Published in final edited form as:

Brain Res. 2008 October 21; 1236: 126–139. doi:10.1016/j.brainres.2008.07.104.

Parallel perceptual enhancement and hierarchic relevance evaluation in an audiovisual conjunction task

Geoffrey F. Potts¹, Susan M. Wood², Delia Kothmann³, and Laura E. Martin⁴

¹University of South Florida, Tampa, FL

²University of Houston, Houston, TX

³St. Edwards University, Austin, TX

⁴University of Kansas Medical Center, Kansas City, KS

Abstract

Attention directs limited-capacity information processing resources to a subset of available perceptual representations. The mechanisms by which attention selects task-relevant representations for preferential processing are not fully known. Treisman & Gelade's (1980) influential attention model posits that simple features are processed preattentively, in parallel, but that attention is required to serially conjoin multiple features into an object representation. Event-related potentials have provided evidence for this model showing parallel processing of perceptual features in the posterior Selection Negativity (SN) and serial, hierarchic processing of feature conjunctions in the Frontal Selection Positivity (FSP). Most prior studies have been done on conjunctions within one sensory modality while many real-world objects have multimodal features. It is not known if the same neural systems of posterior parallel processing of simple features and frontal serial processing of feature conjunctions seen within a sensory modality also operate on conjunctions between modalities. The current study used ERPs and simultaneously presented auditory and visual stimuli in three task conditions: Attend Auditory (auditory feature determines the target, visual features are irrelevant), Attend Visual (visual features relevant, auditory irrelevant), and Attend Conjunction (target defined by the co-occurrence of an auditory and a visual feature). In the Attend Conjunction condition when the auditory but not the visual feature was a target there was an SN over auditory cortex, when the visual but not auditory stimulus was a target there was an SN over visual cortex, and when both auditory and visual stimuli were targets (i.e. conjunction target) there were SNs over both auditory and visual cortex, indicating parallel processing of the simple features within each modality. In contrast, an FSP was present when either the visual only or both auditory and visual features were targets, but not when only the auditory stimulus was a target, indicating that the conjunction target determination was evaluated serially and hierarchically with visual information taking precedence. This indicates that the detection of a target defined by audio-visual conjunction is achieved via the same mechanism as within a single perceptual modality, through separate, parallel processing of the auditory and visual features and serial processing of the feature conjunction elements, rather than by evaluation of a fused multimodal percept.

Correspondence to: Geoffrey F. Potts, Department of Psychology, University of South Florida, 4202 East Fowler Ave., PCD 4118G, Tampa, FL 33620, 813-974-1085, 813-974-2492 (FAX), gpotts@cas.usf.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Multisensory; Attention; Perception; Event-related potentials

1. Introduction

The nervous system has limited information processing capacity. The effective generation of goal-directed behavior requires efficient selection of the most relevant perceptual representations for limited-capacity processing. Attention is the cognitive operation that selects information from high-capacity low computational demand parallel processing systems for access to low-capacity high computational demand processing systems. Most of the experimental studies of attention selection have examined selection based either on a single feature or the conjunction of features within a single sensory modality while most real-world objects impact multiple sensory systems and are thus have multimodal perceptual representations. Event-related potential (ERP) studies indicate that individual perceptual features are processed in parallel in posterior representational cortex and that the task relevance of these features is evaluated individually and serially in frontal cortex rather than as a fused multifeatured percept. The current study used ERPs to examine attention to targets defined by the conjunction of auditory and visual features to investigate whether multimodal task relevance is also evaluated by individual feature or by a conjoined multisensory percept.

One important aspect of attention theory is where in perceptual processing information processing capacity becomes limited. Prior to attention selection, incoming sensory information can presumably receive equal and unlimited processing, in parallel. Subsequent to attention selection, information processing becomes serialized as some perceptual information is prevented from having simultaneous access to privileged processing systems. Two influential theories of attention differ as to the stage of information flow at which attention impacts perceptual processing. Early selection theory holds that attention ‘turns up the gain’ of relevant sensory channels enhancing incoming information, (e.g. Broadbent, 1957). Late selection theory posits that attention functions on formed perceptual representations, allowing access to limited capacity systems to those percepts that meet target criteria (e.g. Deutsch and Deutsch, 1963). In Treisman & Gelade’s (1980) attention theory, the conjoining of primitive features into an object representation is the limiting operation at the boundary between high-capacity, parallel, preattentive processing and low-capacity, serial, postattentive processing. In this model, individual features are identified preattentively, in parallel, but attention is required to bind individual features together into an object representation. Behavioral evidence for this theory is provided by visual search designs in which the participants view stimulus arrays to determine the presence or absence of a target. When selection is based on a single feature the response time to a target is independent of the number of items in the array indicating parallel processing of all simple perceptual features, however, when selection is based on the conjunction of multiple features reaction time increases roughly linearly with the number of items in the array indicating that each item must be attended serially (Treisman and Gelade, 1980). However, behavioral studies cannot follow information flow through neural perceptual system to assess the impact of attention selection at each processing stage; that requires measuring neural activity during attention selection.

ERPs, with their millisecond temporal resolution, allow assessment of the impact of attention on the various stages of perceptual processing. There is a family of attention-sensitive negative voltage polarity components from approximately 150 – 350 ms post-stimulus that have been associated with selection based on both single features and feature conjunctions, including the N2, the N2b, N2pc, and the Processing and Selection Negativities (PN and SN respectively; e.g. Harter and Aine, 1984; Kenemans et al., 1992; Luck et al., 1993; Näätänen, 1992; Potts et

al., 1996; Wijers et al., 1996). The exact relationship between these differently named components has not been fully described but some relationship is apparent. For example Potts and colleagues (Potts et al., 1996; Potts and Tucker, 2001) used the term “N2b” to describe a posterior perceptual channel specific attention-related negativity that has a similar time course (onset 150 – 200 ms, peak 200 – 300 ms) and similar response properties (larger to stimuli containing target defining features) as the SN, while other authors use “N2b” to describe a channel independent central negative attention-related component (Kenemans et al., 1992; Näätänen and Gaillard, 1983; Okita et al., 1985; Wijers et al., 1989b). The bilateral SN to attended foveal stimuli in Kenemans et al. (1992) shares temporal course and psychological responsivity with the contralateral N2pc to attended lateralized stimuli described by Luck et al. (1993). Interpretive differences about what these components index do exist among researchers; for example Luck et al (1997) describe the N2pc as indexing selection of the target item while Smid et al (1999) describe the SN as indexing post-selection processing. We do not currently know to what extent the different component names reflect indices of the same cognitive process under different eliciting conditions, leading to small differences in latency and scalp topography, or reflect indices of different cognitive operations that overlap in time and may share some of the same neural resources. For consistency with most prior ERP studies of feature conjunction, the current report will use the term “SN” to refer to the posterior negative ERP index of attention modulation.

