Auditory Attentional Capture:

Effects of an Abrupt Unexpected Sound

By

Leslie M. Thorne

University of Manitoba

Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

in the

Department of Psychology

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Auditory Attentional Capture:

Effects of an Abrupt Unexpected Sound

BY

Leslie M. Thorne

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Master Of Arts

Abstract

Attentional capture is said to occur when an abrupt, unexpected event draws attention to itself, without intention. While attentional capture by a unique yet irrelevant distractor has been well-studied in the visual modality, less is known about how irrelevant singletons attract attention in the auditory domain. The current research investigated whether an abrupt change in frequency (frequency violation) can capture attention in an involuntary manner (Experiment 1). The logic was that if the frequency violation captures attention, then perception of a target of a different pitch should be impaired when it is presented immediately following the violation. Experiment 2 evaluated whether auditory attention can be captured by an abrupt location change. Finally, Experiment 3 explored the possibility that when a violation has a different pitch and comes from a different location, it more effectively captures attention than does a violation on a single dimension. Experiments 1 and 2 revealed an effect of attentional capture. In Experiment 3, attentional capture was observed, however a violation defined along two dimensions did not capture attention more efficiently than a violation defined along a single dimension. While attentional capture in vision appears to be contingent on attentional set, the results of the current experiments demonstrate that this is not the case with auditory attentional capture. Although the violation was irrelevant to the task, attention was still drawn to it and performance on the target detection task was impaired. The present study show that voluntary orienting is not always successful and that target detection can be impaired by the occurrence of an unexpected sound defined by frequency, location, or both.

Table of Contents

List of Tables

List of Figures

Introduction

Voluntary control of attention is evident in many of our daily activities. In the visual domain, voluntary control includes looking for a lost item, searching for friends in a crowd, watching television, or reading a newspaper. In each case, our attention is consciously focused toward a specific item or location. In the auditory domain, voluntary control can include following a conversation or listening to music, again with attention being consciously directed. In contrast, automatic or involuntary control of attention can occur when there is an unintentional shift of attention. For example, a flash of light or an object that suddenly moves can draw attention towards itself. In audition, a sound such as an alarm, a siren, or any sudden noise can capture attention.

Dichotic Listening Studies

Much of the early research on selective attention used auditory stimuli, which paralleled situations that are commonly encountered in everyday living. An experiment by Cherry (1953) demonstrated that attention could be voluntarily oriented on the basis of spatial location. Cherry presented listeners with two auditory messages presented dichotically (one to each ear). Listeners were instructed to shadow one of the messages (repeat it aloud) and ignore the other message. Cherry observed that subjects found the task fairly easy. He also observed that when subjects were asked to describe the contents of the rejected message, they were not able to say much about it. Thus, in the unattended message, when the speech started out in English, listeners were unable to detect when the message switched to German. In addition, they were not able to notice when English speech was played backwards in the rejected message. However, listeners did notice when the speaker switched gender (pitch) and when speech was replaced with a 400 Hz tone. This early research demonstrated that, when instructed to do so, listeners can easily attend to one location and effectively ignore irrelevant information (e.g. Cherry, 1953; Deutsch, 1986; see Pashler, 1999 for review).

Aside from voluntarily orienting attention based on spatial location, people can also selectively attend to a specific frequency based on frequency differences between attended and unattended messages (Treisman, 1964; Underwood & Moray, 1971). Treisman (1964) asked listeners to shadow one of two messages presented dichotically. Listeners were able to shadow a woman's voice more accurately (74%) when the irrelevant message was in a man's voice than when the irrelevant message was in a different woman's voice (31%). Underwood and Moray (1971) observed similar results when they asked listeners to attend to one of two messages. Their task was to detect target digits in the messages. More targets were detected when one voice was a man's voice and the other a woman's voice than when the messages were spoken by speakers of the same gender. These early experiments using dichotic listening tasks show that listeners can voluntarily orient their attention to a particular spatial location or frequency.

Informative Cues in Vision

In addition to instructions, cues have also been used to provide participants with information about where to orient attention. In the basic cueing paradigm, a cue is presented before a target. The cue is used to signal that a target is likely to occur in the cued location. Typically, the time it takes to respond to a target at a cued location is compared with the time it takes to respond to a target in an uncued location (see Posner, 1978, 1980). In a basic cueing paradigm, two types of cues may be used: exogenous cues and endogenous cues. An exogenous cue is one that is not informative about the subsequent target. Because an exogenous cue has no predictive value about the location of an upcoming target, there is no incentive for the participant to voluntarily orient their attention to the location of the cue (Posner, 1980; Ruz and Lupianez, 2002). In contrast, an endogenous cue signals the likely location of the upcoming target. Because an endogenous cue provides useful information about an upcoming target, there is an incentive for the participant to attend to the location that is signalled by the cue (Jonides, 1981; Posner & Snyder, 1975).

In visual attention, a cueing paradigm has been used to demonstrate that attention can be oriented to a specific location (e.g. Downing, 1988; Jonides, 1981; Posner & Cohen, 1984; Tipper, Weaver, Jerreat, & Burak, 1994). In one such experiment, Downing (1988) presented subjects with 12 stimulus boxes positioned around the circumference of an imaginary circle. Subjects were told in advance about the probability of certain types of stimuli appearing at the different locations. Subjects were instructed either to detect a target stimulus or to make brightness, orientation, or form discriminations about a target stimulus. For the first task of luminance detection, subjects were first presented with a fixation marker located in the center of an imaginary circle of 12 stimulus boxes. After the trial was initiated, subjects saw either an arrow cue located in the centre of the imaginary circle, which pointed to one of the 12 boxes, or they saw a neutral cue, which did not point to any of the boxes. The neutral cue, a circle positioned in the centre of the display, indicated that a stimulus was equally likely to occur in any of the 12 stimulus boxes. The arrow indicated that it was very likely that a stimulus would occur in the indicated location (80% of the trials). With this probability manipulation, subjects were led to expect a target at a cued location more often than any of the other positions. After the cue was presented, stimuli occurred in up to 4 of the boxes and, following a stimulus mask in all 12 boxes, subjects were asked the contents of the cued box and of up to 3 other locations where a stimulus was presented. These other locations where stimuli were presented were equally likely to contain a target or a non-target. Targets in the luminance detection task consisted of a short line segment (either horizontal or vertical) and a longer line segment to one side of the box. Each of the relevant stimulus boxes was randomly probed and subjects were required to respond by pressing one of four keys. Subjects indicated whether the stimulus at the probed location was certainly a target, probably a target, probably a non-target, or certainly a non-target. After each response, subjects received feedback as to whether their response about the stimulus was correct. The discrimination tasks were similar to the luminance detection task except that four stimuli were presented on each trial. In addition, rather than targets and non-targets being defined in terms of presence or absence, some attribute of the stimulus was used instead. For example, in the brightness discrimination task, targets were bright, whereas non-targets were dim. Targets in the orientation discrimination task contained a vertical center line segment, while the line segment for non-targets was horizontal. In comparison, for the form discrimination task targets contained two perpendicular line segments, and non targets contained two parallel line segments. It was hypothesized that perceptual sensitivity would decrease as the distance from the cued location increased. The results revealed a main effect for distance such that as the distance between the probed location and the cued location increased, perceptual sensitivity decreased. That is, if the probed location was far from the cued location, target detection / discrimination was impaired. The results also revealed that more accurate responses were made on cued trials where the arrow cue directed attending to a specific stimulus box versus a neutral cue that indicated that stimuli were equally likely to occur in any of the 12 stimulus boxes.

The results that Downing (1988) observed demonstrate that cues can be used to provide subjects with information about the spatial location of a future target and where to orient attention in order to facilitate target detection and discrimination. These findings are consistent with a number of other studies that also used a visual detection task (e.g. Kurylo, Reeves, & Scharf, 1996). For example, Posner, Nissen, and Ogden (1978) presented subjects with either a plus sign on the display screen or an arrow pointing to the left or the right. When the plus sign was presented, the target to be detected was equally likely to occur to the left or the right of fixation. On trials where an arrow was presented, there was a probability of 0.8 that the target would occur on the side indicated (valid trials) and a probability of 0.2 that the target would occur on the opposite side (invalid trials). The results show that reaction times were faster for a target presented at a cued location and slower for targets presented at an uncued location. Since performance was significantly better for valid trials compared to invalid trials, this indicates that there is a benefit from knowing where a target will occur. In addition, there is a cost involved (impaired performance and slower reaction times) when the target occurs at a position that is not expected. This is what Posner refers to as "attentional costs and benefits" (Posner, et al., 1978; Posner, 1980). These results are taken as evidence that cues can be used to provide information about orienting attention to specific spatial locations.

For both detection tasks and discrimination tasks, the cueing paradigm has been used to demonstrate that attention can be oriented to a specific location. Reaction times are faster when the target appears at a cued location, than when it appears at an uncued location. The same holds true for studies using temporal order judgements. For example, Gibson and Egeth (1994) manipulated the number of target stimuli that were presented following a cue to determine if reaction time was faster for cued locations only when a single target was presented. Cueing consisted of brightening one of the stimulus locations. On half the trials, a single target (T1) was presented randomly at either the cued or uncued location. On the other half of the trials, a second target (T2) was presented at the opposite location. The task in this experiment was to identify the location of T1. The time interval between the onset of T1 and the onset of T2, referred to as target asynchrony, was varied. On any given trial, subjects could not predict whether one or two targets would be presented following the cue. After each trial was initiated, subjects were presented with a stimulus display that consisted of a central fixation point with two squares positioned on either side. After 500 msec, one of the two squares was brightened by increasing the thickness of the lines that made-up the square. After a variable cue lead time, T1 appeared randomly at either the cued or uncued location. On trials where there was a second target (T2), it appeared in the opposite square following a target asynchrony of 14, 43, or 114 msec. Subjects were instructed to press a key corresponding to the location of the first target. If T1 appeared in the left square, subjects were required to press the left shift key and to press the right shift key if T1 appeared in the right square. On trials where only a single target followed the cue, reaction times were faster in the cued condition than in the uncued condition. In addition, subjects were significantly more accurate in the cued condition (32% errors) than in the uncued condition (97% errors).

On trials where both T1 and T2 were presented, subjects were faster and more accurate when T1 appeared in the cued location (first-cued condition) than when T1 appeared in the uncued location and T2 appeared in the cued location (second-cued condition). The results of this experiment suggest that visual information processing is improved at cued locations, relative to uncued locations.

Research using visual information provides evidence that having advance knowledge about where a stimulus is likely to occur facilitates responding at the location that was expected and delays responding when the stimulus occurs at an unexpected location (e.g. Downing, 1988; Gibson & Egeth, 1994; Posner, 1980; Posner & Cohen, 1984). Informative cues provide useful information about where to orient attention. However, recent studies have found that cues not only provide information about where to orient attention, but also when to orient attention (Coull, Frith, Buchel, & Nobre, 2000; Coull and Nobre, 1998; Miniussi, Wilding, Coull, & Nobre, 1999). Orienting in time, according to Coull and Nobre (1998) concerns how information about time intervals can be used to orient attention to a point in time when an event is expected. What Coull and Nobre were interested in was whether stimuli that occurred at predictable cued intervals were detected more efficiently compared to stimuli that did not occur when predicted. Subjects were presented with a central cue with two boxes positioned on each side of the central cue. A target, either an "X" or a "+", appeared in one of the boxes and subjects were required to detect the target as quickly as possible. No discrimination was required. Two types of targets were used to enable compatibility with future studies where discrimination was required. On every trial, only one of the target stimuli was presented. The task used in the experiment manipulated the subject's expectancy about where and when the target would occur using four types of cues as follows: cues that predicted both location and onset time (ST), target location only (S), target onset time only (T), and neutral cues (N). The neutral cues predicted neither the target location nor the onset time. The central cue in the visual display consisted of a diamond with a circle positioned inside and an outer circle surrounding the diamond. Depending on which part of the cue was highlighted, subjects attended to the position of the target (left or right) or to the time of the target (300 or 1500 msec from the presentation of the cue) as follows: In the spatial cueing (S) condition, if the left side of the diamond was highlighted, the target was likely to appear in the left box (the opposite if the right side was highlighted). In the temporal cueing (T) condition, highlighting the inner circle indicated that a target would appear after a short time interval (300 msec), whereas highlighting the outer circle indicated that a target would appear after a longer interval (1500 msec). On trials where the cue predicted both spatial position and onset time (ST), one of the circles as well as one side of the diamond was highlighted depending on which combination of location and interval was to be cued. In the neutral cue condition, the entire cue (inside circle, outside circle, and diamond) was highlighted. Subjects were required to detect the target by pressing a button. As expected, reaction times were faster when informative cues were present than when neutral cues were used. In addition, reaction times were slightly faster in the onset time condition compared to the spatial location condition. Interestingly, reaction times were fastest in the condition with informative cues along two dimensions (spatial and temporal combined). Overall, the results indicate that there is not only an advantage to knowing where a stimulus is going to occur but also when the stimulus is expected to occur. Attentional resources can be directed to specific locations as well as to specific moments in time with the use of cues that provide information about a future target. Expectancies about when an event will occur, as well as where it will occur can be used to optimize responding.

