Object Perception: Separating the Contributions of High and Low Level Visual Processes with Event Related Brain Potentials

by

Carley Graceanna Piatt
B.A., University of Colorado, 2002

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCES

in the Department of Psychology

Object Perception: Separating the Contributions of High and Low Level Visual Processes with Event Related Brain Potentials

by

Carley Graceanna Piatt
B.A., University of Colorado, 2002

Supervisory Committee

Dr. J. W. Tanaka, (Department of Psychology)

Supervisor

Dr. C. B. Holroyd, (Department of Psychology)

Departmental Member

Dr. M. Masson, (Department of Psychology)

Departmental Member

Supervisory Committee

Dr. J. W. Tanaka, (Department of Psychology)

Supervisor

Dr. C. B. Holroyd, (Department of Psychology)

Departmental Member

Dr. M. Masson, (Department of Psychology)

Departmental Member

Abstract

Object recognition was studied by combining a continuous presentation paradigm and event related potentials (ERPs). Using the Random Image Structure Evolution program (RISE), the phase spectrum of object images was parametrically altered to produce a set of continuous noise-to-object sequences. The RISE technique controlled for the low level visual properties of the object image (i.e., luminance, contrast, spatial frequency). Although the stimulus transformation proceeded continuously and smoothly, perceptually, participants reported the abrupt onset of object recognition, distinct from noise, at a critical frame in the sequence. During electrophysiological recording, the critical onset frame was marked by increased activity in posterior-occipital and central-parietal components between 152 and 324 ms post stimulus onset of the critical frame. More broadly, this study also highlights the strength of the continuous presentation paradigm for investigating object adaptation effects with ERPs.

Supervisor: Dr. James Tanaka, (Department of Psychology)

Table of Contents

List of Figures

Acknowledgments

First, I am deeply grateful for the full intellectual and financial support of my advisor, Jim Tanaka, without whom, none of this work would have been possible. I embarked on this journey at a time in which all of the components needed to make our ideas a reality were already in place and am forever indebted to Jim for allowing me to be first to fit the pieces together. This work would not have been possible without the support and collaboration of Javid Sadr who generously allowed me access to the RISE image processing algorithm and who offered creative and supportive comments throughout the entire process.

I am also indebted to Dev Balachandran, Kathryn Morrison, Lara Pierce, Jordan Reems and Davin Shikaze for assisting with data collection, Olav Krigolson for helpful discussion concerning PCA and management of the Brain and Cognition Lab, the members of the Perceptual Expertise Network, especially Tim Curran, for helpful discussion, Sandra Vella for her friendship and e-cards of encouragement, Robin Yates for long walks on the beach and wine, and Peter Lee for his unflagging support, love and comments on a draft of this thesis. I would also like to thank the members of my committee for their insights: Clay Holroyd and Mike Masson and my external examiner, Alexandra Albu. Finally, not least of all, my deepest gratitude to my parents, Sue and Terry Piatt for their enduring love, patience and support in, and through, all things.

This work was supported by grants to Jim Tanaka from the Natural Sciences and Engineering Research Council of Canada and the James S. McDonnell Foundation.

Neural Correlates of the Onset of Object Perception

Under normal viewing conditions, object recognition is effortless, accurate and almost instantaneous. To illustrate: Thorpe, Fize and Marlot (1996) asked participants to decide whether an animal appeared in a briefly flashed natural scene. They found that an exposure of 20 ms was sufficient to produce accurate recognition performance. Recordings of event related potential (ERPs) during the task revealed that differences between object-present and object-absent trials began to emerge at about 150 ms post stimulus onset in frontal recording sites. Based on these findings, they concluded that processing of object recognition is completed within the first 150 ms post stimulus onset (Thorpe, Fize, & Marlot, 1996). Consistent with this view, Grill-Spector and Kanwisher (2005) employed a visual masking paradigm to evaluate the amount of stimulus exposure needed to detect an object ("is an object present?") versus the time needed to categorize an object ("is the object a car?"). Their results indicated that a stimulus duration of 50 ms led to accuracy above chance for both object detection and object categorization and that accuracy did not differ between the two kinds of decisions. These findings suggest that the process of object categorization occurs concurrently with the process of object detection. Taken together, the empirical findings confirm what we know intuitively to be true about object recognition – that it is almost instantaneous and often accurate.

Although this work suggests that object recognition is completed within the first 150 ms of stimulus presentation, what remains elusive about the process of object recognition is a clear understanding of the relative contribution of low and high level visual information in object recognition. Low level visual information refers to the visual properties of the object, e.g. contrast and spatial frequency information, which activate a feed-forward process in which visual information is detected by low level cortical areas, i.e. V1, before being projected to higher level cortices, i.e. inferiortemporal cortex (Tanaka, 1996). In contrast, high level visual information refers to the recognition of visual information as a coherent or meaningful representation of an object. The results of Thorpe et al., (1996) and Grill-Spector and Kanwisher (2005) suggest that within 150 ms post stimulus onset, both low and high level visual information have been accessed to recognize an object. However, the temporal dynamics between low and high level visual remain unknown.

Classically, one way of studying the underlying processes of object recognition behaviourally has been to slow the process down by degrading the stimuli. Bruner and Potter (1964) conducted experiments in which blurred objects were brought into focus until participants reported correctly identifying the object. Objects were initially presented either in a very blurry condition, a somewhat blurry condition or a less blurry condition and the speed with which the images were brought into focus varied. They found that object recognition was adversely affected by the initial blurriness of the image, as well as by longer exposure to blurry stimuli. These findings are interpreted as indicating that the more time available to generate incorrect hypotheses about the object, the more difficult it becomes to correctly recognize it (Bruner & Potter, 1964).

Other studies have used contour deletion (Doniger et al., 2000; Snodgrass & Corwin, 1988; Viggiano, 2000) with a sequential presentation paradigm in which decreasingly degraded images are presented for set durations. At each image, participants are asked if they are able to recognize the object. If they are able to recognize the object, this level of degradation is considered the level of identification. If not, the next least degraded image is shown and they are again asked if they are able to recognize the object, with the sequence continuing until the object is recognized. Though these studies are informative in that they provide information on the effects of interference on visual recognition (Bruner and Potter, 1964) and on phenomena like the Gestalt perceptual closure effect (Doniger et al., 2000), they do not provide insight into effortless object recognition. Similar criticisms have been raised elsewhere (Gauthier, 2000) because slowing down the process of object recognition invokes effortful and prolonged processing which is the antithesis of object recognition under normal viewing conditions.

The paradigms mentioned above fail to address the abrupt nature of object recognition for three reasons. First, studies of object recognition that ask participants to inspect a degraded or incomplete image for a period of time and then ask "Do you recognize the object?" are, by their very nature, invoking effortful processing. That is, participants may be engaged in a hypothesis-testing strategy that is not typical during normal, everyday object recognition. For example, as more information appears on the screen, participants may hypothesize that a particular cluster of pixels looks like a leg of an animal and that the image could be a dog. As a result, the dog hypothesis is tested and the processing that occurs is thus related to confirming the hypothesis. Alternatively, an incorrect hypothesis might interfere with subsequent recognition (Bruner and Potter, 1964). In light of the rapid and accurate nature of object recognition, it seems unlikely that this hypothesis-testing strategy stimulates the processes of object recognition that are reported to occur within 150 ms of stimulus onset.

A second drawback found in previous approaches is that the stimuli used in these experiments fail to control for differences in low level visual properties. For example, studies with the original Snodgrass and Vanderwart set (Snodgrass & Vanderwart, 1980), employed fragmented images (Doniger et al., 2000; Viggiano, 2000) that visibly change in the amount of low level visual information available including changes in contrast, luminance and spatial frequency. Researchers have also used systematically blurred images (Bruner & Potter, 1964) as well as images that deliberately change in spatial frequency (Bar et al., 2006; Norman & Ehrlich, 1987). These manipulations alter the images in important ways that have been shown to influence object recognition. For example, it has been demonstrated that the contributions of high spatial frequency information to object identification are greater than that of low spatial frequency information (Norman & Ehrlich, 1987). Moreover, changing contrast and luminance information is problematic when studying the electrophysiology of object recognition because these stimulus parameters are known to modulate early ERP components (e.g., C1, P1) associated with visual processes (Luck, 2005). Thus, in studies that use degradation techniques without controlling for variability in low level visual properties, unintended consequences associated with that variability may affect both behavioural and neurophysiological results.

The third limitation is that much of the research using paradigms with contour deletion, or any successive presentation of monotonically degraded stimuli, may be confounded by priming effects. That is, studies have specifically used contour deletion to investigate the effects of perceptual priming on object recognition (Biederman & Cooper, 1991). Perceptual priming occurs when a stimulus is specifically primed by visual information. For instance, Biederman and Cooper (1991) investigated the extent to which parts and features primed perceptually equivalent images (e.g. how well does an image with fifty percent contour deletion of, say, a flashlight prime that particular flashlight?) and conceptually equivalent images (e.g. how well does an image with fifty percent contour deletion of a square padlock prime a round padlock?). Their results showed that images with fifty percent contour deletion primed the perceptually equivalent image with the same accuracy and same reaction time as the identical image with no contour deletion. Moreover, their results showed that the images with a fifty percent contour deletion also primed the conceptually similar image enough to decrease the mean reaction time by approximately 100 ms and to decrease errors slightly. Thus, the priming effects are shown to be especially salient in the case of contour deleted images at the perceptual level and to a lesser degree at the conceptual level. Studies interested in teasing apart the effects of low and high level visual processing, would have to consider priming effects, as priming clearly has some influence on both types of processing. Hence, a final goal of this research is to capture the abrupt moment of object recognition in the absence of priming effects.

To date, few studies have investigated the electrophysiological correlates of abrupt, instantaneous object recognition. Electrophysiological studies with degraded stimuli (systematically fragmented versions of the original Snodgrass and Vanderwart set) have found differences in ERP components associated with visual processing and repetition priming effects when objects are recognized within a discrete presentation paradigm of successively more complete images (Doniger et al., 2000; Viggiano, 2000). Doniger and colleagues (2000) recorded ERPs while successively less degraded images were presented. A level of identification was recorded for each image and this level was called the ID level. ERPs associated with the ID level and the three levels prior to identification were analyzed. Their results indicated that early visual ERP components, the P1 and N1, did not show any significant differences with the onset of the ID level or any of the levels prior to identification. They did however find that at approximately 230 ms post stimulus onset, activity in occipito-temporal channels showed a large increase in negativity that peaked at 290 ms post stimulus onset to the onset of the image indicated as the level of identification. These same channels showed increasing negativity as the images approached identification. Doniger et al., (2000) suggested that these effects demonstrate that the onset of perceptual closure in object recognition builds successively over time and does not occur all at once.

Viggiano and Kutas (2000) also used contour deleted images to study the effects of encoding on object recognition. They exposed participants to intact images and asked them to either name the image or draw the image. In the identification phase which followed, participants were asked to recognize contour deleted images that had been previously studied or new images, which were presented from most degraded to least degraded using a similar paradigm as Doniger et al., (2000) while ERPs were recorded. They found that studying images improved the level at which identification occurred, certainty in identification and decreased reaction times, compared with new images. With regard to the temporally early ERP components, they found a main effect of fragmentation (level of contour deletion) for the posterior P100 component that increased in positivity as the images moved towards completion (less fragmentation or contour deletion). A similar effect was found for the posterior P2. In addition, they found that the P2 showed a slight increase in amplitude for unstudied images. The identification level was found to elicit a large positivity between 250 and 1000 ms post stimulus onset across all channels and that this effect was more abrupt for unstudied items and slightly more gradual for studied items (Viggiano, 2000). Their findings also provide neural evidence for the gradual build up of object recognition and further show that this is modulated by prior exposure to the stimuli. However, these studies investigated perceptual closure (Doniger et al., 2000) and the effects of prior exposure on object recognition (Viggiano, 2000). As a result, with regard to understanding the neural mechanisms of effortless object recognition, these studies are subject to criticisms associated with invoking hypothesis testing and priming effects (Gauthier, 2000).