For individual features to be processed in parallel, each perceptual channel may need its own neural processing resources. Harter and Aine’s (1984) neural specificity theory proposed that attention operates on separate neural populations encoding specific stimulus features (e.g. orientation and color) and that attention to a feature enhances the activity of the neurons coding that feature via a ‘gain’ modulation of those neurons making them more responsive to incoming sensory information (early selection). This enhanced neural response is reflected in the ERP as an SN recorded at the scalp area overlying the neural population responsible for encoding the attended feature. Neural specificity has been extended beyond the two perceptual channels originally proposed by Harter & Aine (1984) by demonstrating differences in the scalp distribution of the SN with attention to multiple perceptual features including location, spatial frequency, shape, orientation, and color (e.g. (Kenemans et al., 1993; O’Donnell et al., 1997; Potts and Tucker, 2001), although specificity is not always found, e.g. Smid et al. (1999) found that shape and color selection produced SNs with the same scalp distribution. Hillyard & Mangun (1986) argued that one could not support neural specificity with ERP data due to the low spatial resolution and source localization ambiguity inherent in the method. However, results from hemodynamic imaging studies, which have unambiguous high spatial resolution, have been consistent with neural specificity theory. For example, Corbetta et al. (1990) found that when participants scanned a display for change in a specific feature (shape, color, or movement), different parts of prestriate cortex were activated depending upon the attended feature. Results from dense-sensor array ERP recording, which improves spatial localization, have also been consistent with neural specificity (O’Donnell et al., 1997; Potts and Tucker, 2001). A study combining electrophysiological (both ERP and magnetic evoked potentials) and hemodynamic (functional MRI) methods directly related the blood flow changes in prestriate areas associated with attention to color or motion to topographically distinct selection negativities, further supporting parallel processing of individual perceptual features in distinct prestriate areas (Schoenfeld et al., 2007). However, specificity theory does not attempt to explain how information in specific perceptual channels is selected as relevant given the individual’s goals in the current task. Some theories of neural organization make a distinction between the representative functions of posterior cortex and the executive functions of frontal cortex (Luria, 1973; Pribram, 1973) and the identification of task relevant representations may require those frontal executive functions (Faw, 2003; Posner & Petersen, 1990).

There is a medial prefrontal positivity at approximately the same latency as the SN variously called the Frontal Selection Positivity (FSP: Kenemans et al., 1993), anterior P2 (P2a: Potts et al., 1996; Potts and Tucker, 2001), frontal P3 (P3f: Makeig et al., 1999), and the frontal polar component (FP: Guillem et al., 2001). As with the posterior negativity the relationship among these components is not fully clear, however they also share temporal course and psychological responsivity. While some authors report onset latency (e.g. Kenemans et al., 1993; Smid et al., 1999) and others report peak latency (e.g. Makeig et al., 1999; Potts et al., 1996), comparing the reported latencies and waveform plots, these differently named components all appear to have an onset between about 150 – 200 ms and peak at about 240 – 320 ms, all have a medial frontal scalp distribution, and all are larger to stimuli that possess at least some of the target-defining features, suggesting that these different terms refer to the same or similar ERP components. This positivity has been described as indexing “selection of items containing some relevant features for further processing” (Kenemans et al., 1993), “detection of a task-relevant stimulus” (Potts et al., 1996), and “suppression of interfering items to increase salience of the task-relevant item” (Guillem et al., 2001). All these descriptions imply a stimulus-related evaluation operation, containing the concept of ‘task-relevance’, the value of the stimulus given the task, rather than a response-related function. However, Kenemans et al., (1993) also associated the FSP with response-related functions and Makeig et al. (1999) found that the P3f was earlier in individuals who made faster responses and therefore interpreted the P3f as indexing “activity in brain systems associated with speeded manual responding” indicating that the frontal positivity was motor response production related, not stimulus evaluation related. Potts (2004) reported a frontal positivity (P2a) both on tasks requiring manual (keypress) and covert (silent count) responses indicating that an overt, manual response was not required to elicit the frontal positivity. It remains possible that the frontal positivity could index a more general-purpose response-related function that could include both overt and covert responding.

While the scalp location of the SN appears dependent upon the perceptual channel of the stimulus, the FSP appears to be channel independent (Kenemans et al., 1993; Potts et al., 1998; Potts and Tucker, 2001). Thus the SN and FSP may reflect interaction between modality specific posterior systems of perceptual representation and a modality independent frontal system of relevance evaluation (Potts et al., 1996; Potts and Tucker, 2001). However, the exact nature of this interaction is not known; the SN and FSP could index the bottom-up flow of information where a percept is formed in posterior representational cortex and the relevance of that percept is then evaluated in frontal cortex or they could reflect a top-down modulation in which a frontal executive control system provides the biasing input that enhances the posterior representation. The temporal relationship between the SN and FSP might help clarify the functional relationship, however the temporal order of the components varies by report, e.g. Kenemans et al. (1993) reported an SN onset latency of 220 – 230 ms with the FSP preceding the SN by about 60 ms implying a top-down modulation, while Potts & Tucker (2000) reported an SN peak latency of approximately 260 ms in an object shape selection task with the SN preceding the FSP by about 30 ms implying a bottom-up flow of information. Additionally, temporal order of ERP components cannot fully disambiguate direction of information flow since neural information processing employs both feed-forward and reentrant projections (Di Lollo et al., 2000). In either case, if the FSP and SN index the point in processing where attention selection takes place, then perceptual processing should switch from a parallel to a serial mode at that point.