Informative Cues in Audition

Relative to research on visual attention, less is known about attending to auditory events. However, the cue-target paradigm that has been used in visual research has also been used in studies of auditory perception (e.g. Mondor, Breau, & Milliken, 1999; Mondor, Terrio, & Hurlburt, 2000; Reuter-Lorenz, Jha, & Rosenquist, 1996; Schmidt, 1996). These auditory versions of the cue-target paradigm yield similar results to those found in the visual domain. As with visual stimuli, the cueing paradigm has been used to demonstrate that attention can be oriented to a specific location.

Spence and Driver (1994) presented subjects with a cue-target paradigm and asked them to judge the location of the target sound. The stimulus onset asynchrony or SOA (the time between the onset of the cue and the onset of the target) was varied. Three loudspeakers were placed on either side of the listener. The cue was always presented through the speaker directly to the side of the listener, and the possible target speakers were positioned 45° toward the front of, and 45° toward the rear of the listener. Participants were told that the target would be presented on the same side as the cue on 75% of the trials. These were called valid trials. The remaining 25% of the trials were invalid trials, meaning the target occurred on the opposite side from the cue. Before beginning each trial, subjects were instructed to attend to the first sound since the target was likely to be presented from the same side. The results showed that target detection time improved when the cue correctly predicted the location of the target (valid trials) compared to trials when the cue did not correctly predict the target location (invalid trials).

Similar results have been found regarding expectancies about the frequency of a target sound based on the frequency of a cue that is presented previously. Mondor and Bregman (1994) conducted an experiment to determine whether attention can be oriented to a specific frequency region. In Experiment 1, a 1000 Hz tone was presented in order to cue subjects to attend to that frequency region. A target was then presented that was the same frequency as the cue on 75% of the trials (valid trials) and different on the remaining 25% of the trials (invalid trials). On invalid trials, the frequencies of the cue and target could be either near to, or far from, each other. For example, the cue had a frequency of 1000 Hz and the target could be either 925 or 1075 Hz (invalid-near trials) or the target could be either 600 or 1500 Hz (invalid-far trials). Subjects were asked to judge whether the target was long or short in duration. The results of Experiment 1 showed that performance was better on valid trials than on invalid trials and also that performance on invalid-near trials was better than on invalid-far trials. Mondor and Bregman (1994) suggest "the focus of attention conforms to a gradient with the density of attentional resources declining with increasing frequency separation from the focal point of attention" (p. 272).

When listeners are asked to detect a sound embedded in a white noise background, they are better at detecting expected tones compared to unexpected tones. In one such study, Greenberg and Larkin (1968) presented listeners with two successive intervals of white noise. On each trial, a pure tone was embedded within one of the white noise intervals. Listeners were asked to indicate in which interval the tone was presented. The experimenters led the listeners to expect that only tones of a particular frequency would be presented (the expected tone is referred to as the primary). Although listeners were expecting only tones of a particular frequency, on 30% of the trials the tone that was presented was of an unexpected frequency (referred to as probes). The experimenters observed that detection was best for expected sounds (the primary). Detection was worst for a probe that was distant in frequency. However, probes that were close in frequency to the primary were detected more often than those that were far away in frequency from the primary. The findings by Greenberg and Larkin (1968) suggest that the listener focuses on a critical band which is centered on the primary and is better able to hear tones at frequencies within that band but not those frequencies that fall outside.

Going back to early experiments that used dichotic listening tasks, shadowing one of the messages improved when the messages were presented in different locations or when the pitch of the speaker's voice differed. Dichotic listening tasks showed that both frequency and location were effective in guiding auditory attention (see Woods, Diaz, Alain, Rhodes, & Ogawa, 2001). However, a debate exists about whether either frequency or location has an advantage in guiding auditory selective attention. Naatanen, Porkka, Merisal, and Ahtola (1980) compared the efficiency of frequency and location cues in an auditory selective attention task. Listeners were presented with sequences of tones lasting 5 seconds and were required to detect targets. On some trials, targets were distinguished from the other tones only by their location (left versus right), while others were distinguished based only on frequency (high versus low). The results indicate that listeners were faster at identifying targets when they were distinguished by location than when they were distinguished by frequency. Naatanen et al. (1980) observed that spatial location seems to have an advantage over frequency in directing selective attention. In contrast to the results obtained by Naatanen et al. (1980), Woods, Alain, and Ogawa (1998) found that listeners were faster and more accurate in identifying targets that were distinguished from non-targets based on frequency compared to targets that were distinguished by location. In addition, Woods et al. had participants respond to target tones that were distinguished based on a combination of frequency and location. In this conjunction condition, performance was better than in conditions where targets were based on location alone. In a follow up study, Woods, Alain, Diaz, Rhodes, and Ogawa (2001) had participants identify targets that were distinguished from distractors based on frequency, location, or a conjunction of frequency and location. In Experiment 1 they observed that participants were faster and more accurate in the frequency condition compared to the location condition. However, at long ISIs, selection based on location was superior, suggesting that ISI is critical when determining the relative effectiveness of selection by frequency and location.

The results of these studies demonstrate, overall, that both frequency and location are effective in guiding auditory attention and that attention can be oriented based on frequency or spatial information provided by the cues.

Studies of Attentional Capture in Vision

In visual experiments, it is evident that voluntary processes play an important role in cue-target tasks. Informative cues can be used to provide important information about the location of a subsequent target (e.g. Gibson & Egeth, 1994). These cues provide expectancies about where or when a subsequent target will occur and allow voluntary orientation of attention to that space. Up until this point the focus has been on voluntary processes involved in visual tasks. However, as is evident in daily activities, an interaction exists between voluntary and involuntary processes. Attention may be controlled involuntarily by a perceptual event that draws attention, even when the observer is consciously trying to ignore it (see Yantis, 1996). This effect has been referred to as attentional capture. Attentional capture in the visual modality occurs when an abrupt visual event draws attention to its location, without intention. This results in improved perception of a subsequent target in the same location, relative to a target that appears in a different location (e.g. Gibson & Egeth, 1994).

Attentional capture in the visual domain is commonly referred to as visual capture. In one experiment, Gibson and Amelio (2000) used a spatial cueing paradigm to investigate unintended shifts of attention caused by uninformative cues. On each trial, subjects were presented with a series of three displays: a fixation point, a cue display, and a target display. The fixation display consisted of a central white "+" within a dark grey box. In addition, four dark grey boxes were positioned on each side of fixation as well as one above fixation and one below. In the onset-cue condition, the outline of one of the four boxes became thicker and changed in colour from grey to white. The onset-cue involved changing both thickness and luminance. In the onset-colour condition, three of the four boxes also changed in thickness and luminance from thin and grey to thick and white. The fourth box was transformed from thin and grey to thick and red. In the onset-target condition, either a white H or a white U appeared in one of the boxes. In the colour-target condition, either a red H or a red U was presented in one of the boxes while a white distractor letter (always E, S, and P) appeared in each of the remaining boxes. On half the trials, an onset cue preceded the target and on the other half a colour cue preceded the target. In this experiment, the cue and the target appeared in the same location on 25% of the trials and in different locations on the remaining 75% of the trials, making the cue uncorrelated with the location of the target. Attentional set was manipulated in this experiment such that half the subjects were consistently presented with onset defined targets while the other half were consistently presented with targets defined by colour. Subjects were instructed to determine which of the two target letters was presented on each trial. The goal of this experiment was to show that exogenous attentional orienting is contingent on attentional control settings. The effects of attentional orienting would be evident from faster reaction times when the cue and target were oriented in the same location compared with reaction times in the different location condition. This should be observed only when the cue type and target type are the same since it is hypothesized that attentional orienting is contingent on the attentional set. The reaction time data were consistent with the expectation that onset cues facilitate responding (faster reaction times) when they are consistent with the defining features of the target but not when they are inconsistent with them. In addition, colour cues were not capable of eliciting attentional capture when these colour cues preceded onset-defined targets. Overall, attentional capture occurred to a cue only when the feature that defined the cue was the same feature that defined the target. That is, attentional capture appears to be contingent on the attentional set or the goals induced by the experimental task.

The idea that attentional capture occurs only when the observer has an optimal attentional set for that feature has been studied extensively (Egeth, Folk, Leber, Nakam,

& Hendel, 2001; Folk, Remington, & Wright, 1994; see Ruz and Lupianez, 2002 for a review). Researchers maintain that it is the boundaries of the task that set attention so that a distinct element can draw attention to itself. That is, in order for an involuntary shift of attention to occur, there has to be a relationship between the properties of the cue and the properties relevant to the task. To test this hypothesis, Folk, Remington, and Johnston (1992) explored the conditions under which attentional capture occurs. They proposed that under conditions of spatial uncertainty, an automatic shift of attention would happen depending on whether there was a feature that was common to both the abrupt event and critical to the task, otherwise a shift of attention would not occur. This has been termed the "contingent involuntary orienting hypothesis" (Folk et al., 1992). In order to test this hypothesis, Folk and his colleagues used a spatial cueing paradigm in which the task was letter discrimination. In Experiment 1, a cue appeared in one of four potential target locations and the relationship between the defining features of the cue and target were manipulated. For one group, the target was an abrupt onset that appeared in one of the locations while the other three potential target locations remained empty. The target that appeared for a second group was a unique coloured letter accompanied by a white non-target letter in the remaining three potential target locations. The cue consisted of four small white dots that surrounded the location of one of the four possible locations. The validity of the cue was also manipulated in this experiment. In one condition, the cue always predicted the location of the target (100% valid), while in the other condition the cue never appeared in the same location as the target letter (100% invalid). Since conditions were either 100% valid or 100% invalid, subjects were fully aware of the spatial relationship between the cue and the target. In addition, on some trials a neutral cue appeared in the centre of the screen and on some trials no cue was displayed. On valid trials, the researchers predicted a benefit (faster reaction times) compared to when no cue appeared regardless of the target type. This was expected because on valid trials, the cue provided useful information about the location of the target and should encourage subjects to attend voluntarily to the cued location. What the experimenters were interested in, however, was performance on invalid trials. They predicted that if the cue automatically captured attention, performance on the task should be impaired even when subjects tried not to attend to the cued location. The reaction time data observed were consistent with these predictions. On valid trials, reaction times were faster for both colour and onset targets. However, on invalid trials the results observed proved interesting. Reaction times were slowed when the target was defined by an abrupt onset, however there was neither a cost nor a benefit when the target was defined by colour. In other words, an abrupt onset cue captured attention only when the defining feature of the target was an abrupt onset and not when the target was defined by colour. The results support the contingent involuntary orienting hypothesis proposed by Folk et al. (1992) that involuntary orienting to a cue depends on whether or not that cue is consistent with the task-relevant properties that define the attentional set. According to this hypothesis, once the attentional set is established stimuli that fulfill the criterion of the set generate a signal that involuntarily captures attention.

While some researchers suggest that involuntary shifts of attention are dependent on the relationship between stimulus properties and the properties required for the task (e.g. Folk et al., 1992; Gibson and Amelio, 2000), others researchers have obtained results that are inconsistent with the contingent involuntary orienting hypothesis. In one such experiment, Theeuwes (1991) had subjects perform a search for a stimulus that was unique (in some dimension) in an array that contained a second target that was unique in an irrelevant dimension. According to the contingent involuntary orienting hypothesis, if subjects are searching for a stimulus that is defined by a specific dimension (i.e. colour), then only stimuli that are also defined by that dimension should capture attention. In his experiment, Theeuwes (1991) had participants search for a target (either an E or an H) that was defined by either luminance or colour. Participants were asked to determine whether the target was presented on each trial. On half the trials, an irrelevant distractor in the other dimension was presented. For example, if the relevant stimulus was a luminance target, the irrelevant distractor was a colour stimulus. In the control condition, no target was presented. On each trial, participants were unaware of the specific feature that defined the target (i.e. luminance or colour) and they only knew they were to search for a unique item. Theeuwes found that reaction times increased when an irrelevant distractor was present, regardless of whether the target was defined by colour or luminance. Overall findings of this study show that any salient singleton (a stimulus that differs substantially in one or more visual attributes from its background), when it differs from the target in some salient feature, has the ability to capture attention. That is, salient unique features can lead to attentional capture

One question that is central to the debate on attentional capture in vision is what are the salient stimulus properties that are capable of capturing attention? According to Egeth and Yantis (1997), there are two major categories of stimulus properties that are capable of capturing attention. The first are stimuli that differ in one or more attributes (e.g. colour or motion) from their background (called feature singletons) and the second are abrupt visual onsets.