The focus of this research it to address the limitations outlined above that have encumbered previous studies of object recognition. First, to address the limitations imposed by failing to account for changes in low level visual information, an image processing algorithm called Random Image Structure Evolution (RISE) was used. RISE degrades or alters images parametrically so that only the phase information of an image is changed (Sadr & Sinha, 2001, 2004). The amount of phase information change can be manipulated parametrically while keeping the low level visual properties of an image (i.e., luminance, contrast, spatial frequency) constant as illustrated in Figure 1. Most importantly, by imposing visual continuity on the stimuli, it is possible to invoke a metric that relates the stepwise images to one another along a continuous trajectory (Sadr & Sinha, 2004). The advantage of being able to relate the images to one another parametrically, suggests that discontinuities or categorical changes in responses while moving along the trajectory, might be related to the onset of high level visual events ((Sadr & Sinha, 2004). Thus, RISE allows for the systematic degradation of objects without compromising results related to the perception of an object with processing related to changes in low level visual information.

Second, the current study employs a continuous presentation paradigm in which a sequence of RISE images are shown with no intervening blank intervals or masking stimuli. As shown in Figure 2, each trial begins with a fixation point followed by the onset of the first image followed immediately by a second image and so on until all of the images in a sequence have been presented. Jacques and Rossion (in press) showed that the use of a continuous presentation paradigm with ERPs resulted in the adaptation of early visual components (the P1 and N1). Their study illustrates the advantages of using adaptation to elucidate processes that vary for different stimuli (e.g. different faces) but utilize the same neural, visual system overall. By attenuating the activity in the same overall neural system with the presentation of blocks of stimuli with no intervening blanks, Jacques and Rossion (in press) showed that occipital channels were sensitive to face individuation as early as 130 ms post stimulus onset. Their findings suggest that by using a similar continuous presentation paradigm, similar adaptation effects in channels associated with visual processing may be found and would thus provide insight into the neural underpinnings of visual processing corresponding to the differentiation between the perception of incoherent visual information and object perception. Being able to differentiate between neural processing related to overall visual processes and visual processes related specifically to object perception, would make it possible to study more fine-grained perceptual processes such as distinctions between object categories and individuation within a category, as Jacques and Rossion have done (in press), without introducing variance related to having to re-engage visual processing with every stimulus presentation.

Perceptually, when images processed with RISE are presented continuously, it appears that the first half of the sequence contains only noise frames and then at the critical "onset" frame, an object abruptly emerges. There are three important advantages for studying object recognition using this approach: 1.) The low level visual properties of the images, except their phase spectrum, do not change, 2.) given the abruptness of the object appearance, it is unlikely that participants can effectively use a hypothesis-testing strategy, and 3.) perceptual priming effects are markedly reduced or eliminated given the lack of visual overlap between successive displays. Next, because object recognition occurs in the order of milliseconds (Thorpe et al., 1996; Grill-Spector et al, 2005), event related potentials (ERPs) which have precise temporal resolution, provide an ideal method for investigating the temporal dynamics of object recognition. Combining the strengths of the RISE algorithm with a continuous presentation paradigm obviates the limitations inherent in other studies of object recognition and capitalizes on the temporal resolution of ERPs. Finally, the implication of a continuous presentation paradigm employing RISE with ERPs is that the abrupt onset of object recognition is preserved while the low level visual properties of visual processing will be attenuated by adaptation within the sequence thereby making it possible to study high level visual processing related specifically to the onset of object perception. In addition, because it is apparent to an observer each time a new image appears on the screen, it will be possible to study how an increase in phase coherence affects neural processes.

In the current experiment, the goal was to capture the instantaneous and automatic nature of object recognition and to investigate the neural correlates of the onset of object recognition. Because parametrically changing images were used within a continuous presentation paradigm, it is predicted that ERPs associated with visual processing will show adaptation effects until the onset of the a priori frame of object coherence. In addition, using RISE to produce object images that change incrementally from incoherent to coherent and asking participants to indicate the onset of object perception makes it possible to establish when object perception will occur with ERPs a priori. Thus, with an index of object perception established that is consistent across all participants and all objects, an empirically defined frame of recognition may be used as the a priori frame of interest to study the neural correlates of object recognition.

The first set of experiments in this thesis is concerned with finding the categorical boundary between noise and object coherence behaviourally, and then using this information to look at the neural correlates associated with this boundary. The second set of studies address whether the results of the first experiment could simply be due to detecting a change in visual information or, if the effects are related to a specific kind of information change.

Experiment 1a

The first experiment was designed to find a set of stimuli with a clear categorical boundary between the noise images and the onset of object perception. Establishing this boundary was important for several reasons. First, a primary goal of this research was to develop a paradigm that captured the instantaneous recognition that is characteristic of normal object processes rather than the effortful hypothesis-testing strategies and the attendant priming effects found in previous studies. The second goal of this research is to remove the relative contribution of low level visual information from high level recognition. In this experiment, the phase spectrum of one hundred and eighty line drawings from a new set of images based on the Snodgrass and Vanderwart set (Rossion and Pourtois, 2004) were parametrically manipulated in RISE to yield a ten step, noise-to-object sequence. Participants then viewed the ten frame sequence on a computer screen, with each frame shown for 500 ms. When participants recognized the object, they pressed the "o" key and typed in the name of the target object. The goal of Experiment 1a was to identify those objects in which there was a distinct category boundary between the noise frame and the object frame.

Method

Participants. Seventeen members of the University of Victoria community participated in this study. Nine were male and all participants had normal or corrected-to-normal vision and all were native English speakers.

Materials. Stimuli for the experiment consisted of 180 black and white line drawings with shading and texture information, overlaid on a noise background. The phase spectrum of each original object image was parametrically altered to produce a set of continuous noise-to-object sequences using Sadr and Sinha's RISE algorithm (Sadr and Sinha, 2001, 2004) (Figure 2). Images were 281 x 281 pixels and subtended a horizontal and vertical visual angle of approximately 8 degrees. The parameters of RISE were defined so that the most coherent object (the original image) had a phase coherence of 100% and the least coherent object image had a phase coherence of 10%. The step size chosen resulted in ten successive images increasing in phase coherence by 10% (Figure 2), each moving from noise to object coherence (noise-to-object).

Procedure. Each object sequence began with a fixation point (+) followed by the presentation of the least coherent image, showing no distinguishable object within the noise, for 500 ms followed by the immediate presentation of the next least coherent image for 500 ms and so on until the image reached complete coherence (the original image of the object overlaid on the noise background) (Figure 2). Participants viewed the noise-to-object sequences and indicated with a button press the instant the object could be recognized. The button press terminated presentation, and participants typed in the name of the object in order to later verify accuracy in recognition. The stimuli presentation and data collection were accomplished using Matlab 7.0.4.

Results

The following information was computed for each noise-to-object sequence: average frame of recognition; standard deviations around the mean and accuracy in naming. The data for correctly named objects clustered predominately at and around frame seven. Hence, the seventh frame was identified as the frame at which participants categorically distinguished the object from the noise. From the total set of objects, fifty were identified that were recognized at the seventh frame (S.D. = .25), with at least 90% accuracy (Figure 3c). These fifty noise-to-object sequences were used in the subsequent ERP experiment (Experiment 1b).

Discussion

The goal of this study was to find a clear boundary between the perception of noise and the perception of an object. The results show that a set of stimuli could be found that had a clearly defined boundary between the onset of object perception from noise occurring at the same frame for all fifty images with 90% agreement among participants. The results also show that with the RISE parameters chosen and the continuous presentation paradigm, an index of object perception was established that could be used as the a priori frame of interest in the ERP study (Experiment 1b).

Experiment 1b

This experiment was designed to examine the relative timing and contributions of low level and high level visual processes to object perception. In order to obtain the clearest results possible about the onset of object perception, a passive viewing task in which participants viewed images appear on the screen, was thought to be an ideal task to engage perceptual attention without requiring additional cognitive processing. To successfully investigate the neural processes of object recognition in a passive viewing paradigm of this nature, it was established a priori, in Experiment 1a, when in the sequence of images, participants experienced object perception. In addition, because a focus of this study is to capture the abrupt onset of recognition that is ecologically similar to object recognition in everyday contexts, a final aim was to avoid testing those images with high inter- and intra-subject variability in the ERP study.

Stimuli identified in Experiment 1a with a consistent categorical boundary between noise and the onset of object perception were used to investigate the neural correlates associated with object recognition. With Frame 7 as an index of object recognition in the context of the continuous presentation paradigm, changes in the ERPs at Frame 7 are predicted to show differences and that those differences suggest processing specifically related to the perception of an object relative to noise.

Method

Participants. Fourteen undergraduates from the University of Victoria participated in this study. Nine were female, all were right handed; all participants had normal or corrected-to-normal vision and all participants were native English speakers. The mean age was 19 years. Participants gave written informed consent after the experimental procedure was explained and were received course credit for their participation.

Materials. Stimuli were fifty noise-to-object sequences identified from Experiment 1a (Figure 3a). In addition to the fifty noise-to-object sequences, the same parameters were applied to create fifty unique noise-to-noise sequences were by submitting the noise background only to the RISE algorithm fifty times (Figure 3b). Thus, there were fifty unique noise-to-noise sequences comprised of ten frames each which were designed as catch trials wherein an object did not appear.

Procedure. Participants were instructed to passively watch each image sequence. Each sequence began with a fixation point (+) followed by the presentation of either a noise-to-object sequence (50% of the trials) or a noise-to-noise sequence (50% of the trials). Each frame of the 10-frame sequence appeared for 500 ms. In the case of the noise-to-object sequences, the least coherent image appeared first for 500 ms followed by the immediate presentation of the next least coherent image for 500 ms and so on until the image reached coherence (the original image) (Figure 3a). At the end of each sequence participants were asked to type in the name of the object in order to later verify that the object was correctly recognized or indicate that there was no object by answering "none". The presentation of sequence type (noise-to-object or noise-to-noise) was random and all sequences were presented in their entirety. Participants passively viewed the sequences and typed in either an object name or "none" at the end of the image sequence.

Electrophysiology. EEG was recorded from 40 electrodes using BrainVision Recorder software (Version 1.3, Brainproducts, GmbH, Munich, Germany). An array of 35 Ag/AgCl sintered ring electrodes were fixed into a fitted cap (Easy Cap) with a standard 10-20 layout. Horizontal EOGs were recorded from electrodes placed on the outer canthi of the right and left eyes and a vertical EOG was recorded from below the right eye. Two ear-clip electrodes recorded from the right and left earlobes. The impedances of the electrodes were below 12 kΩ. All data processing was done with Brain Vision Analyzer (Brainproducts, GmbH, Munich, Germany). All recordings were recorded with an average reference and no re-referencing was performed. The signals were sampled at 250 samples/s and were digitized with 16 bit analog-to-digital converter (Brain Vision QuickAmp 72). Editing of the EEG for eye movement and blink artifacts was performed with an artifact detection algorithm (Gratton, Coles, & Donchin, 1983). Channels were also eliminated if they changed by > 50 μV between samples or exceeded ± 100 μV. EEG recordings from the 40 sites were filtered offline with a .01 – 40Hz (24 dB/oct) bandpass filter. The artifact-free EEG was then averaged for each subject across all trials within a condition. There were ten conditions with each of the successive ten frames in the noise-to-object sequence being considered as a condition. In the noise-to-object sequences, Frame 1 corresponded to the least coherent image and Frame 10 corresponded to the most coherent image. EEG was baseline corrected for the 200 ms interval prior to the presentation of each of the picture stimuli at each frame. Segments included data for the 500 ms after stimulus onset resulting in segments that were a total of 700 ms long for each condition. A final grand average was obtained by averaging across the subject averages for each of the ten experimental conditions: noise-to-object frames 1-10. All participants had at least 37 artifact free trials per condition.