Since in the Treisman & Gelade model (1980) feature conjunction requires serial attention, ERPs can be used to study the transition from parallel to serial processing in selection tasks where a feature conjunction defines the target. As noted by Hansen & Hillyard (1983) and Kenemans et al. (1993) ERP selection effects can reveal three types of processing including a distinction between serial and parallel processing: if the ERP effects are additive, i.e. if the

ERP effect to selection based on feature 1 alone plus the effect to attention to feature 2 alone sum to the ERP effect of selection based on the conjunction of features 1 and 2 then this is evidence of parallel independent processing of the two features; if the ERP effects are contingent, i.e. the effect to selection on feature 1 can occur with or without the presence of feature 2 but the effect to selection on feature 2 requires the presence of feature 1 then this is evidence of serial hierarchic processing; finally if the effects are interactive, i.e. the effect to the conjunction of features 1 and 2 is different than the additive effect of the features presented alone, then this is evidence of unique processing of a conjoined feature percept distinct from the processing of the individual features alone. For example, in a study by Hillyard & Münte (1984) stimuli were targets only if they appeared at the correct location and were the correct color. When a stimulus appeared at or near the target location there was enhanced negativity in the ERP compared to stimuli far from the target location and additional negativity if the stimulus was also the target color, but when a stimulus appeared distant from the target location there were no ERP effects due to color. This contingent pattern (effect due to color was contingent upon location) indicated that the selection system employed a hierarchic evaluation process, evaluating location first and then evaluating color if and only if the location evaluation returned 'target' (Hillyard and Münte, 1984). In a similar design using visual grating stimuli where targets were defined by the conjunction of spatial frequency and orientation, Kenemans et al. (1993) reported that an SN was present if either the orientation or spatial frequency were in the target-defining set (additive effects) but that an FSP was present if the stimulus was the correct frequency but not the correct orientation but not present to the correct orientation when the presented frequency was not in the target set (contingent effect), indicating that the SN indexed parallel, independent processing of frequency and orientation in posterior cortex but the FSP indexed serial hierarchic processing in frontal cortex with spatial frequency taking precedence over orientation.

Wijers and colleagues (Wijers et al., 1989a; Wijers et al., 1989b; reviewed in Wijers et al., 1996) have examined parallel versus serial processing in target detection using ERPs in a design where participants had to respond to a subset of the letters in a stimulus set but only if the target letter was presented in the correct color and the correct size. In this design an FSP was present when the stimulus was color-relevant but not size-relevant with additional FSP enhancement when the stimulus was both color and size relevant but than an FSP was not present when the stimulus was only size-relevant, indicating serial, hierarchic processing with color taking precedence (Wijers et al., 1989a). More recently, Smid et al. (1999) reported that the SN in a similar design using color and global and local shape showed initially additive effects of color and shape followed (in some conditions) by interactive effects (depending on discriminability of the features), but that the FSN showed hierarchic effects, indicating that the SN initially indexed parallel processing of the perceptual dimensions while the FSP indexed serial processing of the individual features. Note that in these studies the color and size or global shape features did not fully define a target: in the (Wijers and colleagues 1989a, 1989b) studies letter identity with the final target defining feature while in Smid et al. (1999) local shape was. Color and size or global shape signaled that the stimulus could be a target; for a stimulus to be a target there had to be a conjunction of three features: correct color, correct size or global shape, and correct letter identity or local shape.

Most ERP studies of attention to feature conjunctions have been unimodal, they examined attention effects within or between perceptual channels in a single sensory modality, e.g. color and orientation in the visual modality. However, in the real world we frequently encounter objects that impact multiple sensory systems and identification of a task-relevant item may require conjoining information from more than one sensory modality. Some studies have used ERPs to study when in perceptual processing multimodal integration might take place. In one common design stimuli are presented unimodally in each modality under investigation (e.g. auditory alone and visual alone) and the ERPs from each modality summed (A+V) and that

summed ERP compared to the ERP from the stimuli presented simultaneously (AV; e.g. Giard and Peronnet, 1999; Gondan and Röder, 2006; Talsma et al., 2007). The points at which the AV waveform diverge from the A+B waveform index perceptual processing unique to multimodal perception, showing when in processing perceptual integration takes place (Giard and Peronnet, 1999). Adding an attentional manipulation to this design can indicate how attention impacts the integration of multimodal perceptual information (Talsma et al., 2007). Studies using this method have found evidence of integrated multimodal processing very early in the perceptual processing stream, as early as 80 ms post-stimulus, at the earliest stages of cortical percept formation (Gondan and Röder, 2006), and that attention impacts these earliest stages of integrated bimodal perceptual processing (Talsma et al., 2007).

ERP studies of multimodal attention have largely examined the impact of spatial attention across perceptual modalities, examining how spatial cues in one modality impact the processing of stimuli in another modality in nearby (usually the same side) versus more distant (usually the opposite side) locations (e.g. Eimer, 1996; Green and McDonald, 2006; Teder-Salejari et al., 2005). For example, a visual cue to the left visual field enhances the ERP to auditory stimuli on the left side and an auditory cue on the left side enhances the ERP to visual stimuli in the left visual field (Eimer et al., 2002). These effects are more apparent in the ERP indices of early sensory processing (e.g. P1 and N1) than in the later ERP components thought to index processing on formed perceptual representations (Eimer et al., 2002).

A few studies have used ERPs to study audio-visual attention to non-spatial features, e.g. object shape and/or tone pitch (Talsma et al., 2007; Talsma and Kok, 2001). One study examined the impact of attention on audio-visual processing by presenting visual square wave gratings and auditory sine wave tones either unimodally or bimodally in unattended, attend unimodal auditory or visual, and attend bimodal (both auditory and visual) conditions (Talsma et al., 2007). Targets were defined by perceptual feature (transient intensity change) in either the auditory or visual modality. Results showed both early (P1) and late (SN) effects of attention such that both the posterior visual P1 and SN were larger when attending either to visual targets alone or to both auditory and visual modalities than when attending to the auditory modality alone or ignoring both streams in favor of another (non-AV) visual stream, interpreted as indexing modulation of both early visual sensory and late visual perceptual processing when attending to the visual stream (Talsma et al., 2007). Note that this design examined the impact of redundant multimodal information rather than multimodal feature conjunction. When attending to both auditory and visual streams, all the simultaneous audio-visual targets were both auditory and visual targets, there were never any audio-visual stimuli that were visual, but not auditory targets or visa versa, and in any case, a target in either modality was sufficient to require a response – the conjunction of an auditory and a visual target was not required.

A study by Fort, et al., (2002) did contain a condition that required integration of tone and shape features for identification of a target. In that study different subsets of the participants showed early (approximately 150 – 200 ms) frontal and posterior effects related to feature audiovisual feature integration for target detection. A subset of the participants (6 of 23) who were designated as perceptually biased to the auditory modality, based on faster responses in an auditory compared to a visual target detection task, showed a posterior integration effect (an ERP spatiotemporal pattern unique to the bimodal target) between about 150 – 200 ms, while visually dominant participants (17 of 23) showed a frontal integration effect at the same latency. These early integration effects were interpreted as indexing either sensory channel specific enhancement in the individual's non-dominant modality (i.e. enhancement of the visual signal in the auditory dominant participants) or a nonspecific signal that a multimodal object was present, rather than the identification of a specific multimodal object (and by extension, the selection of that object for preferential processing). These posterior and frontal effects were topographically distinct from the SN and FSP (both effects were right lateralized;

Fort, et al., 2002) and thus likely did not reflect selection related processes associated with those components.