With respect to feature singletons, the results obtained by Theeuwes (1991) suggest that stimulus salience plays a role in attentional capture.

Abrupt visual onsets are the second category of stimulus properties that is said to capture attention. Yantis and Jonides (1984) hypothesized that the visual system is sensitive to abrupt onsets. Attention is involuntarily captured by an abrupt onset when there are no other similar stimuli present. The researchers used a visual search task in which targets and non-targets did not differ with respect to a visual feature. In addition, stimuli served equally often as targets and distractors. Subjects were presented with a display of either two or four stimuli and were asked to determine whether a pre-specified target was present in the display. On every trial, one of the stimuli had an abrupt onset and the other items were revealed gradually by removing camouflaging premasks. On some trials, the target was an abrupt onset, while on other trials it was revealed gradually. The researchers predicted that the abrupt onset item would be identified first on every trial, whether it was a target or not. If it was a target, then a response would be made right away, otherwise scanning for the target in the display would continue. Reaction times were fastest when the target was an abrupt onset indicating that an abrupt onset in the visual display captured visual attention.

In a later study, the researchers observed that abrupt onsets were unique in capturing attention since luminance and colour singletons did not capture attention (Jonides and Yantis, 1988). Using a visual search task, they presented observers with an array of figure 8 shapes (pre-masks) that appeared for 1000 msec. The number of figure 8 shapes depended on the search display size. At the onset of each trial, the participants were presented with the target letter for the current trial, followed by the array of figure 8 shapes. The array of figure 8 shapes was replaced by the test display. In the no-onset condition, camouflaged pre-masks were removed gradually to reveal the appropriate letter stimulus. In the onset condition, test items appeared in a previously blank position. For the intensity and colour conditions, all the letters appeared abruptly with one of them being either a luminance singleton (one letter was bright while the others were dim) or a colour singleton (red versus green). In all the conditions, subjects were instructed to search for the target and to press one key if the target was present and another key if the target was absent. Reaction times were fastest in the onset target condition compared to both the no-onset and absent conditions, indicating that an abrupt onset singleton captures attention. However, in the luminance and colour conditions, where the unique item was defined by a difference in either brightness or colour, the target had no special status compared to other items in the display. Yantis and Jonides observed that luminance and colour singletons had no effect on the allocation of attention, despite claims by other researchers that an abrupt change in the luminance at a location tends to draw attention to that location and that attention is captured at the level of luminance change detection (i.e. Gellatly, Cole, & Blurton, 1999). The fact that Yantis and Jonides did not observe an effect of luminance and colour singletons while other researchers claim that an abrupt change in the luminance captures attention can be explained by the type of luminance change. If the luminance change merely causes a change in brightness of a singleton, this is usually not enough of a change to capture attention. However, if luminance change results in an abrupt onset of a new object, attention is typically drawn to the location of the object.

The results of the two studies conducted by Yantis and Jonides (1984, 1988) indicate that abrupt onsets have a unique ability to capture attention. But what specific property of the onset is responsible for attentional capture? The abrupt onset in the studies by Yantis and Jonides could have captured attention in one of two ways. The first being an increase in luminance and the second being that the abrupt onset captured attention as a result of being a new perceptual object. To further investigate the possibility that an abrupt onset could capture attention in different ways, Yantis and Hillstrom (1994) used a modified version of the paradigm used in previous studies. Observers were presented with an array of figure 8 shapes (placeholder display), similar to those used in the Jonides and Yantis (1988) study. The task in this study was to decide which of two target letters appeared in the display. In the search display that appeared after 1000 msec, one of the letters was always a target (H or U). In addition, all of the letters except one appeared in previously occupied locations in the display. One letter, which is referred to as the new element, appeared in a location that was previously empty. Participants were required to press the left response button if the target was 'H' and the right button for 'U'. The results of this study show that attention was captured by the abrupt appearance of an equiluminant object when the target was new and not when the target was one of the old items. A visible luminance change in an old object was not enough for attentional capture to occur. Therefore, a change in luminance is not sufficient to produce attentional capture. These results provide strong support for the theory that attentional capture occurs as a result of the appearance of a new perceptual object.

Despite the strong evidence that new objects capture attention, Gibson (1996) argued that attentional capture may not be the result of a new object, since the improved performance for onset targets may be due to a “perceptual difference between onset and no-onset conditions caused by masking” (Gibson, 1996). In the studies by Yantis and Jonides (1984, 1988), the onset condition appears to have an advantage because the figure 8 masks the no-onset letters but not the onset letters. In his study, Gibson manipulated the degree of masking such that there was a strong masking condition (bright no-onset) and a weak masking condition (dim no-onset). Reaction times were slower in the bright no onset condition, a condition where masking was expected to be strong compared to the dim no-onset condition where masking was expected to be weak. Based on the results, Gibson concluded that abrupt onsets have the ability to capture attention as a result of their higher visual quality and their status as a new perceptual object. Other researchers have also found evidence that attentional capture in vision is due to the appearance of a new perceptual object (Enns, Austen, Di Lollo, Rauschenberger & Yantis, 2001; Yantis & Jonides, 1996).

According to Gibson (1996), the ability of an abrupt onset to capture visual attention may occur because the task-irrelevant singleton has certain properties that may signal a new perceptual object. Hillstrom and Yantis (1994) further examined the possibility that new perceptual objects capture attention. Their hypothesis was that "when a new object file is created to accommodate the appearance of a new perceptual object, attention is directed there in a stimulus-driven fashion" (p. 405). They investigated whether the appearance of a new object, which was defined by motion, would capture attention. Subjects were required to identify a "global" letter (always an H or S) that was formed using multiple small letters. The small letters that made up the “global” letter were the same, either E’s or U’s, except for one, which was called the unique local letter.

On some trials, one of the local letters moved, creating a new perceptual object. There were three main conditions in the experiment: the first of these being a baseline condition, in which the unique local letter was part of the H and S global letters. In the second condition, the onset condition, the unique local letter appeared in a location of the global letter that had a gap in it. It was expected that attentional capture would occur in this condition. In the third condition, when the display appeared, the unique local letter began to move back and forth and continued to move until the participant made a response. This condition was termed the motion condition and it was also expected to cause an involuntary shift of attention. In general, the results show that the time it took to identify the target letter was fastest in the baseline condition compared to the onset and motion conditions. The slowest reaction times were in the motion condition. The results supported the hypothesis that attention is drawn to a new perceptual object when it appears abruptly in a previously blank location. In addition, a new perceptual object also draws attention when it segregates from its background as a result of motion (see Yantis, 1993).

Attentional Capture in Audition

Attentional capture by a unique yet irrelevant singleton distractor has been well-studied in the visual modality, however less is known about how irrelevant singletons attract attention in the auditory domain. An issue that is central to the debate over attentional capture (either visual or auditory) is automaticity. Is the attention-getting potential of certain stimulus properties independent of one's goals? In order to determine whether the effect from an abrupt sound is involuntary, certain criteria must be met (i.e.

Jones, 2001). In order for attentional capture to be considered automatic, "a stimulus cue (or singleton) should quickly over-ride the voluntary control conferred by instructions or intentions" (Jones, 2001, p. 203). In other words, orienting of attention is said to be automatic when performance is impaired despite the fact that the singleton is inconsistent with the task. This issue of automaticity in auditory attentional capture can be addressed indirectly using results from cue-target paradigms with exogenous sound cues or a monitoring task where a distinctive sound (singleton) is embedded in an auditory sequence.

The general idea in an auditory cue-target paradigm is to study whether a distinctive sound cue can call attention to itself, such that listeners orient attention to its location or pitch, regardless of its cue validity. If there is an automatic effect of orienting attention, information about the singleton can facilitate responses to a subsequent target sound by improving performance when the singleton is similar to the target in terms of frequency or location. In contrast, performance is impaired when the singleton and target differ in terms of frequency or location. An experiment by Spence and Driver (1994) required listeners to ignore an initial cue. The cue did not provide any information about the location of the subsequent target, making it an uninformative cue. The results of the study found that the exogenous sound cue drew attention to its location. When the subsequent target appeared in the same location as the cue, reaction times were faster than if the subsequent target appeared in a different location (at short intervals). Based on the criteria for automaticity mentioned above, the results are consistent with the criteria. Despite the fact that listeners were instructed to ignore the initial cue because it was uninformative, they were unable to do so. Also, the exogenous cue pulled attention towards itself, despite being uninformative. The effects of the singleton were immediate because subsequent target performance was improved even at short intervals.

Similar results have been reported by Mondor and Breau (1999). In their experiments, listeners were presented with a cue, followed by a target sound embedded in noise. The stimulus onset asynchrony (SOA) was varied at 150, 450, and 750 msec. Listeners were required to judge the rise-time of the target by pressing '1' if the target was sharp (a sound that reaches its maximum amplitude quickly), and '0' if the target sound was dull (a sound that reaches it maximum amplitude slowly). Since the judgment listeners were required to make was independent of the relationship between the cue and the target, performance should not be affected by the information provided by the cue. In the location condition, the cue and target could come from either the left or right speaker. In the frequency condition, the cue and target were always presented from the speaker directly in front of the listener, and their frequencies were either the same or different. The cue and target were the same frequency or from the same location on 50% of the trials (repeat trials) and they were different frequencies or from different locations on the remaining 50% of the trials (change trials). Response times in the location condition (150 msec SOA) were fastest on repeat trials, compared to change trials. Therefore, when the cue and target were from the same location, responding was facilitated. In the frequency condition, again reaction times were fastest on repeat trials, compared to change trials. Since the listener's task was to judge rise time, the cue provided no useful information about the response required to the target. Despite this fact, the cue drew attention to its frequency or its location. When the subsequent target appeared in the same location as the cue, reaction times were faster than if the subsequent target appeared in a different location (at short SOAs). The same was found for frequency; when the cue and target were the same frequency, responding was facilitated. Although listeners were to ignore the cue because it was uninformative and inconsistent with the attentional set required for the task, they were unable to do so. The cue, although inconsistent with the task, called attention to its frequency or its location in an involuntary manner.

In another series of experiments, Mondor (1999) presented listeners with an auditory spatial cue followed by a target and asked listeners to make a speeded judgment regarding the location of the target (left or right). Across the experiments, the probability that the target would be presented in the cued location was manipulated. In Experiments 1 and 2, the probability that the cue and target would be presented from the same location (repeat trial) or in a different location (change trial) was .50. Therefore, the spatial relation between the cue and target could not be predicted. The results of both indicate faster reaction times in the repeat condition compared to change conditions. Although the cue provided no useful information about the subsequent target, listeners were unable to ignore it. The results of the study found that the exogenous sound cue drew attention its location. When the subsequent target appeared in the same location as the cue, reaction times were faster than if the subsequent target appeared in a different location (at short SOAs). Attention was allocated to the location of the cue despite the fact that it provides no useful information about the subsequent target. In each of these studies, attentional capture impairs performance, despite the fact that the cue is inconsistent with the task.

Auditory attentional capture can also be studied using results from a monitoring task where a distinctive sound (singleton) is embedded in an auditory sequence. In one such experiment, performance on a target detection task was impaired following the presentation of a distinct sound in an auditory sequence. In a series of experiments, Arnell and Jolicoeur (1999) instructed participants to report the identity of a target and probe. The auditory stimuli presented were streams of spoken letters and also included the numbers 1, 2, 3, and 4. When listening to the rapid serial auditory presentation (RSAP) of spoken letters, listeners were required to perform two tasks: The first being to report whether the letter X was spoken and the second to identify which one of the four possible numbers (1, 2, 3, or 4) was presented. What the researchers observed was that when one of the numbers (probe) was presented immediately following the letter X (target), responding to the probe was significantly impaired. This effect has become known as 'attentional blink'. When stimuli are presented rapidly, participants fail to accurately report the probe when it is presented within approximately 500 ms of the target.

