A Multimodal Paradigm for Investigating the Perisaccadic Temporal Inversion Effect in Vision

Leo G. Trottier (leo@cogsci.ucsd.edu)
Virginia R. de Sa (desa@cogsci.ucsd.edu)
Department of Cognitive Science, 9500 Gilman Drive
La Jolla, CA 92093-0515

Abstract

Unlike the auditory, or (for the most part) olfactory and somatosensory systems, the human visual system takes input in brief and discrete samples. In spite of this, we are able to act smoothly and perceive the visual world as temporally and spatially continuous. While the spatial aspect of the problem has received considerable attention at a neuroarchitectural level of analysis (Colby & Goldberg, 1999), its temporal dynamics (or how we solve the problem's temporal component) has only of late been given serious consideration. That we know relatively little is evidenced by the recent discovery that during a period (~70 ms) immediately before an eye movement, a pair of flashed stimuli will be perceived as having happened in the 'wrong' order (Morrone *et al.*, 2005). We establish experimentally that unimodal auditory stimuli are not subject to this distortion (Figure 3) and then use an audio-visual temporal order judgment paradigm to determine why and how this illusion occurs. By examining changes to the multi-modal point of subjective simultaneity we are able to discern how the processing of a single visual stimulus is altered as a function of when during eye movement generation the stimulus occurs. Establishing the schedule of visual remapping in fine grained detail will help in answering questions of how perceptual continuity is maintained and serve to inspire and constrain computational models of this aspect of visual processing

Keywords: vision; multimodal; temporal inversion; perisaccadic

Introduction

The temporal dynamics of sensory inputs are intimately involved in perception – at the very least for time judgments, but also for perception more generally. Most models of perception, however, assume passive, stimulus-driven computation (Engel *et al.*, 2001). Models that do involve temporal dynamics are typically self-contained motor control theory accounts (at a level of analysis within, rather than among, cognitive subsystems, e.g. Tweed *et al.* (1998)), take a more philosophical tack (e.g. Grush (2005)), or have been psychological models principally concerned with explanations of time perception, rather than the temporal dynamics of perception (e.g. Poppel (1997)). That research addressing the temporal dynamics of sensory processing is still in its early stages is evidenced by the recent finding that the perceived order of stimulus flashes can be reversed in vision immediately prior to a saccade (Morrone *et al.*, 2005). In the paper describing the effect, the authors theorize that this might be explained by appealing to the slowing of a neural 'clock', but provide no workable model. In contrast, our hypothesis is that this and other phenomena are a symptom of systems' attempting to synchronize themselves optimally. This hypothesis is based on the assumption that among multiple cognitive subsystems temporal coordination is required so that the information needs of one subsystem coincide temporally with the provision of information from other subsystems.

The physiological manifestation of information synchrony is likely coherent oscillatory neuronal activity, already implicated in models of top-down processing (Engel *et al.*, 2001). Importantly for perception research, eye movements (McAuley *et al.*, 1999) and attention (Fries *et al.*, 2001) have been associated with rhythmic, synchronous neural firing.

A need for information synchronization is perhaps found most acutely in the visuomotor system: the eye produces information for perceptual analysis but at the same time requires information for the performance of future movements. A simple serial model might posit that each eye movement informs the next, an arrangement that leaves processing idle while the eye is moving, and vice versa. More efficient is for motion and processing to happen in parallel – in this case, however, the problem of information flow between motor and processing systems becomes significantly more complex. Movement instructions must be sent to the motor system in synchrony with the system's capacity to enact them, while at the same time information is flowing back to visual processing for the planning of future eye movements. Fortunately, much of the visual system's operation is exposed, and in contrast to somatosensory or auditory perception, we are constantly and overtly sampling its environment with consistent and easy to monitor eye movements.