The current study examined whether attention election using information from two sensory modalities used assessment of the features in each modality separately or relied on evaluation of a multimodal percept. The study employed a multimodal (audiovisual) conjunction task using simultaneously presented auditory tones (high and low pitch) and visual symbols (checkmark and integral sign) in random and equiprobable combinations and defined the target, in different blocks, unimodally in each modality and as the conjunction of auditory and visual features. In one task block one of the symbols was the target (Attend Visual, auditory stimuli irrelevant), in another block one of the tones was the target (Attend Auditory, visual irrelevant), and in another block the conjunction of an auditory and visual stimulus was the target (Attend Conjunction, both modalities relevant, e.g. checkmark and low pitch tone). If, as in the visual modality, the SN indexes the parallel enhancement of individual perceptual representations when they are relevant to the participant's task, then when an audio-visual stimulus contains only the task-relevant auditory feature the SN should be present at superior frontotemporal recording sites over auditory cortex, when an audio-visual stimulus contains only the relevant visual feature the SN should be present at inferotemporal sites over visual cortex, and when the audio-visual object contains both auditory and visual task-relevant features there should be SNs over both frontotemporal and inferotemporal sites. If the FSP indexes a serial evaluation of the relevance of the auditory and visual features then it should show a contingent pattern, enhanced when the target feature is present in the modality that takes precedence regardless of the stimulus in the subordinate modality but not present in the subordinate modality if the primary modality stimulus is not a target. In contrast, if the SN indexes the formation of a multimodal percept and the FSP reflects the evaluation of that conjoined audiovisual percept, then it should show an interactive pattern: no (or reduced) SN and FSP when only one of the two features is a target but a present or enhanced SN and FSP when both the auditory and visual stimuli are targets indicating that the stimulus requires a response.

2. Results

2.1 Behavioral

Accuracy was high in all task conditions but with significant differences between the conditions evidenced by a main effect for accuracy, $F(2, 38) = 5.5$, $p = .008$, with subjects more accurate in the Attend Visual task (mean = 99.2%, $SD = 1.3$) than in the Attend Auditory (mean = 97.6%, $SD = 2.3$, $t(38) = 2.7$, $p = .010$) or Attend Conjunction (mean = 97.5%, $SD = 2.9$, $t(38) = 2.5$, $p = .018$) tasks. Accuracy between the Auditory and Conjunction conditions was not significantly different. The target reaction time was also significantly different between the task conditions, $F(2, 38) = 28.4$, $p < .001$, slower in the Conjunction condition (mean = 517 ms, $SD = 90.5$) than in the Auditory (mean = 453, $SD = 103.5$, $t(38) = 2.09$, $p = .043$) or Visual (mean = 433 ms, $SD = 79.6$, $t(38) = 3.13$, $p = .003$) conditions (the Auditory and Visual tasks were not significantly different from each other).

2.2 ERP

To reduce the dimensionality of the electrode factor, region-of-interest (ROI) electrode groupings were created for the medial prefrontal FSP, the frontotemporal SN, and the inferotemporal SN, and an additional centroparietal ROI to examine potential P3b effects, one of each ROI in the right and left hemispheres to test for laterality effects (see Figure 1), and the waveforms averaged across the electrodes within each ROI. To help distinguish the relatively small attention effects from the larger stimulus-driven deflections in the combined audio-visual ERP waveforms, difference waves were created by subtracting the nontarget waveform from the target waveform in both the attend Auditory and attend Visual conditions.

In the Conjunction condition difference waves were created by subtracting the conjunction nontarget (neither auditory nor visual stimulus is a target) waveform from the conjunction auditory target waveform (auditory but not visual stimulus is a target; this subtraction result hereafter referred to as A+V-), the conjunction nontarget from conjunction visual target (auditory nontarget, visual target; A-V+), and the conjunction nontarget from conjunction target (A+V+) waveforms.

Tests for component amplitude differences between conditions were done using repeated-measures ANOVAs for each component. The FSP and P3b ANOVAs had as factors Hemisphere (Right, Left), and Condition (Auditory, Visual, Conjunction A+V-, Conjunction A-V+, Conjunction A+V+), and the SN ANOVA had the above factors and an additional ROI factor (Frontotemporal, Inferotemporal) to test for topographic differences between auditory and visual selection. Reported p-values for effects on or interactions with a factor containing more than two levels were corrected for deviations from sphericity using the conservative Geisser-Greenhouse epsilon. We also tested the peak latency of the FSP positive peak and the SN negative peaks at both the frontotemporal and inferotemporal ROIs in a single ANOVA with Condition and Hemisphere as factors with an additional Component (FSP, frontocentral SN, inferotemporal SN) factor. Note that in all task conditions the stimuli remained the same, simultaneous audio-visual presentation of tone and shape, only the target-defining feature(s) changed. The unsubtracted waveforms and difference waves from the four ROIs are shown in Figure 2 & Figure 3 with the analysis windows delineated.

To illustrate the SN and FSP scalp field distribution across all the electrodes in the recording array we created topographic maps of the scalp voltage fields of the difference waves in the five subtractions: Auditory attention, Visual attention, A+V-, A-V+, and the Conjunction targets A+V+. We extracted the mean voltage within the 220 – 320 ms SN/FSP analysis window at each electrode and used the EEGLab package “topoplot” function to created interpolated topographic maps of the voltage fields (Delorme and Makeig, 2004; see Figure 4).

2.2.1 FSP amplitude (220 – 320 ms, prefrontal ROI)—There was a main effect of condition on the FSP, $F(4, 76) = 11.1$, $\epsilon = .71$, $p < .001$, showing that the FSP was largest (most positive) whenever a visual target was present, i.e. in the Attend Visual condition and when a visual target was present in the Attend Conjunction condition regardless of the auditory stimulus (A-V+ and A+V+), (see Table 1 for individual t-test results and Figure 5 for ERP component mean amplitude histograms).