The possibility that selection of a target might impair subsequent processing has been studied not only with spoken auditory stimuli, but using pure tones as well (Mondor, 1998). In two experiments, Mondor (1998) presented listeners with a sequence of tones. Listeners were required to detect both a target and a subsequent probe. The target varied in its position in the sequence and the probe was presented in one of seven or eight possible positions following the target. In Experiment 1, the post-target item could either be the probe or a distractor whereas in Experiment 2, the post-target item was replaced by silence. In both experiments, detection of the probe was impaired when it was presented directly following the target. Attention was directed to the target such that subsequent processing of the probe was impaired when the probe was presented immediately following the target.

Other researchers who study the attentional blink have observed results that are interestingly similar to attentional capture. In the above study, listeners were required to make separate responses regarding the target and the probe. It was observed that probe detection is significantly impaired when it is presented immediately following the target. A study by Shen and Mondor (in press), provided listeners with a typical attentional blink paradigm where they were instructed to make a detection judgement for both the target and the probe sounds in a rapid serial auditory presentation. The researchers observed an attentional blink effect where probe detection was impaired when it was presented following the target sound. Of particular interest, however, was a condition where listeners were instructed to ignore the target and make a detection judgement for the probe sound. They observed an attentional blink effect emerged even though listeners were instructed to ignore the target. This suggests that, with respect to the attentional blink, attention appears to be captured by the target sound, such that probe detection is impaired.

Evidence also exists for attentional capture by an unexpected time change in a sequence. For example, Jones (2001) observed that sounds in auditory sequences that deviate from a temporal regularity can capture attention. If a probe sound is presented following an interruption in temporal regularity, detection of the probe is impaired relative to if the probe was presented in a sequence of temporal regularity. If listeners are set to attend to a regular time pattern and an unexpected time change occurs, processing of a subsequent probe is impaired. This suggests that unexpected temporal changes may also capture attention.

Present Research

While attentional capture has been well-studied in the visual modality, less is known about how irrelevant items (singletons) attract attention in the auditory domain. If a distinct irrelevant distractor sound is embedded in an auditory sequence, involuntary orienting of attention is said to occur if performance (on a target detection task) is impaired despite the fact that the singleton is inconsistent with the task. Cue-target studies have demonstrated that an exogenous cue that provides no information about an upcoming target can still draw attention to the cued location and frequency. As a result, subsequent processing of the target was impaired. Although these studies demonstrated attentional capture, on half the trials a valid cue was presented which was relevant to the task. In these cue-target studies, because of the presence of valid trials, it is not possible to conclude that attentional capture was noncontingent since on half the trials the cue provided accurate information about the target. The present study used an auditory sequence where all the sounds were consistent with the task. For example, the target sound was a low pitch sound presented among a series of low pitch distractor sounds. Therefore, there is incentive for listeners to pay attention to the distractor sounds. If a sound is presented that was inconsistent with the task and attention is drawn to that sound, then we can say that attentional capture was noncontingent. The purpose of the present study was firstly to examine attentional capture in auditory sequences based on a frequency violation. The basic method used an auditory sequence in which listeners were required to detect a target among similar distractor sounds. The presence of a violation that is unique on an irrelevant dimension (e.g., a low frequency violation in a search for a high frequency target) produces auditory attentional capture if it interferes with target detection. The prediction was that if a violation captures attention, then perception of a subsequent target, when presented immediately following the violation, would be impaired.

The second goal was to determine whether attention is oriented towards the violation voluntarily or involuntarily. In order for attentional capture to be considered involuntary, the effects of the violation must be immediate and the effects of the violation should over-ride voluntary intentional control of the search. Voluntary control is brought on by the instructions given. If the target is the same pitch as the distractor elements in a sequence, then the voluntary process of attending to only that pitch should not be impaired by a task-irrelevant pitch. Attentional capture occurs if there is an involuntary shifting of attention to the irrelevant pitch, despite efforts to ignore it. This is indicated by a decrease in target detection accuracy if the target is presented directly after the violation.

A third goal of the present study was to determine whether auditory attention may be captured by frequency, location, or both (conjunction). A conjunction violation was predicted to capture attention more efficiently than a violation defined along a single dimension. A conjunction benefit was predicted for a couple of reasons. Firstly, in vision, a critical factor which determines whether or not a particular singleton will capture attention is the salience of the singleton compared to the target, that is, how much the singleton and target differ (e.g. Theeuwes, 1991, 1992). If the singleton is a violation defined along a single dimension, and it captures attention, then it may be that a singleton defined along two dimensions (frequency and location) will be more salient. Thus, as the salience of an irrelevant sound increases, the degree to which it captures attention may also increase. Secondly, support for the idea that a conjunction violation may capture attention more efficiently comes from studies of visual search (Duncan and Humphreys, 1992). According to Duncan and Humphreys, search time in vision depends on how similar the target is to the distractors and also how similar the distractors are to each other. If the features that define the singleton are different from those that define the target, then the singleton may capture attention more efficiently. In other words, the singleton may call attention to itself as a result of being different from both the target and the other distractors. Taken together, these reasons provide a basis to suggest that a conjunction violation may capture attention more efficiently than a violation defined along a single dimension. If so, target detection should be slower and less accurate when the features that define the target and violation are different from each other. For example, if the target is defined in terms of frequency (e.g. a low pitched sound) and the violation is defined in terms of both frequency and location, attention will be captured to a greater degree resulting in impaired target detection.

Basic Questions/Logic

Experiment 1 was performed to determine whether an abrupt change in frequency (frequency violation) captures attention in an involuntary manner. The logic was that if a frequency violation captures attention, then perception of a target should be impaired when it is presented immediately following the violation. This type of result would indicate that the involuntary effect of an abrupt unexpected sound interferes with target detection despite the person's goal to ignore it. Experiment 2 evaluated whether auditory attention can be captured by an abrupt location change. The third experiment had three main purposes. The first of these was to attempt to replicate the results of Experiment 1. The second purpose was to attempt to replicate the results found in Experiment 2. Aside from replicating the results of Experiments 1 and 2, Experiment 3 explored the possibility that when a violation has a different pitch and comes from a different location, it may more effectively capture attention

Experiment 1

Method

Participants

Twenty undergraduate students (8 male, 12 female) enrolled in an Introduction to Psychology course at the University of Manitoba participated in exchange for course credit. None of the participants reported any uncorrected or corrected hearing impairment. Participants gave informed written consent prior to beginning the experiment.

Materials

Computer and Sound System. The experiment was controlled using a Dell Dimension XPS T550 Pentium computer running the E-Studio Software System (Psychology Software Tools Inc., 1999). Sounds were presented through Sony SRS-PC30 speakers positioned 45° to the left and right of the listener.

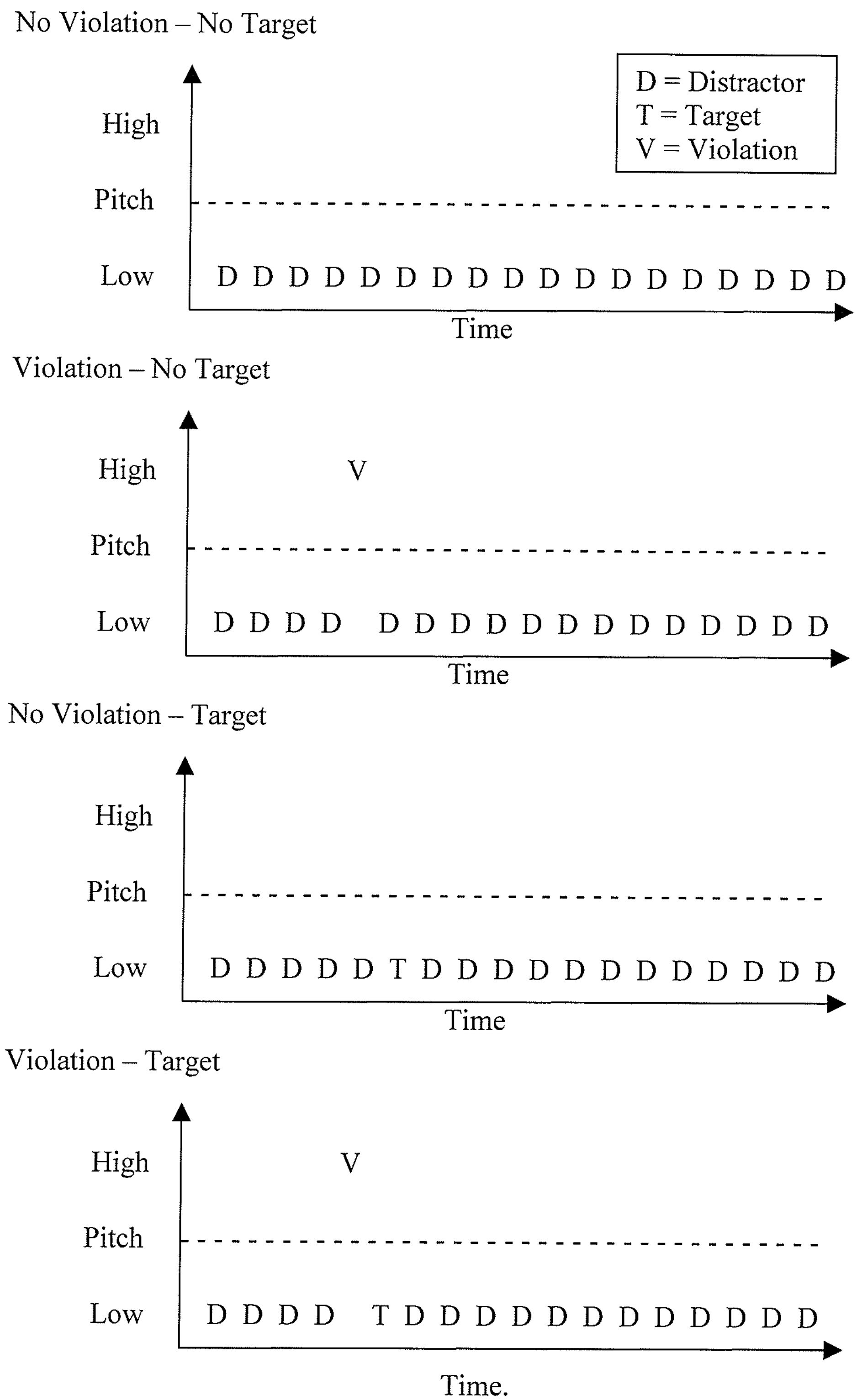
Sounds. Four tones were synthesized using the Cool Edit Software System (Syntrillium Software Corporation, 1999). The frequency of the low pitch distractor and violation sounds was 500 Hz and the frequency of the high pitch distractor and violation sounds was 3000 Hz. The target sound was always the same pitch as the distractor sound, and was composed of three, 10 msec pulses. The result was a target that sounded slightly distorted ('noisy'). Each of the sounds was 30 msec in duration and incorporated 2 msec linear onset and offset amplitude ramps to eliminate onset and offset clicks. All stimuli were presented at a comfortable listening volume (70 dB).

Design and Procedure

Auditory attentional capture was examined using a rapid serial auditory presentation (RSAP) design. On both the practice and experimental trials, participants were presented with a sequence of 18 sounds. A low pitch sequence was made up of all low pitch distractor sounds, while a high pitch sequence was made up of only high pitch distractor sounds. A frequency violation was presented in the sequence half of the time. For low pitch sequences, the frequency violation was a high pitch sound, while a low pitch sound was presented as the violation for high pitch sequences. In addition, on half the trials, a target sound of the same pitch as the distractor sounds was presented. On trials where both a violation and a target were presented, the target could occur in any one of eight temporal positions following the violation. For example, if the target was presented in the position immediately following the violation, this was referred to as 'Position 1'. On trials for which the violation was replaced with a distractor (no violation), the target could be presented in any of the relative positions 1-8. Both the practice and experimental trials consisted of four different trial types, each occurring on 25% of the trials. These trial types are as follows: no violation-no target, violation-no target, no violation-target (relative target positions 1 to 8), and violation-target (relative target positions 1 to 8). Thus, there was no predictable relationship between the violations and the targets. Participants completed 16 practice and 128 experimental trials. A visual representation of the trial types can be seen in Figure 1.