We have performed two experiments to tease apart the details of the temporal inversion illusion: we confirmed that the effect does not occur for sound stimuli (Figure 3) and then, in effect used aural stimuli as temporal 'landmarks'. We contend that two broad classes of mechanisms may be at play. The least general and informative case would be for the illusion to be caused by some interaction of the two stimuli in the visual pathway, meaning that it is in large part an artifact – if this is the case, when flashes are presented on their own there should be little or no temporal distortion. A second possibility would be for visual processing to be selectively distorting temporal perception during particular periods around saccade onset. In this case, the illusion could be caused by slower processing of stimuli appearing earlier in the pre-saccadic interval (during which the first flash would be presented), or faster processing of stimuli appearing later in the interval (during which the second flash would be presented), or both.

Background

The temporal inversion illusion was first reported by Morrone et al (2005). The investigators were interested in examining the effects of eye movement on time perception. The paradigm they used was straightforward (Figure 4) – while wearing an eye tracker, the subject was cued to prepare a saccade, and while doing so was presented two flashes in quick succession, one along the top of the presentation and the other along the bottom. The participant was then to indicate which stimulus had appeared first. Because of the use of the eye tracker, it was possible to after-the-fact determine when, with respect to the eye movement's physical onset, the flashes occurred. Morrone et al observed that when the mean stimulus onset time (MSOT) fell within a particular period before saccade onset, the flashes were consistently registered as having happened in the wrong order (Figure 5). Another study reported in the same paper examined the observer's perception of the interval between the flashes (ISI), and found that even if no inversion was reported, MSOTs closer to the critical interval yielded increasingly distorted perceived ISIs (with observers perceiving ISIs shortened by as much as 50% from the ISI presented).

Methods

Our aural-visual paradigm is identical to the one used by Morrone *et al.* except one of the flash stimuli was replaced with an aural 'click stimulus'. Subjects first fixated a circular black target (1 degree) on the left side of the screen, and indicated that they wished to begin the trial by pressing a button their left thumb. After an unpredictable interval, a target appeared approximately 30 degrees to the right of the fixation point. After the target appeared, stimuli were presented with an unpredictable latency designed to have the visual stimulus fall around the time when Morrone *et al.* found their strongest temporal effects. After saccade onset the subject was instructed to look to the target as quickly as possible. Time of target onset, stimulus order, and ISI were randomly assigned to each trial. Data were collected using an SR Research Eyelink II eye tracker system recording at 500 Hz. The subject indicated which modality was perceived first by pressing one of two buttons on a keypad. Since aural-visual timing judgments can be very plastic, only 90 trials occurred during a given experimental session, and the subject practiced making modality temporal order judgments during fixation. The experiment began only after order judgment accuracy exceeded 80% for 20 randomly ordered trials at fixation with ISIs evenly distributed between 16 and 88 ms. The subject was run in eight sessions, yielding 720 trials.

The aural-visual paradigm allows us to determine which case is most likely since it uses only one visual stimulus (substituting an audible 'click' for a second visual stimulus), thereby eliminating the possible confound caused by multiple visual stimuli. Analysis was performed offline to determine time of saccade onset, and trials were binned according to when around saccade onset the visual stimulus appeared. By observing how the point of subjective simultaneity ('PSS' – the flash-click temporal offset at which the observer is equally likely to report either observation) varies with respect to when the flash was presented pre-saccadically, we can see the effects of saccade preparation on a single visual stimulus.

It is possible to observe the shifting of the PSS in the following way. We represent the flash-click temporal offset along the real number line, setting (for example) the time of the flash as zero (Figure 1), and having the abscissa represent the relative onset time of the other stimulus (with distance from zero representing the amplitude of the ISI) . Along the ordinate we mark the proportion of times the subject indicated that the flash occurred first. We can then plot the perceived ordering of the stimuli as a function of their relative onset, and after fitting a psychometric curve, take the temporal offset at which the curve crosses the 0.5 line as the PSS. We determine what the PSS is under the control condition, and then observe how it changes contingent on when the flash stimulus occurred with respect to saccade onset. For example, if the PSS is -10 ms (i.e. the click needed to occur ten milliseconds before the flash for it to be perceived as simultaneous), then assuming that aural processing is not altered, any changes to the PSS are a result of changes in speed of visual processing. In this case, we might assume that aural processing takes ten seconds under normal conditions. If the PSS moves to -2 ms (rightward), then if aural processing still takes ten milliseconds, it must be that visual processing has been slowed (by eight milliseconds, in this case).