Data analysis. Although Frame 7 was considered the frame of interest a priori for the ERP analyses based on the behavioural data, data from the entire 5 second sequence (Frames 1 through 10) of the noise-to-object sequence were submitted to a temporal PCA for each participant and each electrode using a Varimax rotation and no Kaiser correction (MATLAB PCA toolbox (Dien & Frishkoff, 2005; Dien, Spencer, & Donchin, 2003)). The purpose of this temporal PCA on the 5 second sequence was to identify temporally, when, within the 10 frame sequence, PCA identified an epoch that accounted for a large amount of the variance in the data. Were the temporal PCA to identify an epoch around Frame 7 this would provide independent converging evidence for considering Frame 7 as the frame of interest.

Figure 4 shows the grand averaged waveforms for all 35 channels for Frame 7. An increased negativity in posterior occipital channels and increases in positivity for central-parietal channels can be seen. The activity observed over posterior channels shows an onset and peak amplitude that is similar in character to the N170 but also shows negativity after 200ms which is more similar to an N2 or N250. The activity in central parietal regions upon visual inspection bears a resemblance to a P300. However, the exploratory nature of this research and the nature of the data, which clearly show increases in positivity and negativity that are spread out both in time and over the scalp, make visual inspection of the data an unsuitable method for analyses. Thus, data from the a priori frame of interest, the seventh frame of the noise to object sequences, were submitted to a temporospatial PCA for each participant and each electrode using a Varimax rotation and no Kaiser correction (MATLAB PCA toolbox (Dien & Frishkoff, 2005; Dien et al., 2003)).

A temporal PCA was performed before a spatial PCA because it has been proposed that temporal PCA may lead to a better initial separation of ERP components (Dien & Frishkoff, 2005). Thus, a temporal PCA was performed first on the data from 0 to 500 ms post-stimulus onset from Frame 7 with 490 observations per time point (14 participants x 35 channels). Only Frame 7 was considered because it was the a priori frame of interest. The second step of the tsPCA involved separate spatial PCAs on each of the 10 temporal scores with the data matrix having 35 channels by 14 observations (14 participants). For display purposes, the factor loadings were interpolated with Brain Vision Analyzer software (Version 1.05, Brainproducts, GmbH, Munich, Germany) using the method of spherical splines (Perrin, Pernier, Bertrand, & Echallier, 1989). All statistical analyses were ANOVAs and planned comparisons; all using Greenhouse-Geisser correction when appropriate because the homogeneity-of-covariance assumption is often violated in ERP experiments.

Results

EEG was recorded while participants viewed each of the fifty noise-to-object sequences (Figure3a) and fifty noise-to-noise sequences (Figure3b). As established in Experiment 1a, the seventh frame was considered the frame of interest a priori for the ERP study (Figure3c). The temporal PCA of the 5 second noise-to-object interval returned a solution which accounted for 93% of the total variance with 10 factors. The first temporal factor identified accounted 53% of the variance with loadings that were maximal between 3188 an 4980ms after the onset of the noise-to-object sequence with loadings > .85 where *factor loadings* are correlations between the original variables (time points) in the data and the newly reduced dimensional space (Spencer, Dien, & Donchin, 2001). Frame 7 occurs at 3000ms post sequence-onset and the results of temporal PCA on the 5 second sequence suggest that a large amount of variance within the data can be accounted for by the activity occurring from Frame 7 onward. Thus, with both the behavioural and temporal PCA data providing converging evidence that Frame 7 marks the onset of a change in perceptual activity both behaviourally, and cognitively as evidenced by the ERPs, Frame 7 was considered the frame of interest.

In order to identify both the distribution and timecourse of components that maximally accounted for variance in ERP activity at Frame 7, these data were submitted to a temporospatial principal component analysis (PCA) (Methods and MATLAB PCA toolbox (Dien & Frishkoff, 2005; Dien et al., 2003; Spencer et al., 2001)). The temporal PCA returned a solution which accounted for 97% of the total variance with 10 factors. For subsequent statistical analyses with the ERP data, only the first temporal factor between 152 and 328ms was considered here because the focus of this study is on earlier components in light of the research that objects are recognized within 150 ms (Grill-Spector & Kanwisher, 2005; Thorpe et al., 1996). The remaining eight factors accounted for less than 10% of the total variance, individually. The spatial PCA returned a solution that accounted for 99% of the total variance with 10 factors. Again, only the spatial factors found for the first temporal factor between 152 and 328ms were retained and are shown in Figure 5. For this epoch, the spatial PCA resulted in two spatial factors that accounted for 70% of the total variance combined while none of the remaining eight factors accounted for more than 10% of the total variance individually.

The initial temporal PCA resulted in a temporal factor that accounted for 46.8% of the total variance with loadings that were maximal for the epoch between 152 and 324ms after the onset of Frame 7 with all loadings > .8. To determine the spatial distribution of this factor, a spatial PCA was conducted on the factor scores found in the temporal PCA. Two primary spatial factors for the epoch between 152 and 324ms were found. The first spatial factor had a maximal loading at channel PO8 (Figure 6a) and is consistent with the timecourse and spatial distribution of the N1/N250 complex associated with encoding structural information, object perception and object familiarity (Curran, Tanaka, & Weiskopf, 2002; Schweinberger, Huddy, & Burton, 2004; Scott. L.S., Tanaka, Sheinberg, & Curran, in press; Tanaka, Curran, Porterfield, & Collins, in press). The second spatial factor had a maximal loading at channel CP2 where *factor loadings* refer to the correlation between the original variables in the data (electrodes) and this further reduced dimensional space (Figure 7a). This factor is more ambiguous and has fewer precise counterparts in the literature but it does show a similar spatial distribution and timecourse (although with an earlier onset) as the P300 elicited in classic oddball paradigms (Spencer, Dien, & Donchin, 1999; Spencer et al., 2001).

The ERP analyses were guided by the results of the tsPCA because of the exploratory nature of the research. Separate hemisphere (left, right) x frame (1-10) ANOVAs were conducted on the mean voltages for each of the spatial factors between 152 and 324ms for the noise-to-object sequences only. Thus, one ANOVA was conducted for the channels associated with the posterior-occipital factor, channels PO7 and PO8 between 152 and 324ms, and another ANOVA for the channels associated with the central parietal factor, channels CP1 and CP2, between 152 and 324ms.

The first spatial factor of the temporospatial PCA between 152 and 324ms accounting for 46.4% of the total variance was found to load maximally at channel PO8 (loading at .95) as shown in Figure 6a. Although the noise to noise sequences were not designed to be a condition of comparison, an ANOVA for this temporospatial factor showed a type x frame interaction (F (9, 117) = 27.42, P <.0001, έ = .48). Paired t-tests showed significant differences for frames 1, 7 and 8 only between the noise-to-noise and noise-to-object conditions (P <.01). At Frame 1, the noise-to-object sequences were found to be slightly more positive than the noise-to-noise images. In addition, at Frames 7 and 8, the noise-to-object sequences were significantly more negative than the noise-to-noise sequences at these same frames (Figure 6c).

The hemisphere x frame ANOVA for the noise-to-object sequences showed a main effect of hemisphere (F (1,13) = 6.4, MSE = 46.5, P <.05) where the right hemisphere was significantly more negative than the left hemisphere and a main effect of frame (F (9, 117) = 90.02, P <.0001, έ = .28). There was also a hemisphere by frame interaction (F (9,117) = 4.83, P <.0001, έ = .22) as shown in Figure 6d. Means comparisons revealed that the right hemisphere was significantly more negative than the left hemisphere for Frame 1, Frame 7 and Frame 8 (P <.05). The right hemisphere was only marginally more negative than the left hemisphere for Frame 2 (P =.06) and Frame 9 and Frame 10 (P =.05). There were no hemisphere effects for Frames 3, 4, 5 or 6 (P >.10).

With regard to the main effect of frame, the planned comparisons between sequential frames, i.e. Frame 1 compared with Frame 2, Frame 2 compared with Frame 3 etc., revealed the following: a significant difference between frames 1, 2 and 3, where Frame 1 showed increased positivity while Frame 2 showed a large negativity that had significantly decreased by Frame 3 ($P <.001$). There was a significant difference between frames 3 and 4 as Frame 4 showed activity that was not different from baseline ($P <.05$). There were no differences between frames 4 and 5, and Frames 5 and 6, which all showed activity near baseline ($P>.10$). At Frame 7, a significant increase in negativity compared with Frame 6 ($P<.001$) was followed by increased, but not significant, negativity at Frame 8 compared with Frame 7 ($P>.10$) (Figure 6b). There was a significant difference between frames 8 and 9 as the waveform began to return to activity closer to baseline ($P<.001$). The difference between Frames 9 and 10 was not significant ($P >.10$). The abrupt onset of activity in Frame 7 visually observed by looking at the ERPs at channel PO8 (Figure 6b) is thus confirmed (Figure 6d).

The second spatial factor between 152 and 324ms identified by the temporospatial PCA accounted for 24% of the total variance and had a maximal loading at channel CP2 (loading at .98) as shown in Figure 7a. An ANOVA of the second temporospatial factor of both sequence types showed a type x frame interaction (F (9, 117) = 5.78, $P <.0001$, $\acute{\varepsilon} =.61$) with paired t-tests showing significant differences for Frames 7 and 8 ($P<.01$) and Frame 9 ($P <.05$) only, between the noise-to-noise and noise-to-object conditions (Figure 7c).

The hemisphere x frame ANOVA for the noise-to-object sequences showed a main effect of frame (F (9, 117) = 13.61, $P <.0001$, $\acute{\varepsilon} = .43$). Again, the planned comparisons were between sequential frames, e.g. Frame 1 and Frame 2, Frame 2 and Frame 3 etc. The planned comparisons between Frames 1 and 2 were significant with Frame 2 being more negative than Frame 1 ($P <.01$). Critically, the planned comparisons between Frames 6 and 7 and Frames 7 and 8 were significant ($P <.01$) with Frame 7 showing an increase in polarity compared with Frame 6 and Frame 8 showed the reverse polarity showing a significant negativity relative to Frame 7. No other comparisons were significant ($P>.10$). Again, these results confirm the visual impression that Frame 7 elicits an increase in activity as shown in Figure 7b as well as revealing that this activity changes polarity in Frame 8 (Figure 7c).

Discussion

Experiment 1b was designed to investigate the neural correlates associated with object perception without invoking hypothesis-testing strategies or the influences of perceptual priming. In Experiment 1b, a continuous presentation paradigm was employed in which the neural correlates associated with object perception were investigated independent of the low level visual properties of luminance, spatial frequency and contrast. While manipulating only the phase coherence of the image, Experiment 1a identified the 7th frame as likely to evoke object perception at least 90% of the time. With this experimental procedure the relative contributions of low and high level visual processing to object perception were studied. Importantly, participants engaged in a passive viewing task in which they viewed sequences of images appear and only responded after the complete presentation of each sequence. Finally, a temporospatial PCA was performed on the critical onset frame (i.e., Frame 7) in order to identify the potential time windows and channels associated with object perception.

The results showed that a visual component similar in timecourse and topography to an N170 and an N250 showed an increase in activity at the initial onset of the image sequence (Figure 6c). The initial activity related to the onset of visual activity appeared to return to baseline levels by Frame 4 as shown in Figure 6c. The decrease in amplitude to the presentation of successive, incoherent stimuli as shown by the suppression in ERPs for frames 4 through 6 of the image sequence suggests that even though the change in low level visual information is perceptible, the lack of any new or meaningful visual information results in an early adaptation effect prior to the onset of object recognition. Such adaptation effects are consistent with adaptation effects found in vision including adaptation effects found to light in rod receptors in the human retina (Hood & Birch, 1993) and to single cell recording data from V1 in cats to changes in orientation and contrast (Carandini & Ferster, 1997; Dragoi, Sharma, & Sur, 2000). These results are consistent with data that show that not only does the visual system adapt to stimuli at the level of the retina and single cells, but that adaptation effects can be found at the level of cortical regions as demonstrated by work with repetition priming (Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002; Scott. L.S. et al., in press).