Participants were required to initiate each trial by pressing the space bar. At the end of each sequence, participants were required to respond to the question "Was the target sound presented?" by pressing '1' for yes or '0' for no on the keyboard. Participants were given both oral instructions as well as standard written instructions (presented on the computer screen) prior to beginning the experiment. Before the practice trials, participants were presented with the sounds that were to be used in the experiment in order to become familiar with them.

Figure 1. Visual representation of the four different trial types (Experiment 1). The target in this figure occurs in Relative Position 1, but varies from Position 1 to Position 8 in Experiment 1.



Results

The results obtained in Experiment 1 are summarized in Table 1. The data provide clear evidence that perception of a target sound is impaired when it is presented immediately following a frequency violation. This is supported by the results of a two-way repeated measures analysis of variance (ANOVA) (Preceding Violation X Position) that was used to evaluate performance. Statistical analysis of the accuracy data revealed significant main effects for Preceding Violation, $F(1, 19) = 23.324, p < .001$ and Position, $F(7, 133) = 8.312, p < .001$. Generally, participants made more errors when a violation preceded a target than when a target was presented alone. As the relative position between the violation and the target increased from 1 to 8, target detection accuracy also increased. The analysis also revealed a significant interaction between Preceding Violation and Position, $F(7, 133) = 6.313, p < .001$. Planned comparisons performed to evaluate the effect of Preceding Violation at each position revealed a significant effect at Position 1, $F(1, 19) = 22.294, p < .001$, Position 2, $F(1, 19) = 6.603$, $p = .019$, Position 3, $F(1, 19) = 5.147, p = .035$, Position 4, $F(1, 19) = 7.027, p = .016$, and Position 5, $F(1, 19) = 5.516, p = .030$. Participants made more errors at these positions when a violation preceded the target than when the target was presented alone. There were no significant effects found at Positions 6, $F(1, 19) = 1.305, p = .267$, Position 7, $F(1, 19) = 0.487, p = .494$, or Position 8, $F(1, 19) = 1.652, p = .214$.

On trials where a target was not presented, target detection accuracy was high. When no violation and no target were presented, detection accuracy was .95. When a violation was presented but no target, detection accuracy was .88. False alarm rates were .05 and .12, respectively.

Table 1: Experiment 1 Mean Percent Correct and Standard Errors (SE) for Accuracy data as a Function of Preceding Violation (target presented alone or violation precedes target), and Position (Position 1 to Position 8)

		Position							
		1	2	3	4	5	6	7	8
No violation									
	Mean	97.5	96.3	98.8	98.8	98.8	97.5	97.5	98.8
	SE	1.7	2.0	1.2	1.2	1.2	2.5	1.7	1.2
Violation									
	Mean	58.8	78.8	87.5	87.5	91.3	93.8	95.0	93.8
	SE	7.8	6.6	4.6	3.8	3.3	3.1	2.9	3.6

Discussion

The results obtained in Experiment 1 provide clear evidence that when a frequency violation occurs, perception of a subsequent target is impaired when it is presented in any one of the first five positions following the violation. This experiment suggests that an abrupt frequency change captures auditory attention in a strongly involuntary manner. Attentional capture is said to be automatic if performance is impaired despite the fact that the singleton is inconsistent with the task. Impaired performance in the target detection task indicates that the involuntary effect from an abrupt unexpected sound interferes with target detection despite the person's goal to ignore it.

If attention can be captured by an abrupt unexpected change in frequency, is the same true for an unexpected change in location? Experiment 2 evaluated whether auditory attention can be captured by an abrupt location change.

Experiment 2

Method

Participants

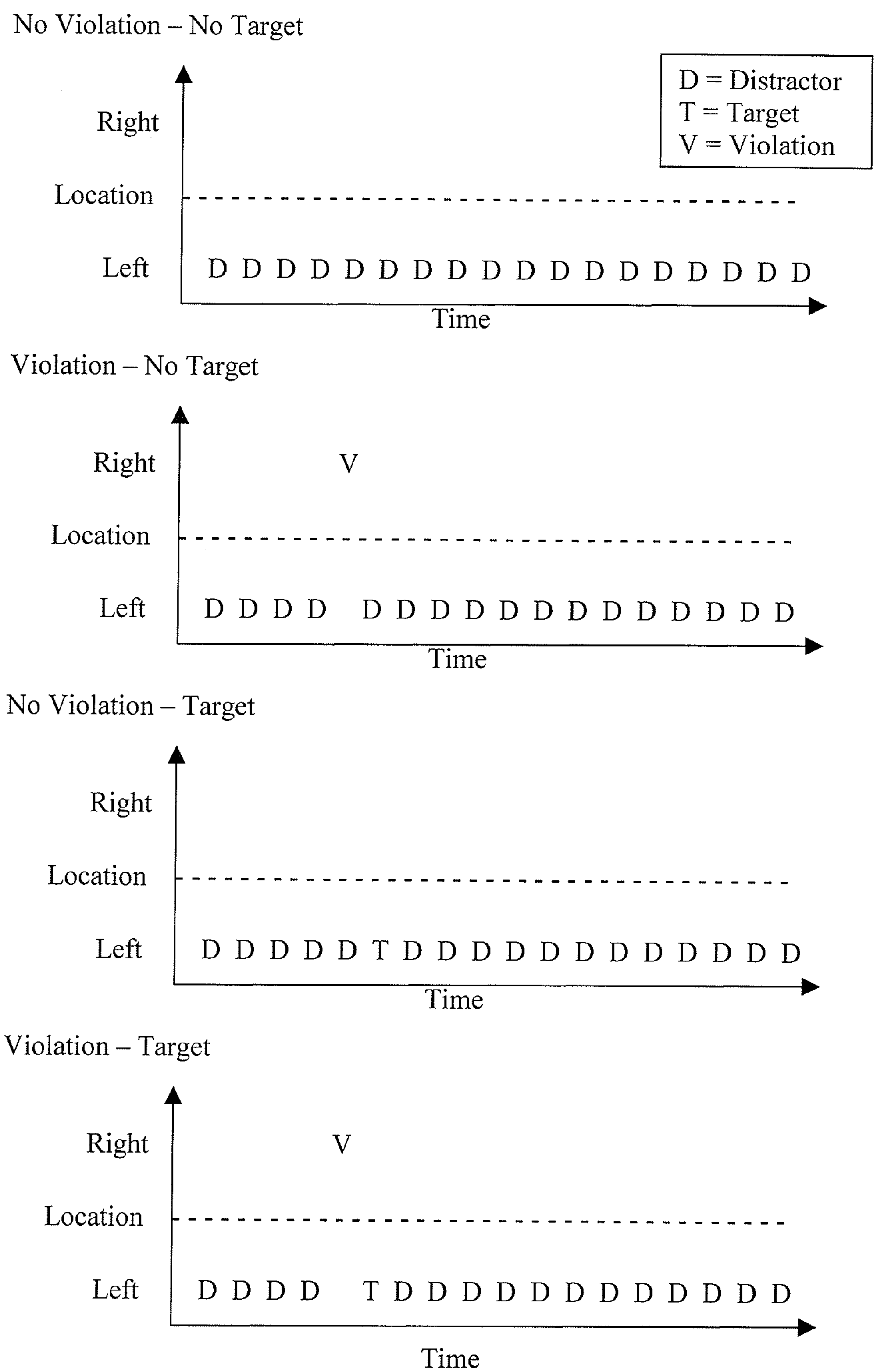
Twenty undergraduate students (11 male, 9 female) enrolled in an Introduction to Psychology course at the University of Manitoba participated in exchange for course credit. As in Experiment 1, none of the participants reported any uncorrected or corrected hearing impairments and informed written consent was given prior to beginning the experiment.

Procedure

Experiment 2 was founded on the four main trial types used in Experiment 1. The procedure was same as that used in Experiment 1, however rather than a frequency violation occurring, a location violation occurred instead. The experiment consisted of a 'Left' condition and a 'Right' condition. In the Left condition, the sound sequence (distractors and target) was always presented from the left with a location violation being presented from the right. In the Right condition, the sound sequence was presented from the right with a location violation presented from the left. An example of each of the trial types for Experiment 2 can be seen in Figure 2. Half of the participants completed the

'Left' condition, while the other half completed the 'Right' condition. In both conditions, the structure of the individual trials was the same as in Experiment 1, except that a location violation was inserted rather than a frequency violation. The violation was always the same pitch as the distractor and target sounds but it was presented from a different location, therefore only a change in location occurred. For each of the two conditions (left and right), an equal number of high-pitched and low-pitched streams were randomly presented. A target sound from the same location as the distractor sounds was presented on half the trials. The relative position of the target ranged from Position 1 to Position 8, as in Experiment 1. The sounds for Experiment 2 were the same as those used in the first experiment. Participants were again presented with 16 practice and 128 experimental trials.

Figure 2. Visual representation of the four different trial types in the Left Condition (Experiment 2). The target in this figure occurs in Relative Position 1, but varies from Position 1 to Position 8 in the experiment.



Results

Mean percent correct and standard errors are described in Table 2. A two-way ANOVA (Preceding Violation X Position) was computed. The results revealed significant main effects of Preceding Violation, $F(1, 18) = 15.041, p = .001$, and of Position, $F(7, 126) = 10.609, p < .001$. The analysis also revealed a significant interaction between Preceding Violation and Position, $F(7, 126) = 6.036, p < .001$. In order to investigate the effect of preceding violation at each of the positions, planned comparisons were performed. A significant effect of Preceding Violation was apparent at Position 1, $F(1, 19) = 12.715, p = .002$, and Position 2, $F(1, 19) = 5.444, p = .031$. Participants made more errors at these positions when a violation preceded the target than when the target was presented alone. There were no significant effects found at Position 3, $F(1, 19) = 0.192, p = .666$, Position 4, $F(1, 19), = 1.000, p = .330$, Position 5, $F(1, 19) = 0.000, p = 1.000$, Position 6, $F(1, 19) = 1.000, p = .330$, Position 7, $F(1, 19) = 1.000, p = .330$, or Position 8, $F(1, 19) = 0.810, p = .379$.

On trials where a target was not presented, target detection accuracy was high. When no violation and no target were presented, detection accuracy was .98. When a violation was presented but no target, detection accuracy was .94. False alarm rates were .02 and .06, respectively.

Table 2: Mean Percent Correct and Standard Errors (SE) for Accuracy data as a Function of Preceding Violation (target presented alone or violation precedes target) and Position (Position 1 to Position 8) in Experiment 2.

		Position							
		1	2	3	4	5	6	7	8
No violation									
	Mean	96.3	90.0	95.0	96.3	98.8	98.8	98.8	95.0
	SE	2.7	4.2	2.4	2.1	1.2	1.3	1.2	4.0
Violation									
	Mean	70.0	72.5	93.8	93.8	98.8	100.0	97.5	98.8
	SE	6.4	6.7	2.4	3.1	1.2	0.0	1.7	1.2

Discussion

The purpose of Experiment 2 was to investigate the possibility that an abrupt, unexpected change in location can capture attention. It was hypothesized that performance on the target detection task would be impaired when an unexpected location violation occurred, similar to when a frequency violation is presented. Despite the fact that the location violation was inconsistent with the goals of the detection task, performance was impaired when the target immediately followed the violation, suggesting involuntary capture of attention by the violation. The results obtained in Experiment 2 indicate that a location violation significantly impairs detection of a subsequent target sound. As seen in Experiment 1, an abrupt frequency change captures attention in an involuntary manner and, as evident in Experiment 2, a location violation seems to act in the same way. Although the location violation captured attention, the location capture effect did not last as long as when a frequency violation captured attention. The results of both Experiment 1 and Experiment 2 suggest that both a frequency and a location violation are capable of capturing attention, a location violation does not appear to capture attention to the same magnitude as a frequency violation does. A location violation results in attention being captured but the impairment appears to be smaller and shorter-lasting than when a frequency violation captures attention. If both a frequency violation and location violation are capable of causing an involuntary shift of attention, would a conjunction violation also cause even greater attentional capture? That is, would a violation defined by both a change in frequency *and* a change in location capture attention to a greater degree? If so, would it capture attention more effectively, resulting in a greater magnitude of attentional capture than violations defined along a single dimension? Experiment 3 explores the possibility that attention can be captured by an unexpected sound defined by both a change in frequency and location. In vision, a critical factor which determines whether or not a particular singleton will capture attention is the salience of the singleton compared to the target (Theeuwes, 1991, 1992). Search time in vision depends on how similar the target is to the distractors and also how similar the distractors are to each other (Duncan and Humphreys, 1992). While searching for a target is voluntary attention, the same mechanisms may be involved for involuntary attention. If the singleton is a violation defined along a single dimension, and it captures attention, it may be that a singleton defined along two dimensions (frequency and location) will be more salient. One would expect that as the salience of an irrelevant sound increases, the degree to which it captures attention might also increase.