Results

Our results suggest that visual processing of stimuli occurring in the first part (80 to 40 ms pre-saccade) of the critical interval (in which temporal inversion was observed) is slowed, while in the later part (40 to zero ms pre-saccade) it appears to have quickened (Figures 1,2).

Discussion

The results, while compelling and statistically significant, are only from one subject, and we are currently collecting data from a second. The data suggest that a means by which we maintain visual continuity is by slowing and speeding processing to cover any temporal 'gap' in information processing. The time windows in which we find the effects are supported by Morrone et al's (2005) findings (Figure 5) – it appears that inversion may take place because processing (or visual latency perception) differs depending on when prior to eye movement onset a visual stimulus is presented. According to our results, Morrone et al's (2005) visual inversion effects are most reliable and strongest in the period around 50 ms pre-saccade because it is at this point when the two stimuli straddle the two time windows in which stimuli are perceived to have different latencies. Stimuli falling in the first window are perceived to have longer than stimuli appearing in the latter window. Most intriguingly, our data thus suggest an abrupt change in how visual information is processed at around 40 ms pre-saccade onset, a change which should be observable electrophysiologically, and begs to be explained neurophysiologically.

Figures

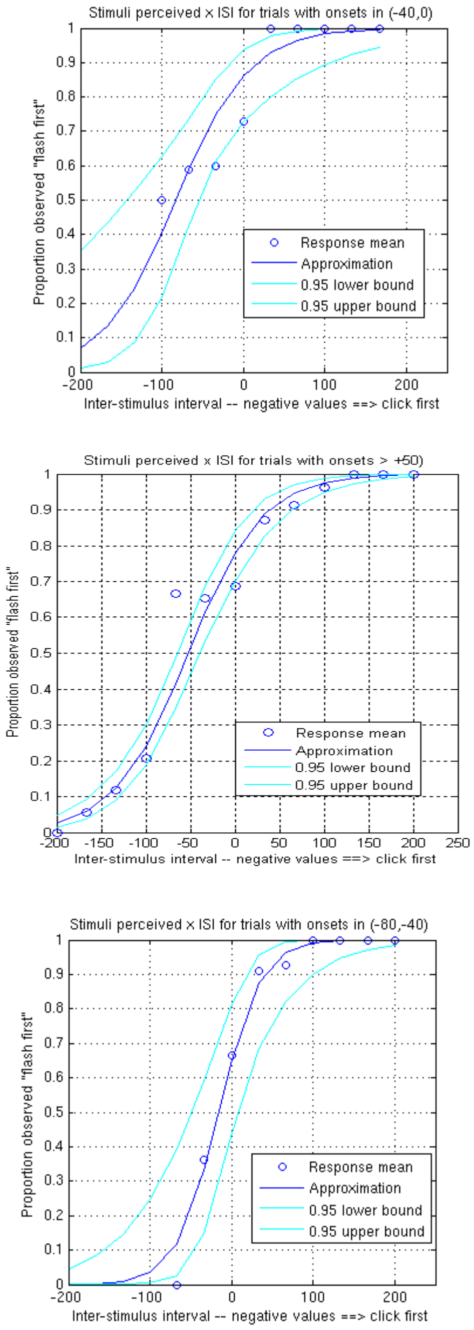


Figure 1: A psychometric curve indicating a shifting point of subjective equality, dependent on window of flash presentation. Top: Note in the later time window (-40,0) the curve shifts leftward, indicating that the subject is perceiving these stimuli more quickly than normal (middle), as greater lead time of the click stimulus is required for perceived simultaneity.

