

The ability to efficiently integrate or segregate sensory information across modalities is a fundamental process in perception. While previous research has extensively explored spatial aspects of multisensory integration, the temporal dimension remains less well understood. The present study investigates the behavioural and neural dynamics underlying audiovisual temporal integration/segregation, focusing on the role of the conflict processing mechanisms. We employed a rate discrimination task where auditory flutter and visual flicker were presented concurrently but at different temporal rates, creating conditions of varying temporal disparity. We propose that temporal disparity between auditory and visual stimuli triggers a competition between two mutually exclusive perceptual hypotheses—common vs. independent causes—engaging conflict detection and resolution mechanisms. To examine this, we conducted two complementary pre-registered studies: a reaction time (RT, N=28) experiment to assess the perceptual decision-making dynamics and an electroencephalography (EEG, N=63) experiment measuring conflict-related theta-band activity at fronto-central electrodes. Consistent with our predictions, RTs decreased as temporal disparity increased, indicating faster evidence accumulation when audiovisual mismatch increased. And critically, EEG results revealed that theta-band activity exhibited a non-linear relationship with disparity, with greater power at intermediate disparities compared to both congruent and highly incongruent conditions. Furthermore, the temporal dynamics of theta activity closely tracked behavioural responses: theta emerged earlier for larger disparities, where decisions were reached more rapidly, and later for intermediate disparities, reflecting prolonged competition between integration and segregation hypotheses. The findings contribute to our understanding of how the brain arbitrates between integration and segregation across sensory modalities over time, offering insights into the neural mechanisms underlying multisensory temporal integration.