Experiment 3

Method

Participants

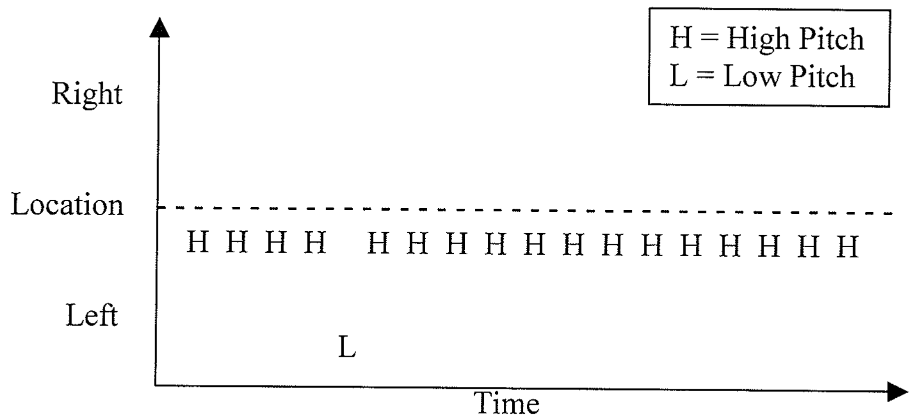
Twenty undergraduate students (10 male, 10 female) enrolled in an Introduction to Psychology course at the University of Manitoba participated in exchange for course credit. None of the participants reported any uncorrected or corrected hearing impairments.

Procedure

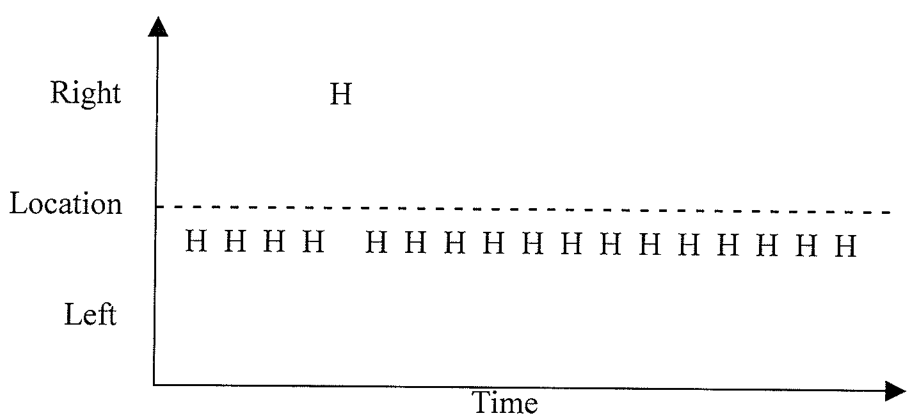
The third experiment was based on a between-subjects design where each participant was presented with one of four types of sequences (high-left, high-right, low-left, low-right). Within each sequence, a frequency-only violation, a location-only violation or a violation based on both frequency and location (conjunction) was presented. An example of each type of violation can be found in Figure 3. Again the frequency of the low pitch distractor and violation sounds was 500 Hz and the frequency of the high pitch distractor and violation sounds was 3000 Hz. As in Experiment 2, sounds were either presented from the left or the right. The target sound was always the same pitch as the distractor sounds and was always presented from the same location as the distractor sounds. The four main trial types; no target-no violation, no target-violation, target-no violation, and target-violation, were the same as in the previous experiments. Each participant completed an experiment that consisted only of one of the four possible basic sequence types (high-left, high-right, low-left, low-right). For each of these sequence, a frequency violation, a location violation, and a conjunction violation could occur, each presented on one-third of the trials. For example, in the High-Left experiment, the distractor sounds and the target were high pitch sounds presented from the left speaker. When a violation occurred, it could have been a low pitch sound presented from the left (frequency change only), a high pitch sound presented from the right speaker (location change only), or a low pitch sound presented from the right (different frequency and different location). This was the case for each of the four basic sequences defined by location and pitch. Violation and target sounds were presented in the same positions in the sequence as in the previous experiments. For this experiment, participants completed 16 practice trials and 384 experimental trials.

Figure 3. Visual representation of the three different violations in the High-Left Condition (Experiment 3). The target in this figure occurs in Relative Position 1, but varies from Position 1 to Position 8 in Experiment 3.

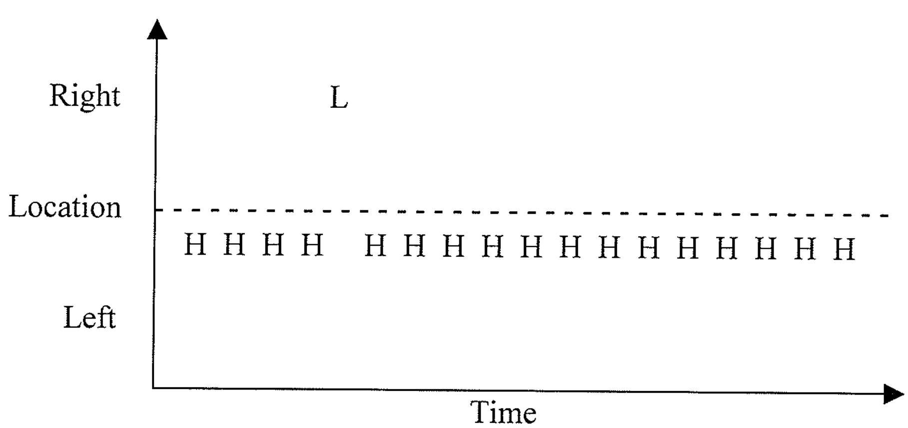
Frequency Only Violation



Location Only Violation



Conjunction Violation



Results

Mean percent correct and standard errors (SE) for accuracy data are described in Table 3. A 3-way repeated measures ANOVA (Preceding Violation X Position X Type of Violation) was performed on the data and the results did not show a significant main effect for Type of Violation, $F(2, 38) = 1.623, p = .211$. In addition, the ANOVA failed to show significant interactions between Preceding Violation and Type of Violation, $F(2, 38) = 0.716, p = .495$, Type of Violation and Position, $F(14, 266) = 1.504, p = .109$, or Preceding Violation, Type of Violation, and Position, $F(14, 266) = 1.327, p = .191$. The analysis did, however, reveal a significant main effect of Violation, $F(1, 19) = 20.917, p < .001$, as well as of Position, $F(7,133) = 29.148, p < .001$. A significant interaction was revealed for Violation and Position, $F(7, 133) = 10.223, p < .001$. The three-way interaction between Preceding Violation, Type of Violation, and Position was examined using two-way ANOVAs to evaluate the effects of preceding violation and position across each condition individually as described below.

Table 3: Mean Percent Correct and Standard Errors (SE) for Accuracy data as a Function of Preceding Violation (target presented alone or violation precedes target), Condition (frequency, location, and conjunction), and Position (Position 1 to Position 8) in Experiment 3.

		Position							
		1	2	3	4	5	6	7	8
No violation									
	Frequency	92.5	80.0	96.3	93.8	95.0	95.0	95.0	91.3
	SE	3.2	5.9	2.7	4.0	2.9	2.3	2.9	3.8
	Location	91.3	85.0	93.8	96.3	96.3	98.8	95.0	90.0
	SE	3.8	5.6	3.1	2.0	2.0	1.2	2.9	3.3
	Conjunction	87.5	87.5	92.5	93.8	87.5	95.0	93.8	92.5
	SE	4.6	5.0	3.2	3.1	3.8	2.3	3.1	2.6
Violation									
	Frequency	56.3	62.5	85.0	93.8	93.8	92.5	93.8	90.0
	SE	8.1	6.4	4.6	3.1	4.0	2.6	3.1	3.3
	Location	73.8	63.8	90.0	97.5	91.3	90.0	95.0	91.3
	SE	6.4	7.4	2.8	1.7	3.3	3.3	2.3	3.8
	Conjunction	52.5	70.0	88.8	93.8	97.5	95.0	95.0	93.8
	SE	7.7	7.4	3.8	2.5	1.7	2.3	2.3	2.5

When analyzing the data for each of the different types of violations separately, it was found that when only a frequency violation was presented, there was a significant main effect of both Preceding Violation, $F(1, 19) = 15.238, p < .001$, and Position, $F(7, 133) = 13.373, p < .001$. There was also a significant interaction effect between these variables, $F(7, 133) = 5.533, p < .001$. Planned comparisons (a one-way ANOVA) performed to evaluate the interaction effect found significant effects of Preceding Violation at Position 1, $F(1, 19) = 15.681, p < .001$, Position 2, $F(1, 19) = 5.444, p = .031$, and Position 3, $F(1, 19) = 5.942, p = .025$. Participants committed more errors at these positions when a frequency violation preceded the target than when the target was presented alone. The analysis did not reveal significant effects at Position 4, $F(1, 19) = 0.000, p = 1.000$, Position 5, $F(1, 19) = 0.073, p = .789$, Position 6, $F(1, 19) = 0.655, p = .428$, Position 7, $F(1, 19) = 0.137, p = .716$, or Position 8, $F(1, 19) = 0.106, p = .0748$.

When only a location violation was presented, there were significant main effects of Violation, $F(1, 19) = 8.504, p = .009$ and Position, $F(7, 133) = 10.928, p < .001$. A significant interaction was found between Violation and Position, $F(7, 133) = 3.039, p = .005$. Planned comparisons performed to evaluate the interaction revealed significant effects of Preceding Violation at Position 1, $F(1, 19) = 5.444, p = .031$, Position 2, $F(1, 19) = 5.431, p = .031$, and Position 6, $F(1, 19) = 7.107, p = .015$. Again, participants made more errors at these positions when a location violation preceded the target than when the target was presented alone. The planned comparisons did not however reveal significant effects at Position 3, $F(1, 19) = 1.305, p = .267$, Position 4, $F(1, 19) = 0.322, p = .577$, Position 5, $F(1, 19) = 2.923, p = .104$, Position 7, $F(1, 19) = 0.000, p = 1.000$, or Position 8, $F(1, 19) = 0.056, p = .815$.

Finally, when a conjunction violation was presented, there were significant main effects of Preceding Violation, $F(1, 19) = 5.945$, $p = .025$, and Position, $F(7, 133) = 15.312$, $p < .001$. The analysis also revealed a significant interaction between the two, $F(7, 133) = 7.804$, $p < .001$. When evaluating the effect of the interaction, planned comparisons revealed significant effects of Preceding Violation at Position 1, $F(1, 19) = 13.591$, $p = .002$, Position 2, $F(1, 19) = 5.144$, $p = .035$, and Position 5, $F(1, 19) = 5.630$, $p = .028$. It did not reveal significant effects of the Preceding Violation at Position 3, $F(1, 19) = 0.810$, $p = .379$, Position 4, $F(1, 19) = 0.000$, $p = 1.000$, Position 6, $F(1, 19) = 0.000$, $p = 1.000$, Position 7, $F(1, 19) = 0.192$, $p = .666$, or Position 8, $F(1, 19) = 0.322$, $p = .577$. Again, participants made more errors at Position 1 and Position 2 when a violation preceded the target than when the target was presented alone.