In the posterior occipital channels identified by the PCA, it was found that the onset of object perception elicited a large negativity between 152 and 324 ms. Such sensitivity to coherent visual information converges with evidence from data on the N170 which is sensitive to the effects of expertise (Gauthier, Curran, Curby, & Collins, 2003; Tanaka & Curran, 2001), individuation of faces (Jacques & Rossion, in press) and the categorical encoding of blobs (Curran et al., 2002), indicating that the N170 may be sensitive to encoding not only visual information but also coherent and possibly, meaningful, visual information.

Importantly, the N1/N2 visual component showed sensitivity to the onset of object perception in the 7th frame as evidenced by the significant increased negativity at Frame 7 shown in Figure 6c. In addition, although the noise-to-noise sequences were not designed for comparison, the results do suggest that the differences found at Frame 7 cannot be explained as an effect of expectancy because similar activity is not observed in the noise-to-noise sequences at Frame 7. This finding is consistent with the results that Vigianno and Kutas (2000), and Doniger et al. (2000) found for the identification level. However, in contrast to previous studies where the increase in negativity was incremental over successive presentations, the negative component in Experiment 1b was abrupt corresponding to the critical recognition frame. These results suggest that when changes in low level visual properties are minimized and when hypothesis-testing and priming effects are reduced, that object perception is abrupt and categorical as evidenced by the behavioural response and ERPs. The results also show a further increase in negativity for Frame 8 in the N1/N2 complex before the visual system has time to adapt a second time (Figure 6c) – this time to an object in the visual field rather than to noise. Finally, the result that the right hemisphere was more negative for this visual N1/N2 like component is consistent with other studies showing an increase in negativity over the right hemisphere during visual processing (e.g., (Gauthier et al., 2003).

The temporspatial PCA identified a second component between 152 ms and 324 ms that was maximal at channel CP2 which accounted for a fair amount of variance (24%) in the data at Frame 7. Results for this component similarly showed significant voltage change with the onset of the seventh frame as shown in Figure 7b and also showed marked activity to Frame 8. An increase in positive going activity was found for Frame 7, the onset of object coherence, while the activity in Frame 8 showed reversed polarity. This is important because it suggests that the activity seen in the central parietal component is not merely a dipole of the activity found in the posterior component in which Frame 8 increases further in negativity after the onset of object recognition at Frame 7 (Figure 7c). The comparison with the noise-to-noise sequence further suggests that the activity in CP1 and CP2 is not the result of an expectancy effect given that the same activity is not observed in the noise-to-noise sequences at Frame 7. However, this component could be a manifestation of the P300 given its spatial distribution and timecourse (Spencer et al., 2001) although the onset of this component is earlier than is typically seen for the P300 and the waveform stays positive well beyond 324 ms. It is possible that the onset of the coherent stimuli within the noise to object sequences could be construed as a novel or rare event, similar to the P300 elicited in an oddball paradigm. However, because the activity switches in polarity for the image following the onset of object perception this activity might not be a P300, suggesting that this activity is not simply a dipole of the activity in the posterior occipital channels. Further investigation is needed to test this interpretation.

Nevertheless, the possibility exists that the abrupt onset of activity for both the N1/N2 component in posterior occipital channels and the activity in CP1 and CP2 could be markers of the detection of perceptual differences. That is, these results may be the result of detecting a change in phase coherence of a certain magnitude regardless of how the information changes. To test this possibility, the same paradigm used in Experiment

1b could run in reverse such that the sequences move from a perceived object to a noise background. It would then be possible to ascertain if the differences in neural processing found at the *offset* of object perception are the result of the offset of coherent visual information, or if they result from any abrupt change in visual information.

Experiment 2 addresses this question by replicating the same procedure described in Experiment 1, but with the aim of studying the neural correlates associated with the offset of object perception. If the results of reversed presentation are the same as Experiment 1, it would show that the obtained N1/N2 component and the activity in CP1 and CP2 have little to do with recognizing objects, but are instead measures of perceptual processes related to change detection in the visual system. However, if the N1/N2 and the P300 components are significantly different, this would indicate that the results of Experiment 1 are associated with visual information related to the recognition of coherent stimuli.

Experiment 2a

Experiment 2a was conducted to find a set of stimuli with a clear categorical boundary between an object and the onset of the perception of noise. In this experiment, the phase spectrum of one hundred and sixty-eight line drawings from a new Snodgrass and Vanderwart-like set (Rossion and Pourtois, 2004) were parametrically manipulated with RISE to yield a six step object-to-noise sequence. It has been shown that the offset of object coherence occurs for an image with less object coherence than the onset of object coherence due to hysteresis effects associated with top down information (i.e. the tendency to hold onto an image in noise)(Sadr and Sinha, 2004). Thus, it was necessary to adjust the step size of the algorithm to a step size of 6 in order to elicit a clear categorical boundary between object perception and noise. Participants viewed the six frame sequences on a computer screen and indicated with a button press when there were no longer able to see the object. They then typed in the name of the object. The goal of Experiment 2a was to identify those objects in which there was a distinct categorical boundary between the object frame and the noise frame.

Method

Participants. Ten University of Victoria undergraduates participated in this study. Two were male and all participants had normal or corrected-to-normal vision and all were native English speakers. Participants received course credit for their participation.

Materials. Stimuli for the experiment consisted of 168 black and white line drawings with shading and texture information (Rossion and Pourtois, 2004), overlaid on a noise background. Again, as in Experiment 1a, the phase spectrum of each original object image was parametrically altered to produce a set of continuous sequences using Sadr and Sinha's RISE algorithm (Sadr and Sinha, 2001, 2004). However, for this study, the sequence presentation was from object-to-noise rather than noise-to-object. As in Experiment 1a, original objects were overlaid on a noise background and were 281 x 281 pixels and subtended a horizontal and vertical visual angle of approximately 8 degrees. The parameters of RISE were such that the original image had a phase coherence of 100% and the least coherent object had a phase coherence of 10%. Adjusting for a step size of 6 resulted in each successive image decreased in phase coherence by 15% as shown in Figure 8.

Procedure. Each object sequence began with a fixation point (+) followed by the presentation of the most coherent image for 500 ms followed by the immediate presentation of the next most coherent image for 500 ms and so on until the image reached incoherence (Figure 8). Participants viewed the object-to-noise sequences and indicated with a button press the instant the object disappeared and they were no longer able to see it. The button press ended the sequence of images and participants typed in the name of the object in order to later verify accuracy in naming. The stimuli presentation and data collection were accomplished using Matlab 7.0.4.

Results

The following information was computed for each object-to-noise sequence: average frame of object offset; standard deviations around the mean and accuracy in naming. The data for correctly named objects clustered predominately at and around Frame 4. Thus, the 4th frame was identified as the frame at which participants categorically distinguished noise from the object. From the total set of 169 objects, fifty were identified in which the offset was clearly at the 4th frame (S.D = .56) as shown in Figure 9c. These fifty object-to-noise sequences were used in the subsequent ERP experiment (Experiment 2b).

Discussion

It was established within Experiment 2a that for fifty objects, 89% of participants agreed that the critical frame that defined the offset of object perception for those fifty objects occurred at the 4th frame. The results of Experiment 2a established an index of object offset at the 4th frame which was used as the a priori frame of interest in the ERP study (Experiment 2b).

Experiment 2b

Experiment 2b was designed to clarify the extent to which neural activity related to the perception of object onset found in Experiment 1b is related the detection of perceptual coherence. In this experiment, the data from Experiment 2a and the results from Experiment 1b were used to guide the design and analyses of the ERP experiment.

Method

Participants. Sixteen undergraduates from the University of Victoria participated in this study. Six were male, all were right handed, all participants had normal or corrected-to-normal vision and all participants were native English speakers. The mean age was 21 years. Participants gave written informed consent after the experimental procedure was explained and were given course credit for their participation or $10. Two participants were not included in the final analyses: one participant had a history of concussions and incomplete electrophysiological data were recorded for one participant due to experimenter error.

Materials. Stimuli were fifty object-to-noise sequences identified in Experiment 2a which had a clear perceptual object offset at Frame 4. The sequences were modified as described next in order to have the same duration and same a priori frame of interest as the sequences in Experiment 1b. Therefore, each object-to-noise sequence used in Experiment 2a was padded with four images in order to create a 10 frame sequence. Three images were added at the beginning and one image at the end of the sequence so that Frame 4, identified in Experiment 2a, now occurred at the seventh frame of the sequence as in Experiment 1b. The three images added to the beginning were the original line drawings superimposed on three different versions of RISE noise. The last image generated from the original image on noise with the RISE algorithm being computed a second time. The least coherent image from this second RISE iteration was inserted as the tenth frame in the object-to-noise sequences. The resulting object-to-noise sequence can be seen in Figure 9a. In addition to the fifty object-to-noise sequences fifty unique object-to-object sequences were created to act as catch trials in which the object did not disappear. Objects not being tested in the object-to-noise condition were superimposed on the noise-to-noise sequences described in Experiment 1b and are shown in Figure 9b.

Procedure. Participants were instructed to watch each image sequence. Each sequence began with a fixation point (+) followed by the presentation of either an object-to-noise sequence (50% of the trials) or an object-to-object sequence (50% of the trials). Each frame of the 10-frame sequence appeared for 500 ms as in Experiment 1b. At the end of each sequence participants were asked to type in the name of the object in order to later verify that the object was correctly recognized. Presentation of object-to-noise and object-to-object sequences was randomized and all sequences were presented in their entirety.

Electrophysiology. The electrophysiologcal methods were identical to those described in Experiment 1b. For the object-to-noise sequences, Frame 1 corresponded to the most coherent image and Frame 10 corresponded to the least coherent image. A final grand average was obtained by averaging across each subject for each frame (1-10) of the object-to-noise sequences. All participants had at least 40 artifact free trials per condition.

Data analysis. Because this experiment was designed to clarify the results of Experiment 1b, the analyses were guided by the tsPCA results from Experiment 1b. Separate hemisphere (left, right) x frame (1-10) ANOVAs were conducted on the mean voltage for the posterior factor (channels PO7 and PO8) and the central parietal factor (channels CP1 and CP2) between 152 and 324 ms that were identified in the tsPCA for Experiment 1b. Additionally, Greenhouse-Geisser adjustments were used in all statistical analyses when appropriate.

Results

EEG was recorded while participants viewed each of the fifty object-to-noise sequences (Figure 9a) and fifty object-to-object sequences (Figure 9b). As established in experiment 2a, the seventh frame was considered the frame of interest a priori for this ERP study. Visual inspection of Figure 10a suggests that there was an effect of frame on the data. Again, even though the object-to-object sequences were not designed as a comparison condition, an ANOVA showed a type x frame interaction (F (9, 117) = 7.72, P <.001, έ =.47) where paired t-tests showed significant differences for Frames 5, 7, 8, 9 and 10 where the object-to-noise sequences were significantly more positive than the object-to-object sequences (P <.01) as shown in Figure 10b.

The visual impression of a frame effect was further confirmed by the hemisphere x frame ANOVA for the object-to-noise sequences which showed a main effect of frame (F (9,117) = 39.04, P <.0001, έ =.19) as well as a hemisphere by frame effect (F (9,117) = 3.39, P =.05, έ =.21). Means comparisons revealed that this interaction was the result of greater negativity in the right hemisphere for Frame 2 (P < .001) and Frame 5 (P <.05) only. No other comparisons were significant (P >.10).

With regard to the main effect of frame in the posterior channels, planned comparisons revealed that the significant differences between frames were between Frames 1 and 2 and Frames 2 and 3 (Figure10b). Frame 1 was significantly more positive than Frame 2 which showed a large negativity. By Frame 3, activity had returned to near-baseline levels. The only significant planned comparisons were between Frames 1 and 2 and Frames 2 and 3 (*P*<.01). Importantly, there were no differences between the frame preceding the offset (Frame 6) and the offset frame (Frame 7).

