“When sounds speak faster than words”: Audiovisual semantic congruency enhances early visual object processing

Yi-Chuan Chen & Charles Spence

Department of Experimental Psychology, University of Oxford

The present study examined crossmodal semantic priming effects elicited by naturalistic sounds or spoken words on early stage of visual picture processing where the picture was detectable but its category had yet to be fully determined. In each trial, an auditory prime was followed by a picture which was presented briefly and then masked immediately. The participants had to detect the presentation of any picture (detection task) or any picture belonging to the category of living- (vs. non-living-) things (categorization task). In the detection task, naturalistic sounds elicited a crossmodal semantic priming effect on picture sensitivity (i.e., a higher *d'* in the congruent than in the incongruent condition) at a shorter stimulus onset asynchrony (350 ms) than spoken words (1000 ms). In the categorization task, picture sensitivity was lower than in the detection task, but it was not modulated by either type of auditory prime. The results therefore demonstrate that semantic information from the auditory modality primed the early processing of a visual object even before knowing its semantic category. The faster crossmodal semantic priming effect by naturalistic sounds than by spoken words was attributable to the former’s accessing meaning directly, whereas the word’s meaning is accessed via lexical representations.

This study is supported by the Arts and Humanities Research Council (AHRC), Rethinking the Senses grant (AH/L007053/1).