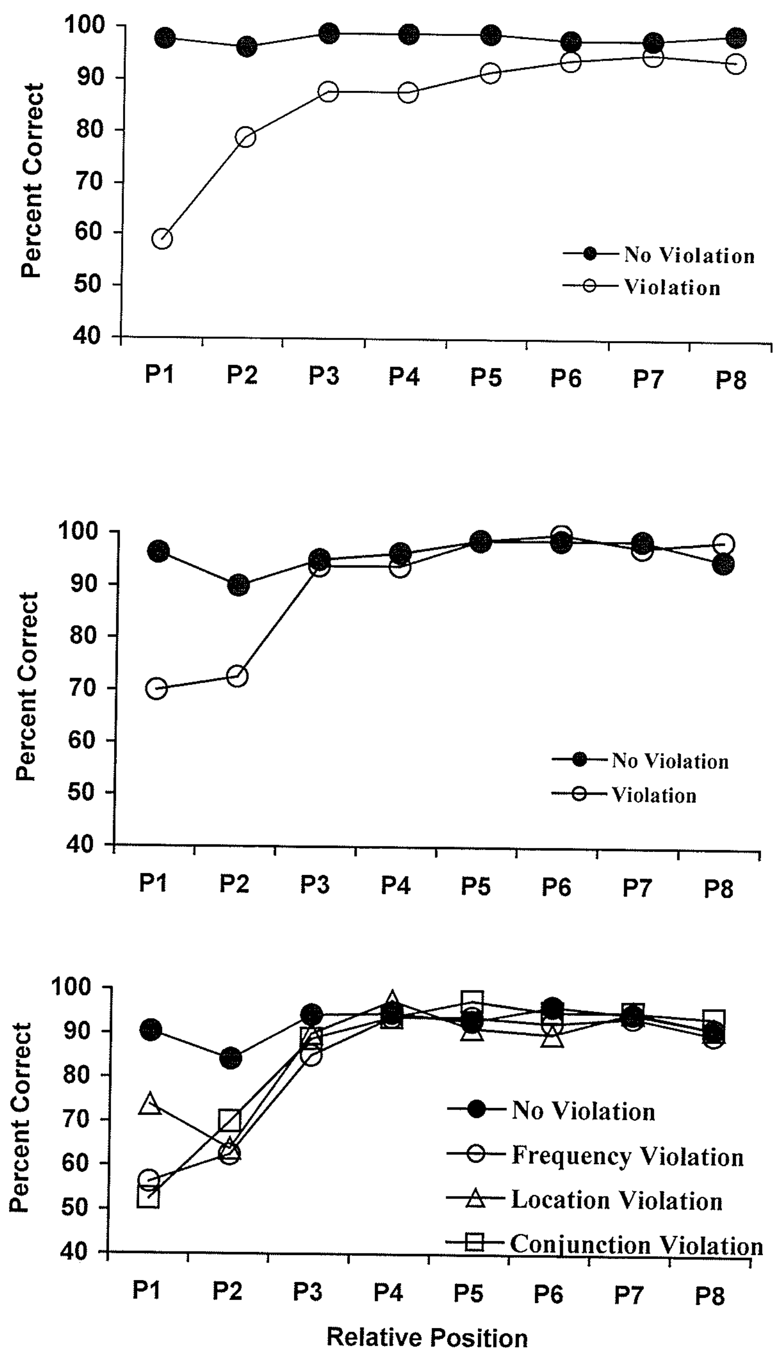
On trials where a target was not presented, target detection accuracy was high. When no violation and no target were presented, detection accuracy was .94. When a violation was presented but no target, detection accuracy was .91. False alarm rates were .06 and .09, respectively.

Discussion

The results obtained in Experiment 3 replicated the results of Experiment 1 that a frequency violation can capture auditory attention in an involuntary manner and impair performance on a subsequent target presented immediately following a frequency violation (i.e. Positions 1 and 2). Experiment 3 also replicated the results of the second experiment where a location violation was presented. Again, target performance was impaired when the target was presented immediately following the location violation. The third goal of Experiment 3 was to evaluate whether a conjunction violation was able to capture attention to a greater degree than a violation defined along a single dimension. The results indicate that a conjunction violation did in fact impair performance when the target was presented in Position 1, Position 2, and Position 5 following the violation. In other words, participants made more errors at these positions when a violation was presented before the target than when the target was presented alone. This suggests that a conjunction violation captures attention. However, the results show that a conjunction violation does not capture attention more effectively than a violation defined along a single dimension. The results of Experiment 3 demonstrate that the involuntary effect from an abrupt unexpected sound interferes with performance despite the voluntary control of the person's goal to ignore the violation, regardless of whether it is a frequency, location, or conjunction violation. The results of all three experiments can be seen in Figure 4.

It was hypothesized that a violation defined by both a change in frequency and a change in location would not only capture attention, but would do so more effectively than a single feature violation. Thus, target detection should be worse for the conjunction condition compared to the frequency only and location only conditions and the effects of the conjunction violation should last longer than the first two positions. The results of Experiment 3 do indicate that a conjunction violation captures attention. However it does not appear that a violation defined by both frequency and location captures attention more efficiently than a violation defined along a single dimension alone.

Figure 4. Mean percent accuracy for no-violation and violation as a function of relative position is described for Experiment 1 (top panel), Experiment 2 (middle panel), and Experiment 3 (bottom panel).



General Discussion

The present study was conducted to examine the effects of an abrupt, unexpected sound on detection of a subsequently presented target sound. Experiment 1 was performed to determine whether an abrupt change in frequency (frequency violation) captures attention in an involuntary manner. The logic was that if a frequency violation captures attention, then perception of a target should be impaired when it is presented immediately following the violation. This type of result would indicate that the involuntary effect from an abrupt unexpected sound interferes with target detection despite the person's goal to ignore it. The prediction for Experiment 1 was that if a violation captures attention, then perception of a subsequent target presented immediately following the violation should be impaired. The results indicate that an abrupt sound of an unexpected frequency is capable of capturing attention, even though it is inconsistent with the task.

The purpose of Experiment 2 was to investigate the possibility that an abrupt, unexpected change in location can capture attention. It was hypothesized that, as was the case when a frequency violation was presented, performance on the target detection task would be impaired when an unexpected location violation occurred. The results of Experiment 2 provide evidence that auditory attention can not only be captured by a sound of an unexpected frequency, but also by a sound presented from an unexpected location.

The third, and final, experiment had three main purposes. The first of these was to attempt to replicate the results of Experiment 1 that attention can be captured by a sound of an unexpected frequency. The second purpose was to attempt to replicate the results found in Experiment 2 that a location violation can capture auditory attention. Aside from replicating the results of Experiments 1 and 2, Experiment 3 explored the possibility that when a violation has a different pitch and comes from a different location, it more effectively captures attention. That is, would a violation defined by both a change in frequency *and* a change in location capture attention to a greater degree? The results of Experiment 3 do indicate that a conjunction violation captures attention. However it does not appear that a violation defined by both frequency and location captures attention more efficiently than a violation defined along a single dimension alone.

Research has established that auditory attention can be voluntarily oriented to a particular frequency or location (Cherry, 1953; Moray, 1959; Treisman, 1964). Much of the early research on selective attention used a dichotic listening technique. An experiment by Cherry (1953) demonstrated that attention could be voluntarily oriented on the basis of spatial location. Aside from voluntarily orienting attention based on spatial location, people can also selectively attend to a specific frequency based on frequency differences between attended and unattended messages (Treisman, 1964; Underwood & Moray, 1971). This research suggested that, when instructed to do so, listeners can easily attend to one spatial location or frequency and effectively ignore irrelevant information (e.g. Cherry, 1953; Deutsch, 1986; see Pashler, 1999, for review).

It was observed, however, that not all irrelevant information was effectively ignored. For example, the listener's own name was often recognized, even though it was presented in the unattended channel (known as the 'Cocktail Party Effect'; Cherry, 1953). Similar to attentional capture by an irrelevant sound, why is it that although attention is being voluntarily oriented to a specific task, certain irrelevant items like a violation or the listeners own name attract attention? Why is voluntary control of attention interrupted by the presence of an irrelevant distractor? The present study on attentional capture provides a starting point in answering such questions.

While attentional capture has been well-studied in the visual modality, less is known about how stimuli attract attention in the auditory domain. In an auditory sequence, if an abrupt unexpected distractor sound is presented, involuntary orienting of attention is said to occur if performance on a target detection task is impaired despite the fact that the singleton is inconsistent with the task. Taken together, the results of the three experiments in the present study provide evidence of attentional capture in the auditory domain. Each of the experiments found significant impairments in the target detection task due to the presence of an irrelevant, unexpected sound. A question that arises from the current research is how to account for attentional capture in audition?

With respect to what causes attentional capture in audition, there are two basic possibilities. The first of these is that attentional capture occurs because attention is drawn to the frequency or location of the new sound. This orienting of attention is automatic. The resulting poor performance in detecting the target occurs as a result of an attentional deficit. That is, attentional resources are allocated to the frequency or location of the violation. This idea that attention is drawn to the frequency or location of a new sound can be explained by the fact that a function of the auditory system, as opposed to the visual system, is to act as an 'early warning' system (Scharf, 1998), and therefore the auditory system is more sensitive to changes. Our attention is drawn to the frequency or location of a new sound because such stimuli could signal an important change in the environment that may require immediate action. For example, with respect to a listener's own name, often somebody calls out your name to get your attention. Sometimes it is to warn you of something, other times it is to just say "hi". In any case, we often react when we hear our name and although it is usually irrelevant to what we are doing at the time, attending to our own name is important because it usually requires action on our part. In addition, irrelevant stimuli like a frequency violation may indicate an important change in the environment and therefore, should be attended to. For example, many abrupt, unexpected sounds in our environment also signal that action may be required. For example, obvious sounds such as sirens or alarms, when heard, require immediate action to be taken. Other sounds such as an alarm clock or a car honking also signal that an immediate response may be required. Since one of the functions of the auditory system is to act as an early warning system, any sound that signals something potentially important is automatically attended, even if it is irrelevant to the task at hand. Attentional capture by an irrelevant violation sound provides an indication of the extent to which auditory search is interrupted by the presence of an irrelevant sound. If the auditory system acts as an early warning system, then you would expect that any unique distractor sounds would capture attention, even if they are irrelevant, in order for the auditory system to rule out the possibility that immediate action may be required.

An alternative account of auditory attentional capture comes from the literature on the attentional blink (Chun & Potter, 1995; Jolicoeur & Dell'Acqua, 1998; Mondor, 1998; Shapiro, Raymond, & Arnell, 1994; Shen & Mondor, submitted; Ward, Duncan, & Shapiro, 1996). In these studies, listeners are presented with a rapid serial auditory presentation (RSAP) of sounds. Within the sequence of stimuli, listeners are presented with two critical sounds, a target and a probe sound. Following the sequence, listeners are required to make a judgement about both the target and the probe. Response accuracy to the probe depends on its relative position to the target sound. When the probe is presented immediately following the target, detection accuracy is significantly impaired. This processing deficit is referred to as an "attentional blink". In one study by Shen and Mondor (in press), an attentional blink was observed when the probe was presented in one of the positions immediately following the probe. Of interest, however, were the results obtained in the condition where listeners were required to ignore the target but still make a detection judgement to the probe. They found a significant attentional blink, even when the target was to be ignored. While there is a probe processing deficit when listeners are required to respond to both the target and probe sounds, this deficit exists even when the listeners are instructed to ignore the target. These results are similar to attentional capture in that attention is drawn to the frequency or location of the new sound, even when it is irrelevant to the task, and perceptual mechanisms are engaged to process it. In other words, there are limited attentional resources that are allocated to the new sound at the expense of subsequent sounds (see Di Lollo, Kawahara, Ghorashi & Enns, 2005 for review). There is a processing deficit such that sounds that follow the violation can only be processed once processing of the violation has been completed. The magnitude of the target detection deficit decreases as the interval between the violation and the presentation of the target increases.

When reviewing the literature, attentional capture is said to occur when attention is drawn away from the location or frequency of a target, thereby interfering with target detection. Attentional capture in vision also has the characteristic that it can facilitate performance if it coincides with the target. Research suggests that attention is not always drawn away from the target and to an irrelevant item, but sometimes attention is drawn toward the irrelevant item and subsequently toward the target as well. For example, Jonides and Yantis (1988) examined the effects of presenting a feature singleton that coincided with the target. They observed that interference occurred if the irrelevant item coincided with a distractor and attracted attention away from the target. In contrast, search was facilitated when the irrelevant item coincided with the target, thereby attracting attention toward the target. Studies that demonstrate that the same irrelevant sound can cause either facilitation or inhibition point to the idea that it is an attentional deficit and not limited attentional resources that causes attentional capture. If limited resources play a role in attentional capture, then one would not observe a facilitative effect. When attention is drawn automatically to the frequency or location of a new sound, any subsequent sound, whether it coincides with the violation or not, will be processed only once processing of the violation has been completed. The target sound will be held in memory until it can be processed, however items following the target can over-write the target. This would result in impaired target detection performance, regardless of whether the violation has an inhibitory or facilitative effect.