As shown in Figure 11a, the ERPs at CP2 do not appear to differ at Frame 7. There was no type x frame interaction in the ANOVA for the object-to-noise and object-to-object sequences. The results of the hemisphere x frame ANOVA for the mean voltages between 152 and 324 ms for channels CP1 and CP2 revealed a main effect of frame significance (*F* (9,117) = 5.31, *P* <.01, έ =.33). Only the planned comparisons between Frames 1 and 2 and between Frames 2 and 3 showed a significant difference (*P*<.01) with frame two showing increased positivity relative to Frames 1 and 3 as shown in Figure 11b. No other comparisons were significant thereby confirming the initial impression of the data. Once again, there was no reliable difference between the frame preceding the object offset and the object offset frame in central parietal channels.

Discussion

In the posterior occipital channels identified by the tsPCA of Frame 7 in Experiment 1b, it was found that there were no significant differences between Frames 6 and 7 related to the offset of object perception between 152 and 324 ms (Figure 9a). There was a similar pattern of activation in the first three frames of the sequence as was found for Experiment 1b where the onset of visual information resulted in increased positivity followed by a large deflection at frame two before the activity began to return to baseline at Frame 3. However, the results did show that with the offset of object perception, activity was significantly decreased and changed polarity for the N1/N2 component relative to the preceding frames as shown in Figure 10b. In addition, although the object-to-object sequences were designed as catch trials, the comparison between the object-to-noise and object-to-object sequences suggests that the pattern of activity observed for the object-to-noise sequences is not simply due to an expectancy effect since the same pattern is not observed in the object-to-object sequences. In central parietal channels identified as the second spatial component between 152 and 324 ms, there were no differences between frames for the seventh frame (Figure 10a). There was however increased positivity for Frame 2.

The results from Experiment 2b suggest that the results found in Experiment 1b are not solely the result of a salient perceptual event. The data show that the N1/N2 component is sensitive to an overall change in information from object to noise but that there is nothing particularly critical about this transition. The lack of differential activity in channels CP1 and CP2 at the offset of object perception, suggests that the increased activity in these channels in the case of object onset is not just the detection of a novel event because the onset of noise is assumed to be as equally novel in this study as the onset of object perception was in Experiment 1b.

General Discussion

This research was designed to study the contributions of low and high level vision to object perception. The goal was to separate out the contributions of high level visual processes from low level visual processes by capitalizing on the advantages of using parametric stimuli within a continuous presentation paradigm. The results of this research suggest that both components associated with visual processes, the N1/N2, and the central-parietal activity between 152 and 324 ms show differential activity with the onset of categorical object perception and that these differences are less confounded within the context of a continuous presentation paradigm. The activity in both components had a latency onset of approximately 150 ms which converges with the findings that object recognition is completed by 150 ms post stimulus onset (Grill-Spector & Kanwisher, 2005; Thorpe et al., 1996). Evidence from the second set of studies examining the offset of object perception suggest that these effects are not related to the detection of novel visual information but may be associated with the onset of coherent information. Finally, both studies on the onset of object perception and the offset of object perception showed visual adaptation effects validating the strength of using RISE and of the continuous presentation paradigm.

Other studies that have researched object completion using segmented line drawings and a discrete presentation paradigm with ERPs, have found differences in ERPs related to object completion (Doniger et al., 2000) and object identification under overt and covert conditions (Viggiano, 2000). Other studies have used degraded stimuli, similar to Richard Gregory's Dalmatian dog (Gregory, 1970) in which the picture first appears to be random black splotches before one perceives a Dalmatian dog in the picture, and ERPs to examine the neural correlates associated with the perception for faces, words and objects (Pegna, Khateb, Michel, & Landis, 2004). In contrast, this study sought to establish how the contributions of high-level vision in object perception can be studied without the noise of other task demands by controlling for low level visual information. That is, while fragmented and noisy pictures might result in hypothesis testing by participants or other task demands, the current paradigm with the "all-or-none" onset of object perception captures the “visual aha!” or the instant of object recognition across various pictures and people to tap into the feeling of “Aha! That is a dog.”

The ERP results of these studies show that event related potentials may be suited to studying this phenomenon in greater depth. The paradigm established here which specifies a clear categorical boundary a priori with behavioural data, and uses that information within the context of a continuous presentation paradigm, is especially suited to a number of lines of research in object perception. One line of research to explore is the extent to which the same activity might be seen for the coherent frame of impossible objects (Williams & Tarr, 1997); novel objects (Gauthier & Tarr, 1997; Shepard & Metzler, 1971) or some other non-object set. The results of these studies also make it possible to test questions regarding the degree to which these components are encoding the structural coherence of a stimulus; their sensitivity to perceptual boundaries (Pegna et al., 2004) and their sensitivity to perceptual learning and expertise (Jacques and Rossion, in press).

It is acknowledged that the step-size chosen within RISE plays an important role in calibrating the onset of object coherence and various interesting questions concerning the step-size and sequence of presentation of the RISE images exist in their own right. For instance, if a smaller step-size were to be used to create a more perceptible build up of information and not an abrupt onset of object perception, it is possible that N1/N2 component would show findings more consistent with Doniger et al. (2000). A further avenue of research would be to vary the order of the images leading up to the frame of object onset. If the order of images preceding object perception were randomized or changed so as not to move in a stepwise fashion from most incoherent to most coherent, would the same adaptation effects and changes in the neural signature of the onset of object recognition be seen?

Finally, this paradigm is promising as a tool for studying the development and processing of object perception across a wide range of populations. The benefits of the passive viewing task may make this type of research more accessible to developmental researchers to study object perception in children and possibly even in infants. Moreover, the task might be suited to studying object perception in patient populations as the paradigm allows for a large amount of participant control (the participant controls the onset of each image sequence) and the time required to record one hundred, 5 second image sequences is approximately twenty minutes. Given the results of this current study, it is suggested that this type of paradigm which both controls for low level visual effects and employs a continuous presentation paradigm, is a powerful way to examine the contributions of high-level vision to object perception with ERPs.

References

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A, 103*(2), 449-454.

Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognit Psychol, 23*(3), 393-419.

Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science, 144*, 424-425.

Carandini, M., & Ferster, D. (1997). A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. *Science, 276*(5314), 949-952.

Curran, T., Tanaka, J. W., & Weiskopf, D. M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cogn Affect Behav Neurosci, 2*(1), 1-18.

Dien, J., & Frishkoff, G. A. (2005). Principal components analysis of erp data. In T. C. Handy (Ed.), *Event-related potentials a methods handbook* (pp. 189-207). Cambridge, MA: MIT Press.

Dien, J., Spencer, K. M., & Donchin, E. (2003). Localization of the event-related potential novelty response as defined by principal components analysis. *Cognitive Brain Research, 17*, 637-650.

Doniger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Snodgrass, J. G., Schroeder, C. E., et al. (2000). Activation timecourse of ventral visual stream object-recognition areas: High density electrical mapping of perceptual closure processes. *J Cogn Neurosci, 12*(4), 615-621.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron, 28*(1), 287-298.

Gauthier, I. (2000). Visual priming: The ups and downs of familiarity. *Current Biology, 10*(5), 1019-1023.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nat Neurosci, 6*(4), 428-432.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Res, 37*(12), 1673-1682.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology, 55*(4), 468-484.

Gregory, R. (1970). *Intelligent eye.*New York: McGraw-Hill Book Company.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychol Sci, 16*(2), 152-160.

Hood, D., & Birch, D. (1993). Light adaptation of human rod recpetors: The leading edge of the human a-wave and models of rod recpetor activity. *Vision Research, 33*(12), 1605-1618.

Jacques, C., & Rossion, B. (in press). The speed of individual face categorization. *Psychological Science.*

Luck, S. J. (2005). *An introduction to the event related potential technique.*Cambridge, MA: The MIT Press.

Norman, J., & Ehrlich, S. (1987). Spatial frequency filtering and target identification. *Vision Res, 27*(1), 87-96.

Pegna, A. J., Khateb, A., Michel, C. M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: Where and when it occurs. *Hum Brain Mapp, 22*(4), 300-311.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol, 72*(2), 184-187.

Rossion, B., & Pourtois, G. (2004). Revisiting snodgrass and vanderwart's object set:The role of surface detail in basic-level object recognition. *Perception, 33*, 217-236.

Sadr, J., & Sinha, P. (2001). Exploring object perception with random structure evolution. *Artificial Intelligence Laboratory Memo #2001-06.*

Sadr, J., & Sinha, P. (2004). Object recognition and random image structure evolution. *Cognitive Science, 28*(2), 259-287.

Schweinberger, S. R., Huddy, V., & Burton, A. M. (2004). N250r: A face-selective brain response to stimulus repetitions. *Neuroreport, 15*, 1501-1505.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Res Cogn Brain Res, 14*(3), 398-409.

Scott. L.S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (in press). Distinguishing perceptual experience and perceptual expertise with erps. *Journal of Cognitive Neuroscience.*

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(972), 701-703.

Snodgrass, J. G., & Corwin, J. (1988). Perceptual identification thresholds for 150 fragmented pictures from the snodgrass and vanderwart picture set. *Perceptual and Motor Skills, 67*(1), 3-36.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *J Exp Psychol [Hum Learn], 6*(2), 174-215.

Spencer, K. M., Dien, J., & Donchin, E. (1999). A componential analysis of the erp elicited by novel events using a dense electrode array. *Psychophysiology, 36*, 409-414.

Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late erp responses to deviant stimuli. *Psychophysiology, 38*(2), 343-358.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychol Sci, 12*(1), 43-47.

Tanaka, J. W., Curran, T., Porterfield, A., & Collins, D. (in press). The acquisition and activation of face representations: The n250 erp as an index of face familiarity. *Journal of Cognitive Neuroscience.*

Tanaka, K. (1996). Inferiortemporal cortex and object vision. *Annual Review of Neuroscience, 19*, 109-139.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520-522.

Viggiano, M. P., and Kutas, M. (2000). Overt and covert identification of fragmented objects inferred from performance and electrophysiological measures. *Journal of Experimental Psychology:General, 129*(1), 107-125.

Williams, P., & Tarr, M. J. (1997). Structural processing and implicit memory for possible and impossible figures. *J Exp Psychol Learn Mem Cogn, 23*(6), 1344-1361.

Figure Captions

Figure 1. An object image processed with the RISE algorithm. An object from the new set of Snodgrass and Vanderwart-like objects overlaid on a noise background starting with a phase coherence of 10% (left-most image) and moving to coherence (right-most image) in 10% increments.

Figure 2. Noise-to-object sequence. An example of a noise-to-object sequence shown in the continuous presentation paradigm.

Figure 3. Continuous presentation paradigm and frame of object onset for ERP experiment. (a) An example of a noise-to-object trial and (b) a noise-to-noise trial. (c) The seventh frame was identified in behavioural sessions as the frame of object onset. The histogram represents the percent identification for the 50 objects used in the ERP study.

Figure 4. Grand average ERPs of all 35 channels for Frame 7 of the noise-to-object sequence.

Figure 5. Topographic maps of the factor loadings for the spatial factors of the temporal epoch between 152 and 324ms identified by the temporal PCA. The percentage of variance accounted for by each spatial factor is shown.

Figure 6. (a) The first spatial factor accounted for 46.4% of the total variance and had a maximal loading at channel PO8 (loading at .95). (b) Grand averages ($N = 14$) of the event-related potentials for each frame in the noise-to-object sequences for channel PO8. (c) Mean voltages for each frame between 152 and 324ms for both the noise-to-object and noise-to-noise sequences. Shown with 95% confidence intervals. (d) Mean voltages

for each frame between 152 and 324ms for the noise-to-object sequences for both left and right hemispheres. Shown with 95% confidence intervals.

Figure 7. (a) The second spatial factor for 24% of the total variance and had a maximal loading at channel CP2 (loading at .98). (b) Grand averages (*N*=14) of the event-related potentials for each frame in the noise-to-object sequences for channel CP2. (c) Mean voltages for each frame between 152 and 324ms for the noise-to-object and noise-to-noise sequences. Shown with 95% confidence intervals.

Figure 8. Object-to-noise sequence. An example of an object-to-noise sequence shown in the continuous presentation paradigm with a step size of 6.

