 2.2 ± 0.3 (respective IAPS means 7.9, 5.0, and 1.8), and for arousal 4.0 ± 0.3, 3.5 ± 0.3, and 7.1 ± 0.4 (respective IAPS means 4.9, 3.0, and 6.6). A one-way ANOVA revealed a significant effect of slide category on valence ratings $F(2, 12) = 121.8$, $p < 0.01$. Post-hoc tests showed that all the pairwise differences were significant ($\alpha = 0.05$). For arousal ratings ANOVA also showed a significant effect of slides $F(2, 12) = 72.6$, $p < 0.01$. Post-hoc tests showed that the negative slides were experienced significantly more arousing than neutral and positive slides. The difference between positive and neutral slides was not significant.

Fig. 1 shows all the electrode sites and grand-average MMN waveforms for the midline electrodes during all visual conditions. The peak latencies of individual MMNs varied between 158–194 ms. In all slide viewing conditions, the deviants elicited a clear MMN. The MMN was of equal amplitude during the reading condition and during the neutral and negative slide viewing. However, the MMN was clearly attenuated during the viewing of positive slides. This pattern was also repeated on more lateral electrodes.

We first tested whether there were any differences between the MMN amplitudes in reading and neutral slide viewing conditions for the midline electrodes. A $5 \times 2$ (electrode $\times$ condition) two-way ANOVA showed a significant main effect for electrode $F(4, 24) = 17.4$, $p < 0.001$. This result is due to the fact that the MMN amplitudes are different at different channels. Importantly, no significant effect was found for the condition or the interaction of the
main effects. A similar analysis for the N100 revealed no statistically significant main or interaction effects.

Next we tested whether there were differences in the MMN amplitudes during the viewing of different slide categories. A $5 \times 3$ (electrode \times slide category) two-way ANOVA revealed a significant main effect of electrode $F(4, 24) = 6.5, p < 0.01$ and slide category $F(2, 12) = 6.6, p < 0.02$. The interaction was not significant. The main effect of slide category is obviously due to smaller MMN amplitudes during positive than other slide viewing conditions (see Fig. 1). The N100 component was analyzed similarly for the standard sounds. No statistically significant main or interaction effects were found (see Fig. 2).

Table 1 shows the relative amounts of theta and alpha activity at Pz during the different slide viewing conditions. One-way ANOVAs revealed no significant effects of the slide categories on these frequency bands.

<table>
<thead>
<tr>
<th>Power band</th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theta</td>
<td>0.23</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.25</td>
<td>0.27</td>
<td>0.28</td>
</tr>
</tbody>
</table>

4. Discussion

Our results showed that the auditory MMN was very similar during passive viewing of neutral slides as well as during active text-reading control condition. However, during the viewing of positively valenced slides the MMN amplitude was significantly attenuated compared to the MMN during the viewing of other slide categories. The N100 deflection was stable in all experimental conditions.

The MMN reflects memory-based, relatively automatic detection of change in auditory stimulus stream [14,17–19]. When subjects are required to pay very focused attention to another auditory or visual stimulus stream, the amplitude of the MMN to deviants in ignored sequence may be suppressed [10,17,18]. The present visual tasks did not require such focused attention. Also, we measured pitch deviance elicited MMNs which are known to be resistant to changes in direction of attention [18]. Therefore, we are confident that the present attenuation of the MMN is not due to our subjects paying more attention to positively valenced slides than to other visual stimuli.

It was possible that the changes observed in the MMN were due to changes in EEG arousal and not directly related to the emotional stimuli. However, the analyses of relative theta and alpha power spectrums between different slide categories revealed no such changes; these frequency bands are known to reflect well changes in general arousal [4]. Studies which followed changes of N100 and MMN in a continuum from alert/relaxed wakefulness to sleep found significant attenuation of these ERPs only after the subjects reached the level of stage 1 sleep [7,21]. Our subjects were not sleepy, at most they were in a state of relaxed wakefulness.

Recently, it was found (with stimuli drawn from the same picture set as in our study) that P300 deflection to auditory startle probes was attenuated during highly arousing (regardless of valence) pictures. In contrast to P3, the authors found that the amplitude of blink reflex varied as a function of emotional valence of the slides. The blinks were strongest during negatively and smallest during positively valenced pictures [23]. This result, with resemblance to our ERP findings, has been replicated in many other experiments [3]. The P300 modulation by affective arousal was suggested to index the mobilization of attentional and metabolic resources. The blink modulation by the valence was suggested to index motivational priming for appetitive or aversive behaviors [11,23].

The attenuation of MMN seems to be explained by the behaviorally rated valence of the stimuli because the positive and neutral slides did not differ significantly from each other in arousal dimension. In addition, although negative slides were rated significantly more arousing than neutral slides the MMNs were very similar between these visual conditions. It is noted, however, that our positive slides were low in arousal dimension. Thus, it is possible that the two dimensions interact so that emotional stimuli which are high in positive or negative valence and high in arousal create motivational priming for approaching or avoiding behaviors, respectively. Emotionally positive but only mildly arousing stimuli may serve as a cue of an environment which is neither interesting nor evoking explorative (approach) behavior. Thus, the automatic detection of change does not have to be as effective as possible. Possibly, if we had used positive slides rated to be high in arousal no attenuation of the MMN would have been observed.

The amygdala has been shown to be a central structure in the processing of emotionally meaningful information [13,15,24]. The primate amygdala is richly interconnected with the neocortex. It receives modality specific information from most sensory cortical areas. Some amygdalofugal projections to auditory areas have been found and projections to visual regions are strong [1,2]. Therefore, it has been suggested that the amygdala may have modulatory effects on relatively early stages of sensory information processing [2].

We propose a model where the amygdala is considered a critical mediator in early information processing. It is activated as a function of valence of emotionally meaningful information. Positive emotional information attenuates the amygdalar activity which, in turn, attenuates the mechanism responsible of detection of change in the auditory cortex [22]. The fact that no increase in the MMN was
found during the negative slides may be because the MMN was already maximal in the reading and neutral slide viewing condition. When changes in the MMN are observed, they usually are attenuating rather than accentuating by nature [10,14,17–19].

The results of this study showed that the auditory MMN was attenuated by positively valenced and little arousing visual emotional stimuli. These stimuli served as a biological signal of nonthreatening and nonappetitive environment. Thus, the need for automatic detection of change was not vital and this was reflected in the attenuated MMN. As the amygdala has been shown to be an important structure in the processing of emotionally meaningful stimuli, we suggested that relatively early auditory information processing is modulated by the amygdala.

Acknowledgements

This research was supported by the Academy of Finland. We thank Heikki Lyytinen for his constructive comments.

References