2.2.2 SN amplitude (220 – 320 ms, frontotemporal and inferotemporal ROIs)—The SN amplitude was significantly different between the conditions, $F(4, 76) = 10.79$, $\epsilon = .73$, $p < .001$, with the greatest negativity to the conjunction targets (A+V+), intermediate amplitude when either one of the auditory or visual stimuli were target features in the conjunction task (A+V- and A-V+), and smallest in the Attend Auditory and Attend Visual conditions. Condition interacted with ROI, $F(4, 76) = 18.61$, $\epsilon = .57$, $p < .001$, indicating that the SN at the frontotemporal and inferotemporal ROIs differed in amplitude depending on which target-defining feature was present. The SN at the frontotemporal ROI was more negative in the Attend Auditory than in the Attend Visual condition and most negative when an auditory target was present in the conjunction task regardless of the visual stimulus (A+V+ and A+V-). The SN at the inferior occipitotemporal ROI was largest in the Attend Visual condition and when only the visual target was present in the conjunction condition (A-V+), next largest to the conjunction targets (A+V+), and smallest in the Attend Auditory condition and when only the auditory target was present in the conjunction condition (A+V-), (see Table 1 and Figure 5; note the difference between A+V+ and A+V- was only marginally significant). There was also an effect of hemisphere, $F(1, 19) = 8.10$, $p = .010$, indicating that the SN was

more negative over the left hemisphere, and condition interacted with hemisphere, $F(4, 76) = 3.8$, $\epsilon = .81$, $p = .013$, suggesting that in all conditions where a visual target was present (all condition except Attend Auditory and A+V-) the SN was larger over the left hemisphere (individual contrasts not explicitly tested).

2.2.3 P3b amplitude (320 – 480 ms, centroparietal ROI)—The P3b differed among the conditions, $F(4, 76) = 3.15$, $\epsilon = .77$, $p = .031$, with the largest P3b in the Attend Visual and the Conjunction Target (A+V+) conditions, although only the contrasts with the Attend Auditory condition reached significance (the P3b in the Attend Visual condition was marginally larger than the A+V- and A-V+ conditions).

2.2.4 FSP/SN peak latency—There was no effect of Component on the latency of the FSP and frontal and posterior SNs indicating that there was no latency difference between the positive FSP peak and either the frontotemporal or inferotemporal negative SN peaks. There was a small effect of Condition, $F(4, 19) = 3.04$, $\epsilon = .73$, $p = .038$, indicating the earliest peak latency in the Attend Auditory task, (mean = 271 ms, SD = 26.8) and the latest in the Conjunction condition when only the visual stimulus contained the target feature (A-V+: mean = 284 ms, SD = 24.8; individual contrasts not tested). There were trends for Condition $\times$ Hemisphere ($p = .095$) and Condition $\times$ Hemisphere $\times$ Component ($p = .061$) but these were not further explored.

3. Discussion

It was more difficult for participants to determine if a stimulus was a target defined by the conjunction of features in the auditory and visual modalities than to detect a target in a single perceptual modality, as evidenced by the slower reaction times in the conjunction condition (about 65 – 85 seconds slower on average). Accuracy was also slightly lower in the conjunction condition than in the Attend Auditory, but not the Attend Visual, condition, but the difference was small compared to the overall high accuracy rate (~99% to ~97.5%).

In the ERP analysis of the critical conjunction condition, when participants were to respond if and only if the presented audio-visual stimulus contained both the auditory target feature (pitch) and the visual target feature (shape), when the audiovisual stimulus contained only the auditory but not the visual target feature (A+V-), there was an SN in the frontotemporal ROI, but not the inferotemporal ROI, consistent with auditory, but not visual, perceptual enhancement. When the audiovisual stimulus contained the visual, but not the auditory target feature (A-V+), there was an SN in the inferiortemporal ROI but not frontotemporal ROI consistent with visual, but not auditory, perceptual enhancement. When the audiovisual stimulus contained both the auditory and visual target features (A+V+), i.e. it was a target requiring a response in the conjunction task, there were SNs over both frontotemporal and interior occipitotemporal sites consistent with enhancement of both the auditory and visual perceptual representations (see Figure 4 and Figure 5). This SN pattern is additive in the spatial sense (rather than the more commonly employed temporal sense) and indicates that the auditory and visual features of the stimuli were processed separately, in parallel, in distinct neural areas.

The FSP did not show an interactive pattern in the conjunction condition, i.e. present only when both the auditory and visual features were targets. Rather the FSP was present when the stimulus contained the visual target feature with or without the auditory target (although there was no additional FSP enhancement from the addition of an auditory target feature to the visual target feature) but was absent when the stimulus only contained the auditory target feature. This is more consistent with a serial pattern with the visual features taking precedence over the auditory features. This SN-additive/FSP-contingent pattern of results in bimodal

conjunction evaluation is similar to that found in prior studies of unimodal conjunction target detection.

In visual studies of targets defined by the conjunction of spatial frequency and orientation (Kenemans et al., 1993), and color and shape (Smid et al., 1997), the posterior SN was present to stimuli that contained a target-defining feature in either perceptual dimension, spatial frequency or orientation, color or shape. Similarly, in the current study an SN was present if either the auditory or the visual stimulus was a target and that SN was in different scalp locations consistent with parallel processing of the individual perceptual features in specific neural areas. In contrast, in the earlier unimodal studies if the stimulus contained only one of the target defining features the FSP presence depended upon which feature it was; an FSP was present if only the spatial frequency was a target but not if only orientation was a target but there was an additional FSP enhancement if both frequency and orientation were targets (Kenemans et al., 1993). Similarly, an FSP was present for the color target feature alone but not shape alone but was additionally enhanced if both color and shape were target features (Smid et al., 1997). The FSP pattern in the current data is similar, but not identical, showing an FSP to the visual, but not the auditory, target feature. Thus the frontal evaluation in the current data is driven by the visual information.

The SN and FSP results in the current data are not consistent with a fused audiovisual percept at attention selection. While prior ERP studies indicate that audiovisual integration can take place very early in perceptual processing and that attention can influence that early integration (Gondan and Röder, 2006; Talsma et al., 2007), when a target is defined by the conjunction of auditory and visual features the current data indicates that the auditory and visual perceptual representations are selected separately, in parallel, in the distinct neural areas responsible for creating those auditory and visual percepts. The frontal assessment of overall stimulus relevance to the task is evaluated serially, dominated by the visual modality.

Since the FSP and the SNs exhibited different patterns of results on the conditions, they represent distinct ERP entities (as opposed to opposite poles of the same component). However, since the FSP and frontotemporal and inferotemporal SNs all had the same peak latency the current data do not indicate whether the SN/FSP represents bottom-up or top-down neural interactions between posterior areas of perceptual representation and a frontal evaluative system.

In the visual unimodal attention condition, when the target in the audio-visual stimuli was defined by visual feature alone, there was an SN over inferior temporal leads accompanied by a medial prefrontal FSP to visual targets consistent with coactivation of the ventral visual object feature “what” pathway (SN) and medial frontal cortex (FSP). This spatial field distribution is similar to that described in Potts et al. (1996), Potts and Tucker (2001) and Kenemans et al. (1993) and is consistent with the interpretation of interaction between a frontal system of relevance evaluation and a posterior system of perceptual representation when the visual feature is the sole target-defining feature (Potts et al., 1996; Potts and Tucker, 2001).