Figure 9. Continuous presentation paradigm and frame of object offset for the ERP experiment. (a) An example of an object-to-noise sequence and (b) an object-to-object sequence. (c) The fourth frame identified as the frame of object offset was shifted to the seventh frame in the ERP study and is defined as the a priori frame of interest in the ERP study. The histogram represents the percent identification for the 50 objects used in the ERP study.

Figure 10. Mean voltages between 152 and 324ms at channels PO7 and PO8. (a) Grand averages (*N*=11) for the event-related potentials for each frame in the object-to-noise sequences for channel PO8. (b) Mean voltages for each frame between 152 and 324ms for both the object-to-noise and object-to-object sequences. Shown with 95% confidence intervals. (c) Mean voltages for each frame between 152 and 324ms for the noise-to-object sequences for both left and right hemispheres. Shown with 95% confidence intervals.

Figure 11. Mean voltages between 152 and 324ms at channels CP1 and CP2. (a) Grand averages (*N*=11) of the event-related potentials for each frame in the object-to-noise sequences for channel CP2. (b) Mean voltages for each frame between 152 and 324ms for the object-to-noise sequences. Shown with 95% confidence intervals.

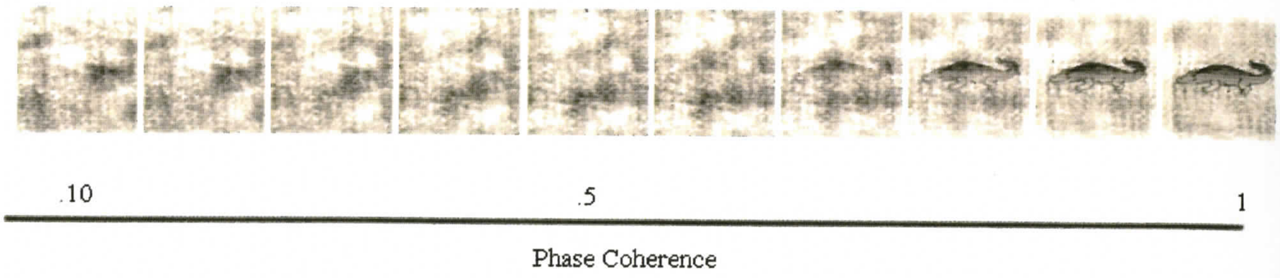

Figure 1

500ms
+
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms

Figure 2

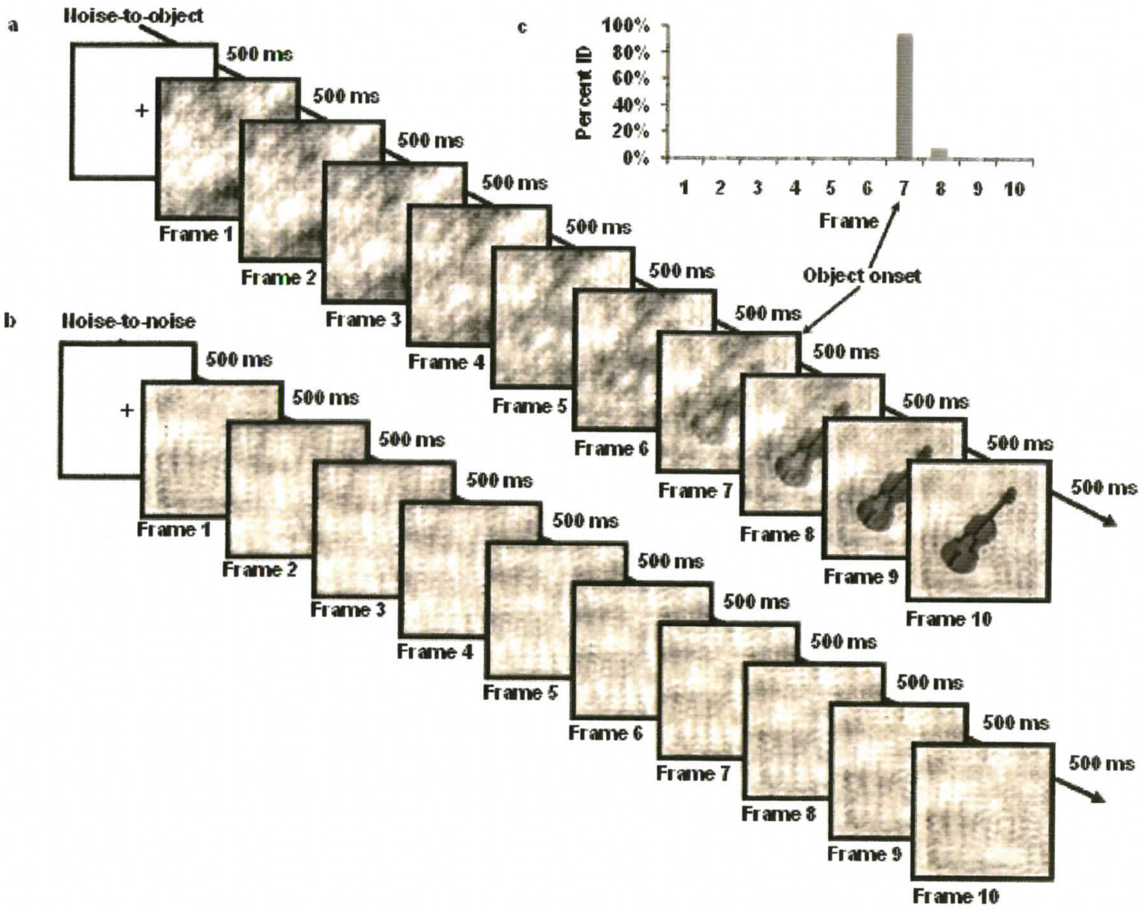
a
Noise-to-object
+
500 ms
Frame 1
Frame 2
Frame 3
Frame 4
Frame 5
Frame 6
Frame 7
Frame 8
Frame 9
Frame 10
b
Noise-to-noise
c
Percent ID
100%
80%
60%
40%
20%
0%
1
2
3
4
5
6
7
8
9
10
Frame
Object onset

Figure 3

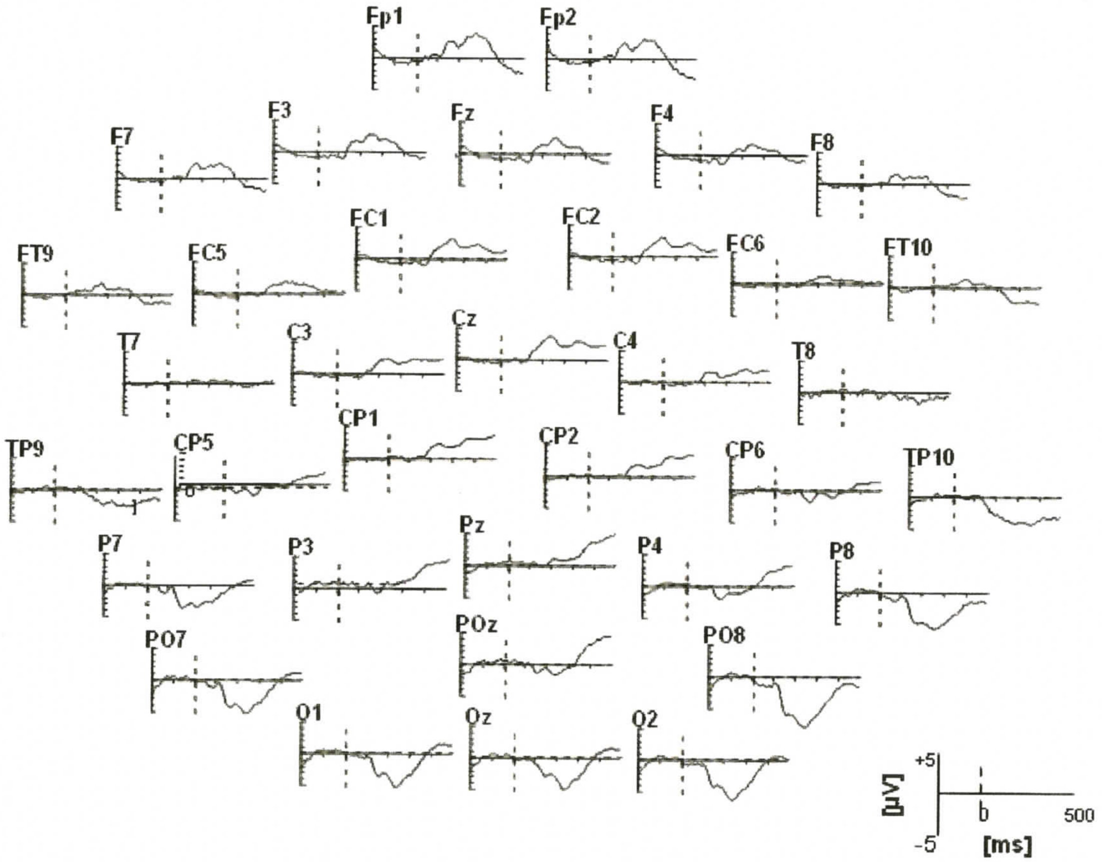
Fp1
Fp2
F7
F3
Fz
F4
F8
FT9
FC5
FC1
FC2
FC6
FT10
T7
C3
Cz
C4
T8
TP9
CP5
CP1
CP2
CP6
TP10
P7
P3
Pz
P4
P8
PO7
POz
PO8
O1
Oz
O2
+5
[µV]
-5
0
500
[ms]

Figure 4

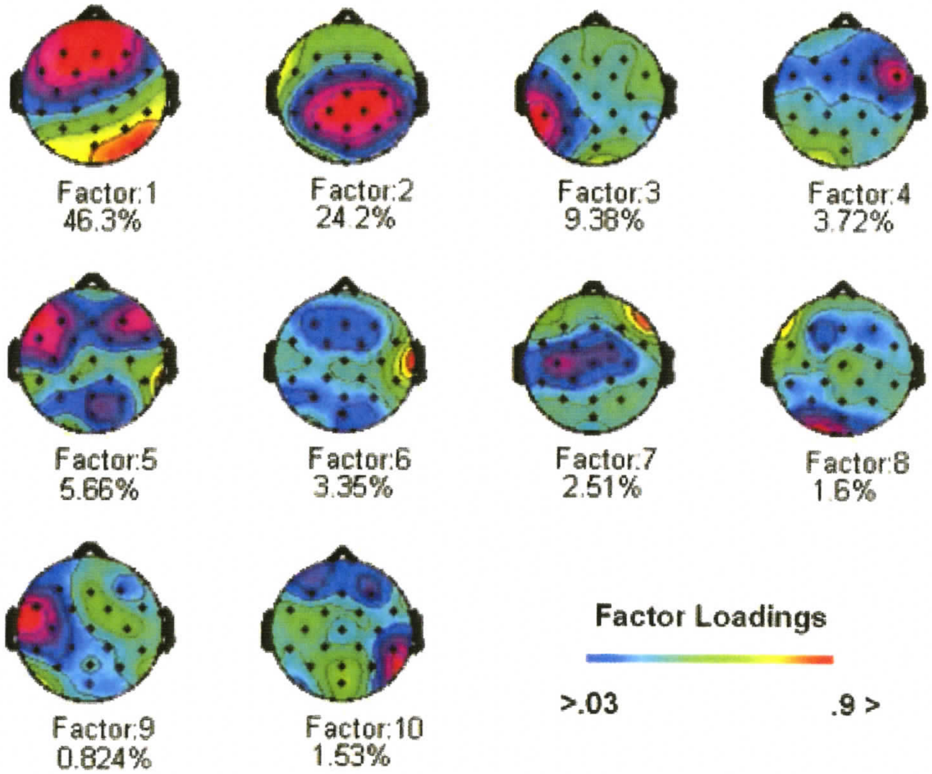
Factor:1
46.3%
Factor:2
24.2%
Factor:3
9.38%
Factor:4
3.72%
Factor:5
5.66%
Factor:6
3.35%
Factor:7
2.51%
Factor:8
1.6%
Factor:9
0.824%
Factor:10
1.53%
Factor Loadings
>.03
.9 >