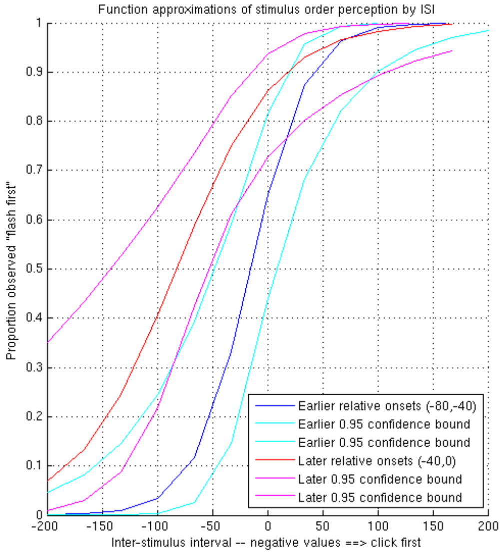


Figure 2: When overlaid, the psychometric curves show the significance of the differences in PSS. The curve for visual stimuli in (-80,-40) is significantly to the right of the curve for the later (-40,0) window, implying that visual stimuli in the (-80,-40) window are processed significantly slower than those in the later one.

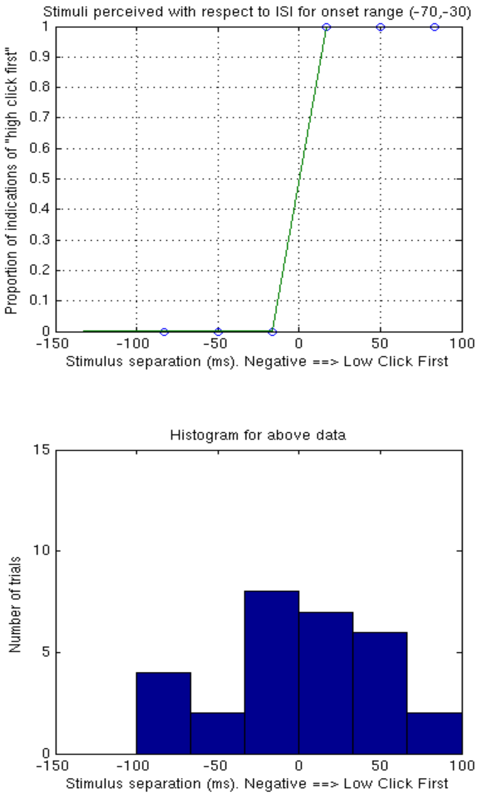
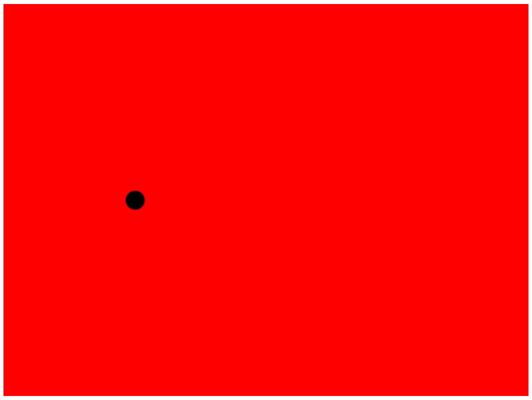


Figure 3: Above plots show that auditory stimuli are not subject to temporal inversion. Performance remains perfect throughout the pre-saccade time window.