When attending to the auditory features alone there were no attention-related effects apparent in the ERP; the target-minus-nontarget difference wave in the Attend Auditory condition was essentially flat. This lack of effect could be in part because the auditory condition was easier, requiring less resource allocation. However, since there was no reaction time difference between the auditory and visual unimodal attention conditions, and only a very small accuracy difference, other factors likely contribute. One possible contribution is the relative frequency of the targets and nontargets in the unimodal attention conditions. Since there were only two auditory and two visual stimuli and the stimuli within each modality were presented equiprobably, the unimodal target probability was .5 (compared to the conjunction target

probability of .25). The impact of target frequency on the P3 is well documented: within an attended stream of auditory stimuli the P3 is largest to the less frequent stimulus (Duncan-Johnson and Donchin, 1977; Squires et al., 1977; Squires et al., 1976). With the equiprobable distribution of the targets and nontargets in the auditory unimodal condition, no P3 would be expected. While the probability distribution of the targets and nontargets was the same in the Attend Visual condition as in the Attend Auditory condition, a P3 attention effect was apparent in the Attend Visual condition. While the visual P3, like the auditory P3, is larger to less frequent stimuli, unlike the auditory P3 and visual P3 is elicited to instructed visual targets that are equiprobable (or even more frequent) than nontargets (Squires et al., 1977).

Similar frequency effects may also differentially impact the earlier FSP and SN attention effects that were also absent in the auditory unimodal attention condition but present in the visual unimodal condition. The auditory system may be more sensitive to stimulus frequency than the visual system, reflected in the ERP. The pre-attentive auditory mismatch negativity (MMN) is an ERP component specifically sensitive to stimulus frequency and does not require attention (i.e. the MMN can be of the same amplitude to deviant stimuli that occur either in an attended or unattended stream) (Näätänen, 1992; Näätänen et al., 1978). While the auditory MMN is well established (see Näätänen, 1992 for a review) there is limited evidence for the existence of a visual MMN (but see Heslenfeld, 2003). The visual MMN, if present, may be later and require a larger feature discrepancy between the frequent and deviant stimuli for elicitation than the auditory MMN (Pazo-Alvarez et al., 2003; Tales et al., 1999). This suggests that an early, automatic, stimulus feature based novelty detection mechanism in the visual system is less sensitive than the one in the auditory system indexed by the MMN.

The current study indicated that selection of a task relevant multimodal stimulus for preferential processing is based on the individual evaluation of the component features in each modality rather than by an evaluation of a multimodal percept. However, the experimental design could have contributed to this result. In the current study, each audio-visual item did not consist of a unique set of perceptual features, i.e. each tone was paired with each visual stimulus so that all possible pairings of the two tones and two shapes were presented. This may have contributed to a dissociation between the auditory and visual features, rather than to a fusing, potentially influencing the way the features were processed. Further studies including unique audio-visual stimuli, rather than just a task that is dependent on a specific conjunction within all possible pairings of visual and auditory features, may help clarify whether target determination is based solely on distinct separate feature processing in separate modality specific areas of cortex or in some conditions by forming and evaluating a multimodal audio-visual percept. Additionally, a design containing stimulus probability as a manipulation will help determine if the auditory attention effects to audio-visual stimuli are dependent upon the relative frequency of occurrence of the stimulus in the same manner as seen in unimodal presentations.

4. Experimental Procedures

4.1 Participants

Participants ($n = 20$) were recruited from ads in local newspapers in Houston, Texas. All participants were male, 16 were right-handed, and ranged in age from 24 – 53 (mean = 39.2, SD = 8.76; note: data from the attend auditory and attend visual conditions but not the attend conjunction condition, for eight of the participants has been previously presented as part of the control group in a study of attention disruption in schizophrenia (Wood et al., 2006)). Participants were screened for neurological insult or illness, alcohol or drug abuse in the last 5 years or lifetime history of addiction, and for alcohol or psychoactive drug use 24 hours prior to testing. All participants exhibited a desire to participate as evidenced by giving their written informed consent.

4.2 Stimuli and Tasks

Visual stimuli were a checkmark or an integral symbol appearing at the center of a CRT computer monitor. Auditory stimuli were high (1.5 kHz) or low (1.0 kHz) pitch tones presented over speakers placed to the right and left sides of the monitor. On each trial, one visual and one auditory stimulus were selected randomly and equiprobably and presented simultaneously for 100 ms with a variable inter-trial interval of 1000 – 1250 ms. There were three task-blocks of 320 trials each. In the visual task block subjects were instructed to respond to one of the two visual stimuli, in the auditory block to one of the two tones, and to ignore all stimuli in the modality that did not contain the target. In the conjunction block subjects were to respond to a specific combination of the visual and auditory stimuli (e.g. the checkmark paired with the high-pitched tone). The target stimuli and task order of the auditory and visual unimodal attention blocks were counterbalanced across subjects; the conjunction task was always that final task block. Participants were instructed to make a keypress to the targets using the index finger of their dominant hand. Note that the stimuli were the same in all the tasks, only the instructed task-relevant feature(s) were different.

4.3 EEG Recording and Signal Processing

The EEG were recorded with a 128-channel Geodesic Sensor Net system (EGI, Eugene, OR), vertex referenced, analog filtered from .1 – 100 Hz, and digitized at 250 Hz. The EEG was digitally filtered offline at 20 Hz lowpass, segmented into epochs spanning 200 ms before to 800 ms after combined audio-visual stimulus onset, the segments screened for eye blinks or eye movements or other non-cephalic artifact, and the artifact-free segments averaged to created ERPs for the following conditions: attend auditory nontargets and targets, attend visual nontargets and targets, and attend conjunction nontargets (neither auditory or visual stimulus is a target), conjunction task auditory targets (the tone is the target tone but the symbol is not), conjunction condition visual target (the symbol is the target symbol but the tone is not the target tone), and conjunction target (both symbol and tone are targets; note that this is the only stimulus combination requiring a response in the conjunction task). The ERPs were baseline corrected using the 200 ms prestimulus period and rereferenced into an average reference representation.

4.4 Data Analysis

Reaction times and percent correct for the targets in the auditory, visual, and conjunction conditions were analyzed using repeated-measures ANOVAs. Differences between the individual means were tested with two-tailed t-tests.

