Crossmodal Visual-Auditory Interference in Object Recognition Process

Shlomit Greenberg, Leon Y. Deouell

Hebrew University of Jerusalem, Department of Psychology

Yuvalsh@pob.huji.ac.il, msleon@mscc.huji.ac.il

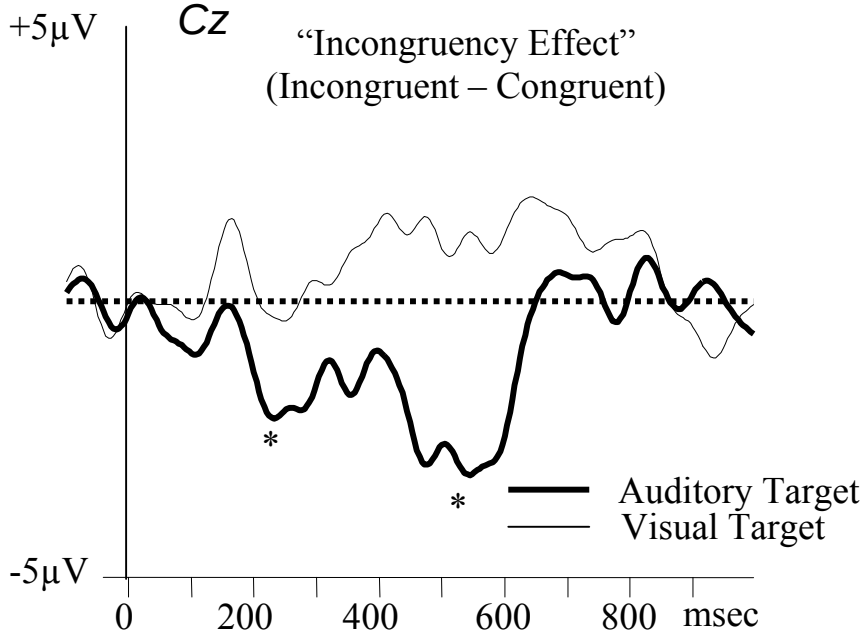
Classical cognitive neuroscience normally views the visual and the auditory systems as separate mechanisms, which operate independently. In the past few years an ever-increasing number of researches studied the interaction between the different modalities. The present research investigates how the auditory and the visual modules interact with each other during the process of object recognition.

The subjects were presented with bi-modal stimuli, composed of concurrently appearing pictures and voices of common and easy-to-identify animals (e.g., cow, rooster). The task was to identify either the picture ("visual condition") or the voice ("auditory condition") by answering a forced-choice question (e.g., “Rooster?”) presented after the auditory-visual presentation. In one third of the trials the picture and the voice belonged to the same animal ('congruent' trials), in another third the picture and the voice belonged to different animals ('incongruent' trials). In the remaining trials, a neutral stimulus, not identified as an animal, was presented as the non-target stimulus. Reaction times and accuracy of responses were analyzed.

Significant interference effects were found in incongruent trials in both the auditory and the visual conditions but were larger in the auditory condition. This 'incongruency effect' indicates that task-irrelevant objects of an unattended modality are processed and even recognized. This implicates a difficulty in suppressing object recognition processes of an unattended modality, despite its detrimental effect on performance. Another finding was the facilitation of object recognition in congruent trials of the auditory but not of the visual condition. That is, recognition of animal voices was facilitated by concurrent presentation of a congruent picture, whereas recognition of the pictures was unaffected by concurrent vocalization of the same animal. This finding, suggests a superiority of the visual system in the process of object-

recognition. Because of the design of the task, this facilitation was probably not due to simple response (yes/no) selection, but to genuine recognition or naming processes.

Event related potentials were recorded in 14 subjects during a similar design. Results showed that incongruent trials elicited a more negative response then congruent trials in latencies of circa 200-650 ms post stimlus onset when the target was the auditory stimuli, reminiscent of an N400 effect. In contrast, the incongruent trials were somewhat more positive than the congruent trials when the target was the visual stimulus. The figure shows this 'incongruency effect', calculated by subtracting the waveforms elicited during congruent trials from those elicited during incongruent trials. The difference-waves in the auditory condition shows a double peaked negativity which had a fronto-central distribution. Both peaks were significantly different from zero in a random effect analysis across the 14 subjects. In contrast, none of the peaks in visual incongruency waveform was significantly different from zero.



Thus, similarly to the behavioral data, the ERP data also shows an asymmetrical audio-visual interaction. The similarity between the behavioral measures and the electrophysiological measures suggests that the observed incongruency negativity is indeed related to the audio-visual interaction in object recognition.