Figure 5

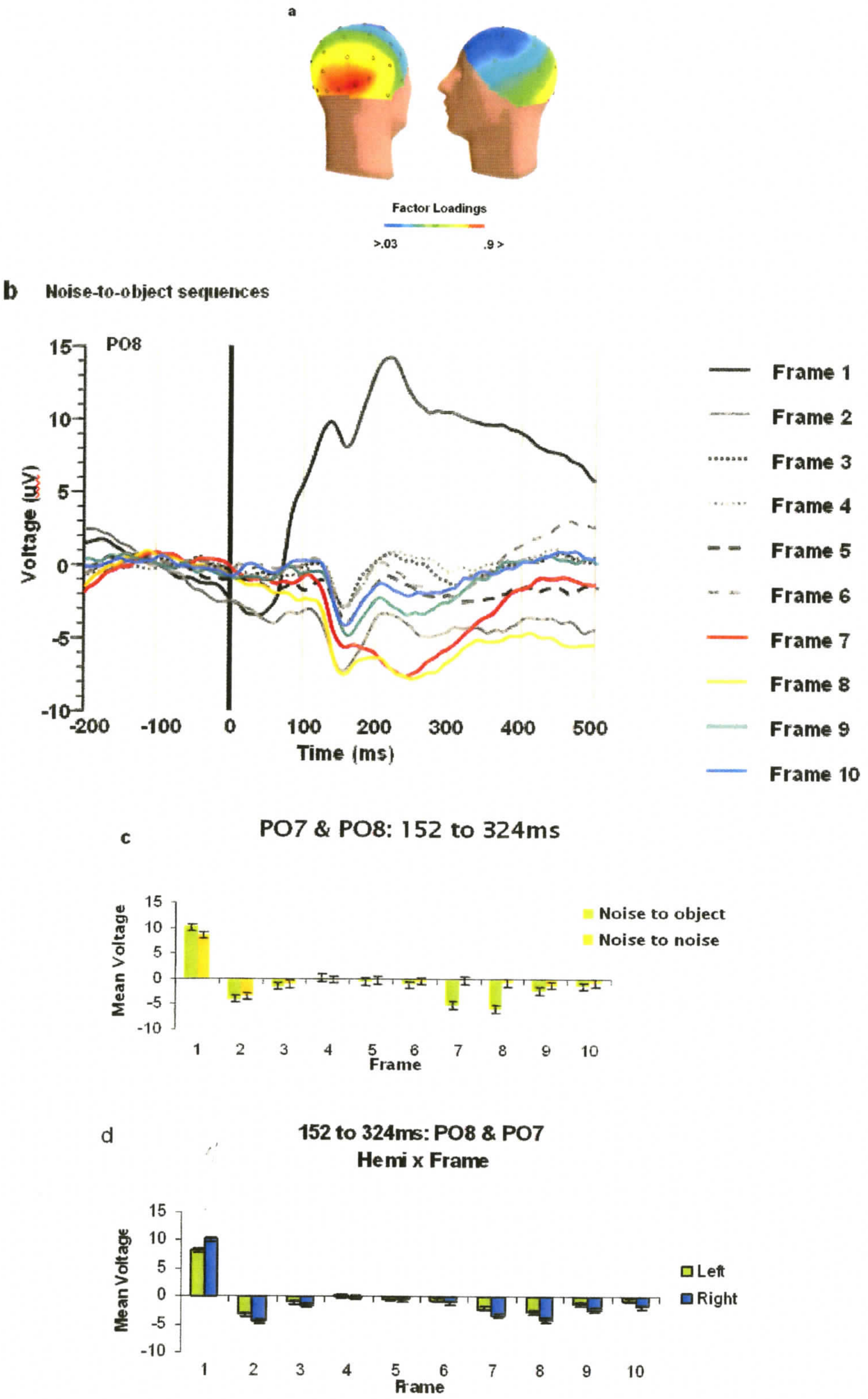
a
Factor Loadings
>.03
.9 >
b
Noise-to-object sequences
PO8
Voltage (µV)
15
10
5
0
-5
-10
-200
-100
0
100
200
300
400
500
Time (ms)
Frame 1
Frame 2
Frame 3
Frame 4
Frame 5
Frame 6
Frame 7
Frame 8
Frame 9
Frame 10
c
PO7 & PO8: 152 to 324ms
Mean Voltage
15
10
5
0
-5
-10
1
2
3
4
5
6
7
8
9
10
Frame
Noise to object
Noise to noise
d
152 to 324ms: PO8 & PO7
Hemi x Frame
Mean Voltage
15
10
5
0
-5
-10
1
2
3
4
5
6
7
8
9
10
Frame
Left
Right

Figure 6

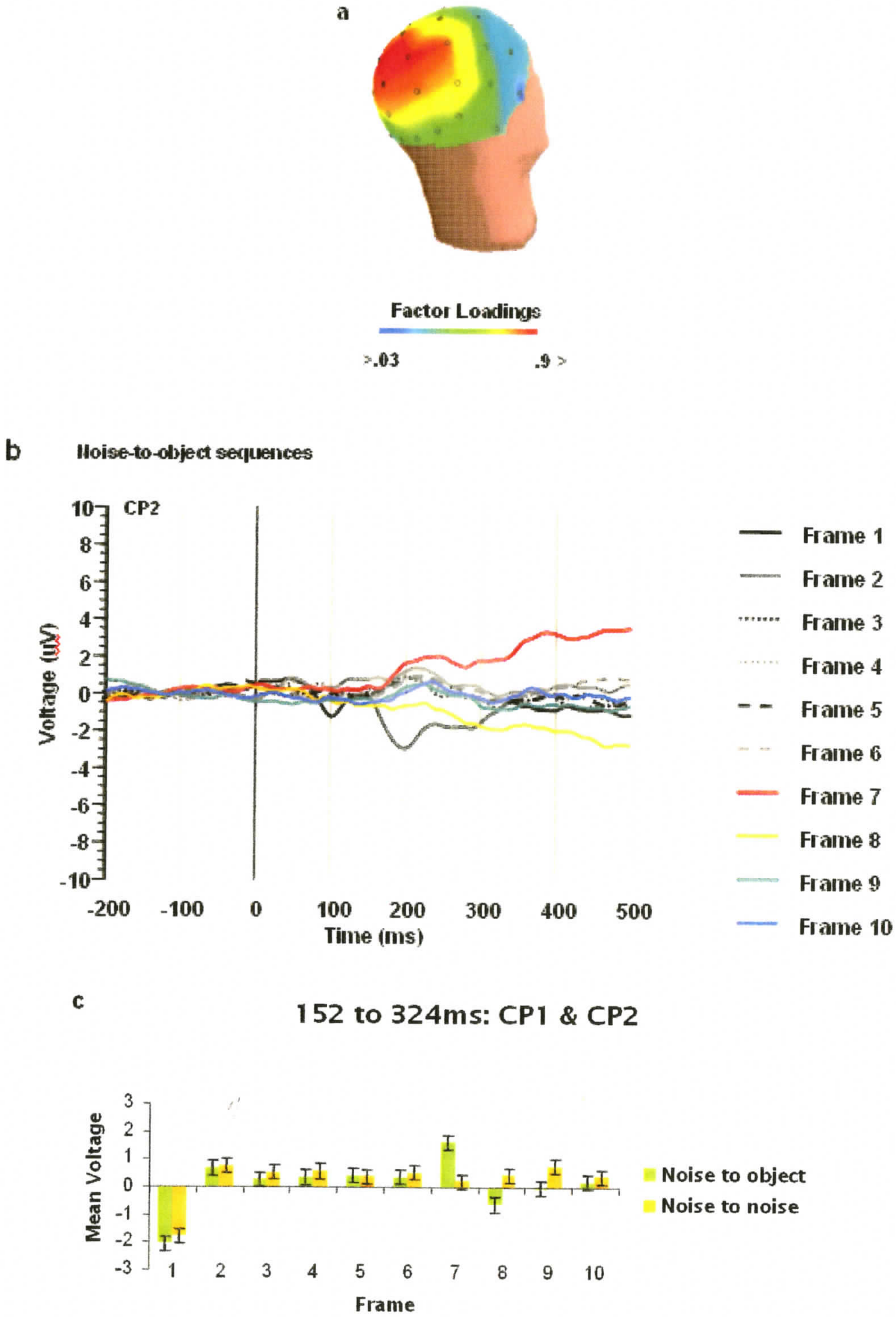
a
Factor Loadings
>.03
.9 >
b
Noise-to-object sequences
CP2
10
8
6
4
2
0
-2
-4
-6
-8
-10
Voltage (uV)
-200
-100
0
100
200
300
400
500
Time (ms)
Frame 1
Frame 2
Frame 3
Frame 4
Frame 5
Frame 6
Frame 7
Frame 8
Frame 9
Frame 10
c
152 to 324ms: CP1 & CP2
3
2
1
0
-1
-2
-3
Mean Voltage
1
2
3
4
5
6
7
8
9
10
Frame
Noise to object
Noise to noise