Dependent measures for the ERP were the mean amplitudes of the difference waves within temporal windows and ROI electrode groupings (see Figure 1 for the ROI electrode groupings). The temporal window for the FSP and SN was 220 – 320 ms and the P3b window was 320 – 480 ms (see Figure 2 and Figure 3 for the temporal windows). The FSP was extracted from the prefrontal ROI, the SN from the frontotemporal and inferotemporal ROIs, and the P3b from the centroparietal ROI. The mean amplitudes from these windows and across these ROIs were cast into three repeated-measures ANOVAs, one for the prefrontal FSP, one for the SN at both the frontotemporal and inferior occipitotemporal ROIs, and one for the P3b at the centroparietal ROI. Each ANOVA contained as factors Hemisphere (Right, Left), and Condition (Auditory, Visual, Conjunction A+V-, Conjunction A-V+, Conjunction A+V+), with the addition of an ROI factor for the N2b (Frontotemporal, Inferior Occipitotemporal). Individual contrasts critical to the hypotheses within significant ANOVA results were tested with t-tests. In addition, the peak latency of the most positive peak in prefrontal ROI and the most negative peaks in the frontotemporal and inferotemporal ROIs between 220 – 320 ms was extracted and tested by repeated-measures ANOVA with Condition and Hemisphere as factors and an

additional Component factor (FSP, Frontotemporal SN, Inferotemporal SN) to examine the temporal relationships between the SN and FSP.

Acknowledgements

The authors thank Estate Schokhadze for his assistance in data collection. Support provided by NIH R01-DA14073 (Potts)

References

- Broadbent DE. A mathematical model for human attention and immediate memory. *Psychol Rev* 1957;64:205–215. [PubMed: 13441856]
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE. Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 1990;22:1556–1559. [PubMed: 2360050]
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Meth* 2004;134:9–21.
- Deutsch JA, Deutsch D. Attention: Some theoretical considerations. *Psychol Rev* 1963;70:80–90. [PubMed: 14027390]
- Di Lollo V, Enns JT, Rensink RA. Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J Exp Psychol Gen* 2000;129:481–507.
- Duncan-Johnson CC, Donchin E. On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology* 1977;14:456–467. [PubMed: 905483]
- Eimer M. The N2pc component as an indicator of attentional selectivity. *Electroencephalogr Clin Neurophysiol* 1996;99:225–234. [PubMed: 8862112]
- Eimer M, van Velzen J, Driver J. Cross-Modal Interactions between Audition, Touch, and Vision in Endogenous Spatial Attention: ERP Evidence on Preparatory States and Sensory Modulations. *J Cognit Neurosci* 2002;14:254–271. [PubMed: 11970790]
- Faw B. Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Conscious Cognit* 2003;12:83–139. [PubMed: 12617864]
- Fort A, Delpuech C, Pernier J, Giard MH. Early auditory-visual interactions in human cortex during nonredundant target identification. *Cognit Brain Res* 2002;14:20–30.
- Giard MH, Peronnet F. Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *J Cognit Neurosci* 1999;11:473–490. [PubMed: 10511637]
- Gondan M, Röder B. A new method for detecting interactions between the senses in event-related potentials. *Brain Res* 2006;1073:389–397. [PubMed: 16427613]
- Green JJ, McDonald JJ. An event-related potential study of supramodal attentional control and crossmodal attention effects. *Psychophysiology* 2006;43:161–171. [PubMed: 16712586]
- Guillem F, Bicu M, Hooper R, Bloom D, Wolf MA, Messier J, Desautels R, Debruille JB. Memory impairment in schizophrenia: A study using event-related potentials in implicit and explicit tasks. *Psychiatr Res* 2001;104:157–173.
- Hansen JC, Hillyard SA. Selective attention to multi-dimensional auditory stimuli. *J Exp Psychol Hum Percept Perform* 1983;9:1–19. [PubMed: 6220115]
- Harter, MR.; Aine, CJ. Brain mechanisms of visual selective attention. In: Parasuraman, R.; Davies, DR., editors. *Varieties of attention*. London: Academic Press; 1984. p. 293–321.
- Heslenfeld, DJ. Visual Mismatch Negativity. In: Polich, J., editor. *Detection of Change: Event-related Potential and fMRI Findings*. Berlin: Springer; 2003. p. 41–60.
- Hillyard SA, Mangun GR. The neural basis of visual selective attention: A commentary on Harter and Aine. *Biol Psychol* 1986;23:265–279. [PubMed: 3828423]
- Hillyard SA, Münte TF. Selective attention to color and location: An analysis with event-related brain potentials. *Percept psychophysics* 1984;36:185–198.

