Signal compatibility as a modulatory factor for audiovisual multisensory integration

Cesare Valerio Parise

The physical properties of the signals activating our senses are often correlated in nature; it would therefore be advantageous to exploit such correlations to better process sensory information. Stimulus correlations can be contingent and readily available to the senses (like the temporal correlation between mouth movements and vocal sounds in speech), or can be the results of the statistical co-occurrence of certain stimulus properties that can be learnt over time (like the relation between the frequency of acoustic resonance and the size of the resonator). Over the last century, a large body of research on multisensory processing has demonstrated the existence of compatibility effects between individual features of stimuli from different sensory modalities. Such compatibility effects, termed crossmodal correspondences, possibly reflect the internalization of the natural correlation between stimulus properties.

The present dissertation assesses the effects of crossmodal correspondences on multisensory processing and reports a series of experiments demonstrating that crossmodal correspondences influence the processing rate of sensory information, distort perceptual experiences and lead to stronger multisensory integration. After a brief introduction to the topic of multisensory processing and crossmodal correspondence in Chapter 1, in Chapter 2, the literature on crossmodal correspondences is critically reviewed. Based on a large body of research on crossmodal correspondences accumulated over more than a century, an inventory of the defining features of crossmodal correspondence is provided. Next, a taxonomy of crossmodal correspondence is developed. Finally the literature on the effects of audiovisual correspondence on human information processing is reviewed.

In Chapter 3, novel evidence for the effect of crossmodal correspondences on the speed and accuracy of human behavior is presented. A number of well-known examples of crossmodal correspondence, including the Mil-Mal effect, the Takete-Maluma effect, and the correspondence between auditory pitch and visual size are investigated using a modified version of the Implicit Association Test (IAT). Moreover, evidence is provided for two new crossmodal correspondences, namely the association between pitch and size of angles, and between the waveform of auditory signals and the roundedness of visual shapes.

In Chapter 4, psychophysical evidence is presented that crossmodal correspondences operate on a perceptual level, and systematically distort perceptual experiences. Human observers sometimes find it easier to judge the temporal order in which two visual stimuli have been presented if one sound is presented before the first visual stimulus and a second sound is presented after the second visual stimulus. This phenomenon has been term temporal ventriloquism. A manipulation of the crossmodal congruency between the visual and the auditory stimuli revealed a systematic modulation of the magnitude of this perceptual effect: Temporal sensitivity was higher for pairs of congruent auditory and visual stimuli than for incongruent pairs of stimuli. These results therefore provide the first empirical evidence that crossmodal correspondences operate on a perceptual level, and systematically distort perceptual experiences.

In Chapter 5, a series of psychophysical experiments showing that crossmodal correspondences modulate multisensory integration are described. Observers were presented with pairs of asynchronous or spatially discrepant visual and auditory stimuli that were either crossmodally congruent or incongruent, and had to report the relative temporal order of presentation or the relative spatial locations of the two stimuli. Sensitivity to spatial and temporal offsets between auditory and visual stimuli was lower for pairs of congruent as compared to incongruent audiovisual stimuli. Recent studies of multisensory integration have demonstrated that reduced sensitivity to perceptual estimates regarding intersensory conflicts constitutes the marker of a stronger coupling between unisensory signals. These results therefore indicate a stronger coupling of congruent vs. incongruent stimuli and provide the first psychophysical evidence that crossmodal correspondences promote multisensory integration.

In Chapter 6, an experiment investigating the role of the similarity of the temporal structure of visual and auditory signals for multisensory integration is presented. Inferring which signals have a common underlying cause, and hence should be integrated (i.e., solving the correspondence problem), is a primary challenge for a perceptual system dealing with multiple sensory inputs. Here the role of correlation between the temporal structures of auditory and visual signals in causal inference is explored. Specifically, it is tested whether correlated signals are inferred to originate from the same event and hence integrated optimally. In a pointing task with visual, auditory, and combined audiovisual targets, the improvement in precision for combined relative to unimodal targets was statistically optimal only when the audiovisual signals were correlated. These results therefore demonstrate for the first time that humans use the similarity in the temporal structure of multiple sensory signals to solve the crossmodal correspondence problem, hence inferring causation from correlation.
In Chapter 7, a Bayesian framework is proposed to interpret the present results whereby stimulus correlations, represented on the prior distribution of expected crossmodal co-occurrence, operate as cues to solve the correspondence problem, that is, to bind those signals that likely originate from the same environmental source while keeping separate those signals that likely belong to different objects/events. Finally, the findings of the present thesis are interpreted on the light of multisensory perceptual learning and development, and the relation between crossmodal correspondences and synesthesia is thoroughly discussed.

In spite of a century of research, the role of signal correlation on multisensory processing was largely unknown. Taken together, the present results demonstrate for the first time that human observers exploit the statistical correlation between multiple signals to solve the multisensory correspondence problem, and to more effectively process multisensory information.

Publications resulting from the present work:


