Attention and multimodal integration in natural stimuli

Christian Muehl, M.Sc.
Human Media Interaction, EEMCS
University of Twente

Attention is a well-known phenomenon. Attending towards a certain feature or location in the environment is known to increase accuracy of decisions and to shorten the reaction time. While most of the fundamental research on cognitive phenomena like attention has been focusing on unimodal stimuli, it becomes more and more evident that in an integrated system like our brain the interactions between information derived from different modalities play an important role. In our daily experience we are surrounded by multimodal objects and the interaction with isolated entities present in only one modality is the exception.

In this respect recent studies delivered evidence for the theory that attention encompasses whole objects (Desimone and Duncan, 1995) rather than only isolated unimodal object features. Especially interesting is the fact, that also features of objects in unattended modalities are integrated into the focus of attention when another object feature is attended. In their EEG study Busse et al. (2005) presented visual stimuli in an alternating manner in one hemifield of the screen or the other. While one side was attended due to a block-wise assignment of task-relevance, the other was not. To maintain attention subjects were required to react to rare target stimuli. Simultaneously, a task-irrelevant sound was played from a central speaker. The event-related contrast between the condition where the sound appeared in an attended context, that is together with an attended visual stimulus, versus that where it appeared in an unattended context revealed differences in the processing of the auditory stimuli. The modulation originated in the primary auditory cortex and was expressed as a late (beyond 200 ms) attention-related negativity over fronto-central regions of the head. These results suggest that attention is spread over modalities and space to encompass other features of an attended object.

However, a serious limitation of most studies in the field of neurophysiological research is the application of simple artificial stimuli, like isolated shapes or pure tones. While it is a convenient method to control the dimensionality of the stimulus features, it compromises the generalizability of the results for a more natural context.

We set out to replicate the above mentioned study with two important modifications. Firstly, we used complex natural stimuli, namely pictures and sounds of animals, to examine the effect of object attention in an ecologically more valid setting. Secondly, we manipulated the multimodal object congruency by the presentation of matching and mismatching audio-visual stimulus pairs, to examine its effect on the spread of attention over modalities.

Though we observed several indicators for a mechanism of object-based attention, we found no significant difference in the main contrast of sounds presented in an attended versus an unattended context. However, there was evidence for a exogeneous reorientation of attention towards visual stimuli on the unattended side of the screen, which could eliminate the effect of the endogeneous attentional bias necessary for the late effect on auditory processing. Taking into account the human capability to categorize natural objects in the near absence of attention (Li et al., 2002), the lengthy presentation times (700 ms) are supporting a reorientation of attention due to the low attentional load they impose on the subject. As sustained attention to the visual modality is a prerequisite for the observation of the spread of attention we suggest several modifications of our experimental setup to yield decisive evidence regarding object-based attention in natural stimuli. Specifically, shorter presentation times and a decrease in the quality of visual stimuli might heighten the attentional load.

Nevertheless, we observe two remarkable and unexpected phenomena. On the one hand, we see a surprisingly early effect of semantic congruency on the integration of multimodal stimuli, starting already at about 70 ms past stimulus onset. This effect is seen over occipital regions for task-relevant stimuli. This is in line with other studies showing that task-relevance or attention is necessary for mechanisms of early multimodal integration. On the other hand, we find indicators for processes of multi-modal semantic integration even for unattended stimuli. In the ERP analysis we

observe a mismatch negativity in the contrast of congruent and incongruent presentations of unattended stimuli around 400 ms past stimulus onset over fronto-central cortices, a so-called N400.

Especially the very early effect of semantic integration for natural stimuli indicates that the importance of ecological relevance of stimuli should not be underestimated. This first attempt to find object attention in natural stimuli revealed marked differences to the processes observed with artificial ones. This could be due to the stimuli applied or to the methodological differences to the replicated experiment. Modifications of the experimental design have been suggested to optimize the conditions for a further investigation of object-based attention in natural stimuli. Overall, our findings indicate that the strive for ecological relevance is a promising trail to be followed in subsequent studies.

Multimodal integration and attention in the context of BCI

While the experiment described above was not conducted in the context of brain-computer interfacing, the study of attention and multimodal integration nevertheless bears some importance for this domain.

Especially the application of BCI technology in real-world settings, virtual and natural ones, is dependent on the understanding of the mechanisms of perception of complex multimodal stimuli. For example a system based on a P300-selection by visual attention on a virtual object (for example Bayliss et al., 2004) might be susceptible to potentials elicited due to attention on auditory stimuli in the environment. The knowledge about such interferences could improve the design of these virtual environments. On the other hand, the study of the intricate interactions between mechanisms of attention and multimodal integration could help to render current BCI approaches more robust or even to create new approaches. Especially with respect to a sensory feedback of the BCI performance, a more natural object-like feedback might support patients and non-disabled users in the acquisition of the skill of BCI usage.

Bibliography

Bayliss, J. D. (2003). “Use of the evoked P3 component for control in a virtual apartment,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), 113–116.

Busse, L., K. C. Roberts, R. E. Crist, D. H. Weissman and M. G. Woldorff (2005). "The spread of attention across modalities and space in a multisensory object," *Proceedings of the National Academy of Sciences*, 102, 18751-6.

Desimone, R & Duncan, J. (1995). Neural mechanisms of visual selective attention. *Annual Review of Neuroscience*, 18, 193-222.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, USA, 99,* 9596–9601.