- Kenemans JL, Kok A, Smulders FT. Event-related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirements. *Electroencephalogr Clin Neurophysiol* 1993;88:51–63. [PubMed: 7681391]
- Kenemans JL, Verbaten MN, Melis CJ, Slangen JL. Visual stimulus change and the orienting reaction: Event-related potential evidence for a two-stage process. *Biol Psychol* 1992;33:97–114. [PubMed: 1525300]
- Luck SJ, Fan S, Hillyard SA. Attention-related modulation of sensory-evoked brain activity in a visual search task. *J Cognit Neurosci* 1993;5:188–195.
- Luck SJ, Girelli M, McDermott MT, Ford MA. Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology* 1997;33:64–87. [PubMed: 9212722]
- Luria, AR. The frontal lobes and the regulation of behavior. In: Pribram, KH.; Luria, AR., editors. *Psychophysiology of the frontal lobes*. New York: Academic Press; 1973. p. 3-26.
- Makeig S, Westerfield M, Jung TP, Covington J, Townsend J, Sejnowski TJ, Courchesne E. Functionally independent components of the late positive event-related potential during visual spatial attention. *J Neurosci* 1999;19:2665–2680. [PubMed: 10087080]
- Näätänen, R. *Attention and Brain Function*. Erlbaum; Hillsdale, New Jersey: 1992.
- Näätänen, R.; Gaillard, AWK. The orienting reflex and the N2 deflection of the event-related potential (ERP). In: Gaillard, AWK.; Ritter, W., editors. *Tutorials in ERP research: Endogenous components*. Amsterdam: North-Holland; 1983. p. 119-141.
- Näätänen R, Gaillard AWK, Mäntysalo S. Early selective attention effect on evoked potential reinterpreted. *Acta Psychol* 1978;42:313–329.
- O'Donnell BF, Swearer JM, Smith LT, Hokama H, McCarley RW. A topographic study of ERPs elicited by visual feature discrimination. *Brain Topogr* 1997;10:133–143. [PubMed: 9455604]
- Okita T, Wijers AA, Mulder G, Mulder LJM. Memory search and visual spatial attention: An event-related brain potential analysis. *Acta Psychol* 1985;60:263–292.
- Pazo-Alvarez P, Cadaveira F, Amenedo E. MMN in the visual modality: a review. *Biol Psychol* 2003;63:199–236. [PubMed: 12853168]
- Posner MI, Petersen SE. The attention systems of the human brain. *Annu Rev Neurosci* 1990;13:25–42. [PubMed: 2183676]
- Potts GF. An ERP index of task relevance evaluation of visual stimuli. *Brain Cognit* 2004;56:5–13. [PubMed: 15380870]
- Potts GF, Dien J, Hartry-Speiser AL, McDougal LM, Tucker DM. Dense sensor array topography of the event-related potential to task-relevant auditory stimuli. *Electroencephalogr Clin Neurophysiol* 1998;106:444–456. [PubMed: 9680158]
- Potts GF, Liotti M, Tucker DM, Posner MI. Frontal and inferior temporal cortical activity in visual target detection: Evidence from high spatially sampled event-related potentials. *Biol Psychol* 1996;9:3–14.
- Potts GF, Tucker DM. Frontal evaluation and posterior representation in target detection. *Cognit Brain Res* 2001;11:147–156.
- Pribram, KH. The primate frontal cortex - executive of the brain. In: Pribram, KH.; Luria, AR., editors. *Psychophysiology of the Frontal Lobes*. New York: Academic Press; 1973. p. 293-314.
- Schoenfeld MA, Hopf JM, Martinez A, Mai HM, Sattler C, Gasde A, Heinze HJ, Hillyard SA. Spatio-temporal Analysis of Feature-Based Attention. *Cereb Cortex* 2007;17:2468–2477. [PubMed: 17204821]
- Smid HGOM, Jakob A, Heinze HJ. The organization of multidimensional selection on the basis of color and shape: An event-related brain potential study. *Percept Psychophysics* 1997;59:693–713.
- Smid HGOM, Jakob A, Heinze HJ. An event-related brain potential study of visual selective attention to conjunctions of color and shape. *Psychophysiology* 1999;36:264–279. [PubMed: 10194973]
- Squires KC, Petuchowski S, Wickens C, Donchin E. The effects of stimulus sequence on event related potentials: A comparison of visual and auditory sequences. *Percept Psychophysics* 1977;22:31–40.
- Squires KC, Donchin E, Herning RI, McCarthy G. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science* 1976;193:1142–1146. [PubMed: 959831]

- Tales A, Newton P, Troscianko T, Butler S. Mismatch negativity in the visual modality. *Neuroreport* 1999;10:3363–3367. [PubMed: 10599846]
- Talsma D, Doty TJ, Woldorff MG. Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? *Cereb Cortex* 2007;17:679–690. [PubMed: 16707740]
- Talsma D, Kok A. Nonspatial intermodal selective attention is mediated by sensory brain areas: Evidence from event-related potentials. *Psychophysiology* 2001;38:736–751. [PubMed: 11577897]
- Teder-Salejarvi WA, Di Russo F, McDonald JJ, Hillyard SA. Effects of Spatial Congruity on Audio-Visual Multimodal Integration. *J Cognit Neurosci* 2005;17:1396–1409. [PubMed: 16197693]
- Triesman A, Gelade G. A feature integration theory of attention. *Cognit Psychol* 1980;12:97–136. [PubMed: 7351125]
- Wijers, AA.; Mulder, G.; Gunter, TC.; Smid, HGOM. Brain potential analysis of selective attention. In: Neumann, O.; Sanders, AF., editors. *Handbook of perception and action*. Vol. Vol 3. Tullamore, Ireland: Academic Press; 1996. p. 333-387. Attention
- Wijers AA, Mulder G, Okita T, Mulder LJ. Event-related potentials during memory search and selective attention to letter size and conjunctions of letter size and color. *Psychophysiology* 1989a;26:529–547. [PubMed: 2616701]
- Wijers AA, Mulder G, Okita T, Mulder LJ, Scheffers MK. Attention to color: an analysis of selection, controlled search, and motor activation, using event-related potentials. *Psychophysiology* 1989b; 26:89–109. [PubMed: 2922460]
- Wood SM, Potts GF, Hall J, Ulanday J, Netsiri C. Event-related potentials to auditory and visual selective attention in schizophrenia. *Int J Psychophysiol* 2006;60:67–75. [PubMed: 16009438]

Figure 1.

Map of the 128 channel Geodesic Sensor Net with 10/20 system locations noted for reference and the Region-Of-Interest electrode groupings for the Prefrontal (FSP: circles), Frontotemporal (auditory SN: squares), Inferotemporal (visual SN: diamonds), and Centroparietal (P3b: triangles) marked (note that electrodes 60 and 86 are in both the Inferotemporal and Centroparietal ROIs).

Figure 2. Unsubtracted waveforms from the Prefrontal, Frontotemporal, Inferotemporal, and Centroparietal ROIs in the unimodal auditory and visual attention (left column) and conjunction (right column) conditions with the analysis windows delineated.

Figure 3.

Difference waveforms from the Prefrontal, Frontotemporal, Inferotemporal, and Centroparietal ROIs in the Attend Auditory and Attend Visual conditions (left column: target minus nontarget) and Attend Conjunction condition (right column: auditory target feature minus nontarget, visual target feature minus nontarget, conjunction target minus nontarget).

Scalp Field Topography Averaged Across the SN/FSP Window (220 - 320 ms)

Figure 4.

Scalp field topographies averaged across the SN/FSP window (220 – 320 ms) of the target minus nontarget difference (see text for details of the subtractions) from the attend auditory only (Auditory), attend visual only (Visual), and the attend conjunction conditions when only the auditory stimulus was a target (A+V–), when only the visual stimulus was a target (A–V+) and when both stimuli were targets making a conjunction target (A+V+).

Mean Voltage in the SN/FSP Window (220 - 320 ms)

Figure 5.

Histograms of the mean voltage (bars are standard error) in the SN/FSP analysis windows at the prefrontal, frontotemporal, and inferotemporal ROIs.

Table 1

T-test results for individual contrasts

	FSP				Frontotemporal SN				Inferotemporal SN				P3b			
	vis	A+V-	A-V+	A+V+	vis	A+V-	A-V+	A+V+	vis	A+V-	A-V+	A+V+	vis	A+V-	A-V+	A+V+
aud	***	ns	***	**	**	***	***	**	***	ns	***	^	*	ns	ns	*
vis		**	ns	ns		***	^	***		***	ns	ns		^	^	ns
A+V-			***	*			***	ns				^			ns	ns
A-V+				ns				***				*				ns

* p < .05
** p < .01
*** p < .001
^ .05 < p < .1

Brain Res. Author manuscript; available in PMC 2009 October 21.