Figure 7

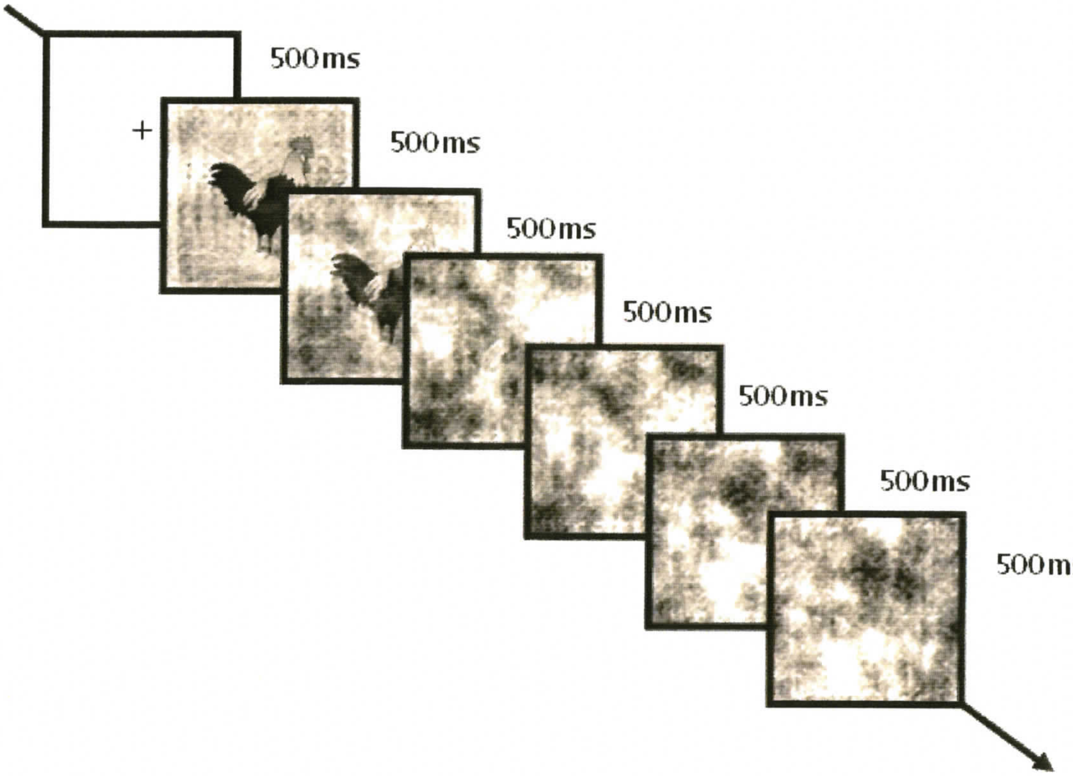

Figure 8

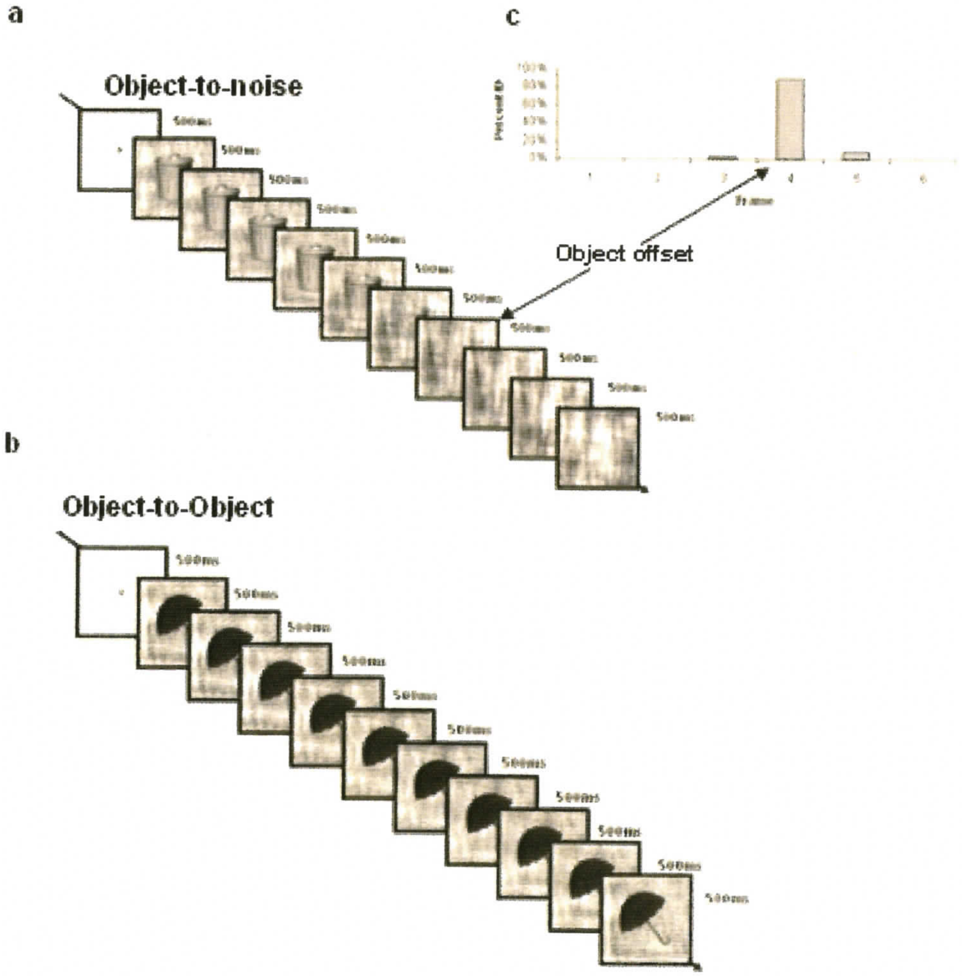
a
Object-to-noise
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
c
100%
80%
60%
40%
20%
0%
Percent ID
1
2
3
4
5
6
Frame
Object offset
b
Object-to-Object
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms
500ms

Figure 9

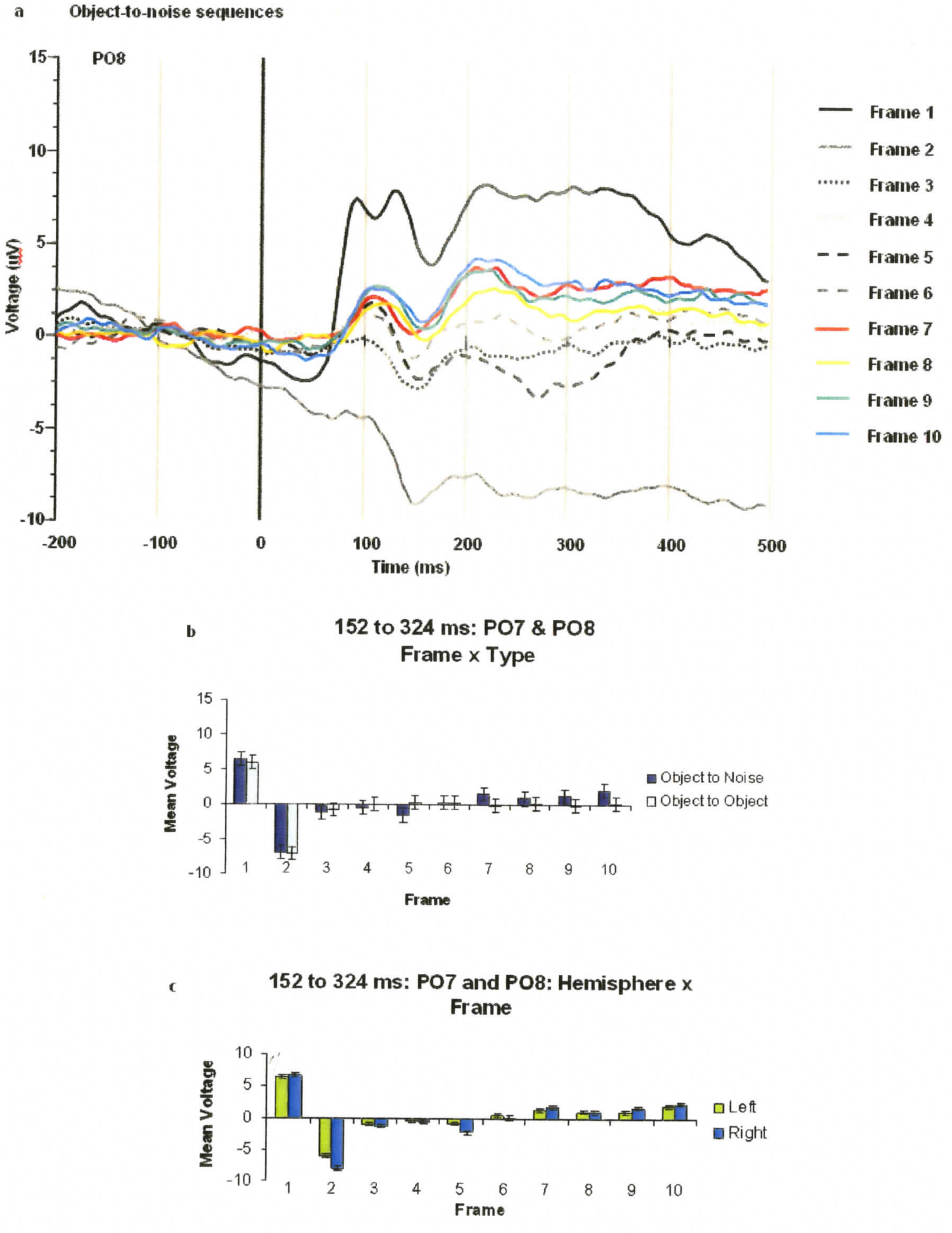
a
Object-to-noise sequences
PO8
15
10
5
0
-5
-10
Voltage (µV)
-200
-100
0
100
200
300
400
500
Time (ms)
Frame 1
Frame 2
Frame 3
Frame 4
Frame 5
Frame 6
Frame 7
Frame 8
Frame 9
Frame 10
b
152 to 324 ms: PO7 & PO8
Frame x Type
Mean Voltage
15
10
5
0
-5
-10
1
2
3
4
5
6
7
8
9
10
Frame
Object to Noise
Object to Object
c
152 to 324 ms: PO7 and PO8: Hemisphere x Frame
Mean Voltage
10
5
0
-5
-10
1
2
3
4
5
6
7
8
9
10
Frame
Left
Right

Figure 10

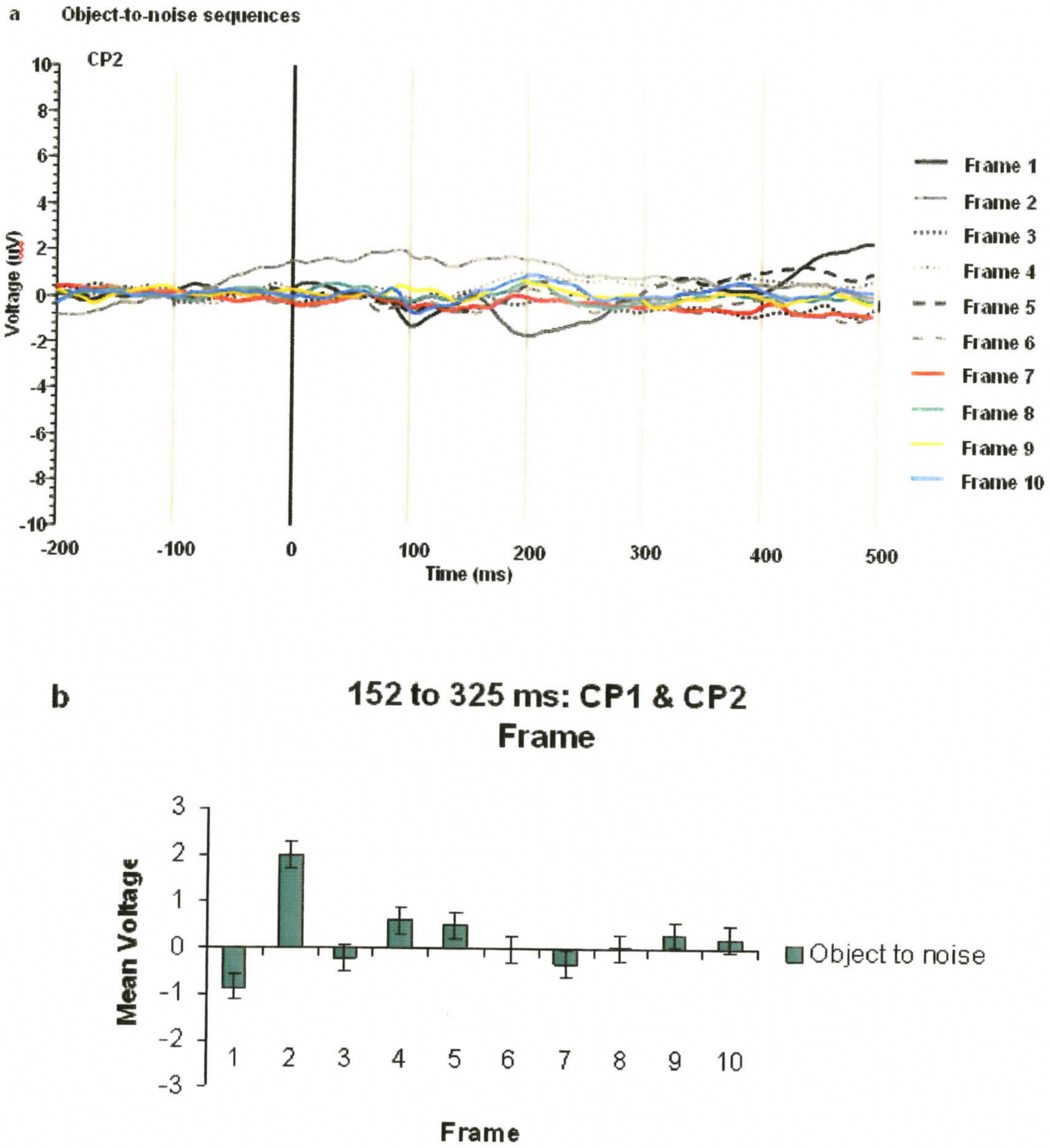
a
Object-to-noise sequences
CP2
Voltage (μV)
10
8
6
4
2
0
-2
-4
-6
-8
-10
-200
-100
0
100
200
300
400
500
Time (ms)
Frame 1
Frame 2
Frame 3
Frame 4
Frame 5
Frame 6
Frame 7
Frame 8
Frame 9
Frame 10
b
152 to 325 ms: CP1 & CP2
Frame
Mean Voltage
3
2
1
0
-1
-2
-3
1
2
3
4
5
6
7
8
9
10
Frame
Object to noise

Figure 11