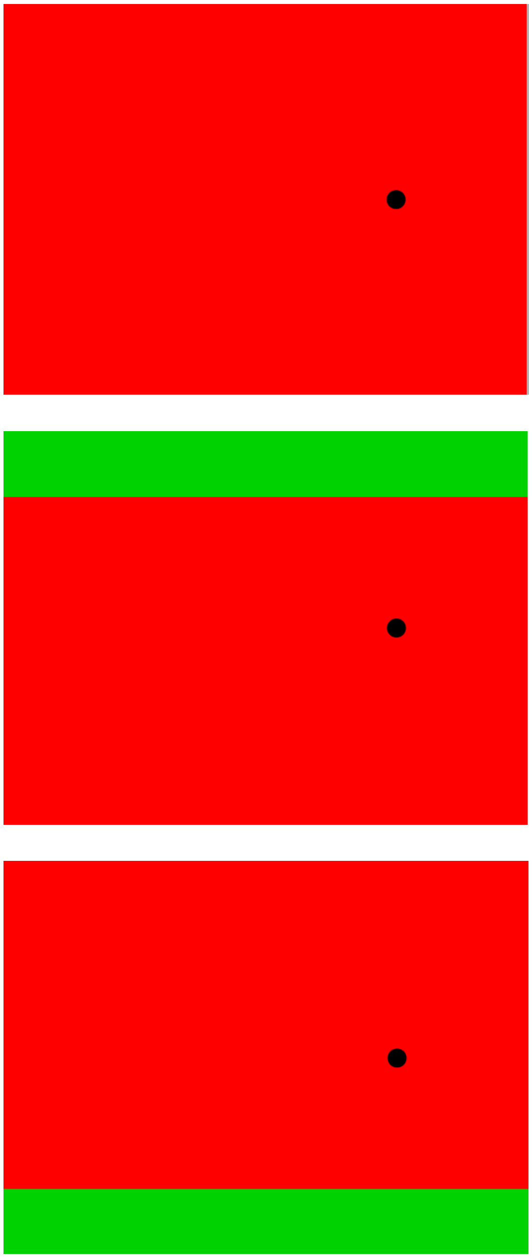


Figure 4: The four events in the experimental sequence of Morrone et al (2005). The subject fixated a point on the left side of the screen. A target appeared to the right side of the screen. While the subject was preparing to saccade, two 8 ms flashes with varying ISIs were presented at the top and bottom of the screen. The multimodal paradigm is identical, save that one of the flashes is replaced by a audible 4 ms ‘click’.

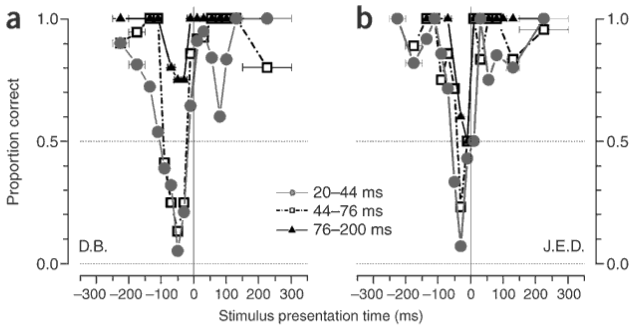


Figure 5: From Morrone et al (2005). For two subjects, proportion correct for order judgments as a function of presentation time. Data are plotted for three different ISIs (20-44, 44-76, 76-200). Note that at around -40 ms subjects are nearly always ‘wrong’ (perceiving flashes as happening in the reverse order).

Reprinted by permission from Macmillan Publishers Ltd: *Nature Neuroscience* (Morrone et al), copyright 2005.

Acknowledgments

This work is supported by an NSERC PGS-M to L.G. Trottier and NSF Career Grant 0133996 to V.R. de Sa and NSF Science of Learning Center grant to G.W. Cottrell. An earlier version of this work was presented at the graduate student conference COGNITIO (Trottier, 2006) and will be presented as a poster at COSYNE 2007.

References

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. Nature Reviews Neuroscience, 2(10), 704-716.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. Science, 291(5508), 1560-1563.

Grush, R. (2005). Brain time and phenomenological time. In A. Brook & K. Akins (Eds.), Cognition and the brain: The philosophy and neuroscience movement (pp. 160-207). Cambridge; New York: Cambridge University Press.

McAuley, J. H., Rothwell, J. C., & Marsden, C. D. (1999). Human anticipatory eye movements may reflect rhythmic central nervous activity. Neuroscience, 94(2), 339-350.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. Nature Neuroscience, 8(7), 950-954.

Poppel, E. (1997). A hierarchical model of temporal perception. Trends in Cognitive Sciences(1), 56-60.

Trottier, L. (2006). An exploration of the temporal inversion effect in vision, COGNITIO. Montreal, Canada.

Tweed, D., Haslwanter, T., & Fetter, M. (1998). Optimizing gaze control in three dimensions. Science, 281(5381), 1363-1366.