Impaired target detection performance seems likely to arise because of attention being drawn to the frequency or location of a new sound. When performance is impaired, attention is automatically drawn away from the task towards an irrelevant violation. If the violation has a facilitative effect, attention is drawn to the frequency or location of the target, resulting in improved performance on the target detection task. Taken together, evidence from the current experiment, as well as studies demonstrating a facilitative effect, seem to provide evidence that attentional capture arises because of attention being drawn to the new sound, and not as the result of limited attentional resources. However, if auditory attentional capture does occur because attention is drawn to a new sound, then one could argue that any abrupt-onset sound should be capable of capturing attention. This raises the possibility that each distractor sound also served to capture attention. However, since all the distractor sounds in the present study were identical and simply repeated themselves regularly throughout the sequence, they may have been treated as repetitions of one sound and not as new events. In this case, a sound could be considered new if it was different from the regular pattern of the sequence (in terms of frequency or location). In the current study, the effect of the distractors was constant so there is no way to know its effects on performance. However, it seems likely that the violation essentially interrupted the regular pattern of frequency or location, it was treated as a new sound and attention was drawn to it.

Attentional capture in audition is similar to attentional capture in vision in a number of ways. However, at the same time, auditory attentional capture also appears very different. In vision, attentional capture appears to be contingent on attentional set (Folk et al., 1992). That is, an object will only capture attention if it is consistent with the task at hand.

In contrast, some studies in vision have found that attentional capture may possibly be non-contingent. For example, attentional capture can also occur due to the presence of a feature singleton. This is a nontarget that has a unique feature making it stand out from the rest of the nontargets. An example in vision would be a red nontarget among green nontargets (e.g., Theeuwes, 1992). Even though nontargets in a visual search task can be ignored since they are irrelevant, research indicates that feature singletons capture attention and interfere with performance. This occurs even though the singleton is irrelevant to the task at hand which suggests that attention is shifted involuntarily to the singleton. There is no incentive for the participant to pay attention to it yet it still interferes with target detection. While attentional capture in vision appears to be contingent on attentional set, in audition a sound is capable of capturing attention even when it is inconsistent with the task.

The current study further supports this idea of non-contingent attentional capture. The results of the current study demonstrate that a sound from an unexpected and irrelevant frequency is capable of capturing attention in an involuntary manner, even though the frequency violation is inconsistent with the target detection task. Not only is a frequency violation capable of capturing auditory attention, a location violation and a conjunction violation capture attention as well. In each case, the violation is non-contingent on the attentional set, such that voluntarily orienting of attention is disrupted and attention is deployed to the frequency or location of the violation sound.

Attentional capture in vision is said to occur only as a result of abrupt visual onsets or the appearance of a new perceptual object. It seems plausible that such objects can cause attentional capture. A unique object can attract attention because it is different and may indicate a change in the environment that is important and requires action. In the rapid serial auditory presentation that was used in the current study, such an explanation of attentional capture is likely since the violation in the sequences was an abrupt, unexpected sound. Listeners had no incentive to pay attention to the violation since it was irrelevant to the task but the unexpected sound captured attention regardless. The fact that a violation draws attention to an irrelevant frequency or location is likely the result of the auditory system acting as an early warning system which evaluates whether sounds are important or to be ignored. A similar warning system may act in vision with abrupt visual onsets.

In vision, a critical factor that determines whether or not a particular singleton will capture attention is the salience of the singleton compared to the target. For example, Theeuwes (1992) found that in a visual search for an odd-shaped object, colour singletons attract attention and interfere with voluntary search. However, when searching for an odd colour, a singleton defined by shape may not interfere with search. Although, if the shape singleton is made so that it is more salient (more discriminable) than the colour singleton, shape singletons now capture attention while colour singletons do not capture attention. Similarly, Pashler (1988) found that when searching for an odd shape, the presence of a colour singleton in the visual array captures visual attention. However, if the elements in the array vary in colour, the variation is not enough to capture attention. It appears that relative salience may be an important factor that determines whether a particular singleton will capture attention. In the current study, the violation captured attention in all three experiments, regardless of whether it was defined by frequency, location, or both. Experiment 3 evaluated whether a violation defined by both frequency and location captured attention to a greater degree than a violation defined along a single dimension. It was expected that a conjunction violation would capture attention more efficiently as a result of it being more salient than a violation defined by frequency or location alone. The fact that a conjunction violation did not capture attention to a greater degree can be explained by the idea that one of the functions of the auditory system is to act as an early warning system and that any sound that signals something potentially important is automatically attended to. If the auditory system acts as an early warning system, then you would expect that any unique distractor sounds would capture attention, even if they are irrelevant and there is no incentive to pay attention to them. In the current study, any violation was prone to capture attention, and it appears salience may not play the role it does in vision or it may be that a conjunction violation is not sufficiently salient. Any violation captured attention and it did not matter if it was more salient or not, since it still attracted attention and impaired performance. As long as the features that define the violation are different than those that define the target, the violation will capture attention. In other words, it appears that attention is captured by the occurrence of a new object and the degree to which attention is captured does not depend on the degree to which a new sound differs from sounds that are being voluntarily attended. Therefore, it appears that salience does play a role in attentional capture but it does not seem to matter how much the target and violation differ since a violation defined by both frequency and location did not capture attention to a greater magnitude

The current study had similarities to studies in vision where attention was captured by a feature singleton (a stimulus that is different along a single dimension such as colour). The task was to search for a target defined along a single dimension (e.g., low frequency). Both the target and nontarget sounds were low pitched, however an abrupt sound of an unexpected pitch would occur on some trials. Since the task was to voluntarily attend to low pitch sounds, a high pitch sound was irrelevant and should therefore be ignored. This was not the case. In all the experiments, the violation interfered with target detection even though there was no incentive for listener's to attend to it since it did not predict the target in any way. Voluntary search of low pitch sounds was interrupted by the presence of a sound of a different frequency. Attention was captured by the violation. This could be explained in terms of an attentional template (e.g., Mondor, Zatorre, and Terrio, 1998). When attending to sounds, a template is formed that holds specific information about the features to be attended. Each sound is then compared to the template and those that match are processed easily. If a sound does not match the attentional template, it interferes with search since it is not processed as easily.

Attentional capture in vision also has the characteristic that it can facilitate performance if it coincides with the target. Attentional capture is said to occur when attention is drawn away from the location or frequency of a target, thereby interfering with target detection. However, performance on a target detection task is improved if the irrelevant singleton draws attention to its frequency or location if it is the same frequency or location of the target. For example, if the feature singleton is a sound presented from the left, attention is drawn to that spatial area. If a target sound is presented in the same location immediately following the distractor sound, target detection will be facilitated. An example of this facilitative effect was observed by Jonides and Yantis (1988). They observed that interference occurred if the irrelevant item coincided with a distractor and attracted attention away from the target. In contrast, search was facilitated when the irrelevant item coincided with the target, thereby attracting attention toward the target. Dalton and Lavie (2004) also observed that the same singleton is capable of either facilitating performance or impairing it, depending on whether it coincided with the target or a distractor.

The results of studies that show that the same irrelevant sound can cause either facilitation or inhibition provide evidence that it is not some general property of the singleton that causes attentional capture. It is the irrelevant object itself that is responsible for capturing attention and not the fact that it has a lower probability of occurring than the other nontarget sounds. An example of an irrelevant object having a lower probability is if all of the distractor sounds are high-pitched and there is one low-pitched sound in the sequence. This low-pitch sound has a much lower probability of occurring than a high-pitched sound. Attentional capture is therefore the result of the singleton itself.

While the current experiments did not investigate a facilitative effect, it seems likely that this would occur. For example, in a cue-target study, Gibson and Egeth (1994) found that a cue captures attention so that the perception of a target presented subsequently in the same location is improved, relative to perception of a target appearing in a different location. In addition, it has been established that in a search for a sound of a low frequency in a sequence, a high-pitched violation captures attention. If a target of a low frequency is presented immediately following the frequency violation, target detection is significantly impaired as a result of attention being drawn to the high-pitched frequency violation and away from the frequency of the target. If the current study were modified so the target was a high-pitch sound embedded in a sequence of low pitch distractor sounds and the frequency violation remained a high pitch sound. It would be highly likely that if attention is drawn to the high-pitched violation, a target that is also high-pitched presented immediately following the violation would be detected quickly. When attention is attracted to the frequency region of the violation sound, a target that is presented in the same frequency region immediately following the violation would be detected easily. Taken together, the facilitation and inhibitory effects of the same singleton strengthen the claim that it is the irrelevant singleton sound that captures attention.

The findings of the current study are consistent with the results of visual search studies that have demonstrated impaired performance (or facilitation effects) as the result of attentional capture by an irrelevant singleton (e.g., Jonides & Yantis, 1988; Pashler, 1988; Theeuwes, 1992).

While attentional capture in vision appears to be contingent on attentional set, the results of the current experiments demonstrate that this is not the case with auditory attentional capture. Although the violation was irrelevant to the task, attention was still drawn to it and performance on the target detection task was impaired. The target and distractor sounds were defined along a specific dimension and the violation was defined along a different dimension making it irrelevant to the task. There was no incentive for the listeners to attend to the violations since they did not predict the target in any way, making the violations irrelevant.

The results of the present study show that voluntary orienting is not always successful and that target detection can be impaired by the occurrence of an unexpected sound defined by frequency, location, or both. This involuntary orienting of attention as a result of the violation occurs even though the violation is completely irrelevant to the task at hand, suggesting that auditory attentional capture is noncontingent. The results establish that the involuntary effect from an unexpected violation can interfere with a listener's ability to detect a target, despite the voluntary control of the listener's goals.

Target detection is impaired when the target immediately follows the violation, suggesting that auditory attentional capture is involuntary.

Not only do the results of the current research provide information about how selective attention works, it can be used as a basis for further research in the area of how attention is captured and such things as warning alarms. If researchers can establish effective ways of capturing auditory attention, design of such things as warning alarms (used in hospitals, etc.) can be improved so that they will capture attention most effectively. This would result in alarms being responded to more quickly and accurately (e.g. Mondor & Finley, 2003).

Although the current research focused on attentional capture in the auditory domain, auditory attentional capture shares characteristics with its visual counterpart. Attention can be captured by a sound that is unexpected, even if the sound is completely irrelevant to the task, resulting in impaired performance on a target detection task. Since a role of the auditory system is to act as an early warning system, any sound that is unexpected captures attention and salience does not appear to matter. As long as the auditory system deems that the unexpected sound could be potentially important, attention will be involuntarily oriented to the frequency or location of the sound, even if it means task performance will be temporarily impaired.

References

Arnell, K. M., & Jolicoeur, P. (1999). The attentional blink across stimulus modalities: Evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception & Performance, 25(3),* 630-648.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975-979.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109–127.

Coull, J. T., Frith, C. D., Büchel, C., & Nobre, A. C. (2000). Orienting attention in time: behavioural and neuroanatomical distinction between exogenous and endogenous shifts. *Neuropsychologia, 38,* 808-819.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience, 18,* 7426-7435.

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance, 30(1)*, 180-193.

Di Lollo, V., Kawahara, J., Ghorashi, S., & Enns, J. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research, 69(3),* 191-200.

Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance, 14(2),* 188-202.

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18(2)*, 578-588.

Dysan, B., & Quinlan, P. (2003). Feature and conjunction processing in the auditory modality. *Perception & Psychophysics, 65(2),*254-272.

Egeth, H. E., Folk, C. L., Leber, A. B., Nakama, T., & Hendel, S. K. (2001). Attentional capture in the spatial and temporal domains. In C. L. Folk and B. S. Gibson (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture. Advances in psychology, 133.* (pp. 191-229). New York, NY: Elsevier Science.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269-297.

Enns, J. T., Austen, E., Di Lollo, V., Rauschenberger, R., & Yantis, S. (2001). New objects dominate luminance transients in setting attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1287-1302.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030-1044.

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and colour. *Journal of Experimental Psychology: Human Perception and Performance, 20(2),* 317-329.

Gellatly, A. R. H., Cole, G. G., & Blurton, A. (1999). Do equiluminant object onsets capture visual attention? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1609-1624.

Gibson, B. S. (1996). Visual quality and attentional capture: A challenge to the special role of abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 22(6),* 1496-1504.

Gibson, B. S., & Amelio, J. L. (2000). Inhibition of return and attentional control settings. *Perception & Psychophysics, 62,* 496-504.

Gibson, B. S., & Egeth, H. (1994). Inhibition and disinhibition of return: Evidence from temporal order judgments. *Perception & Psychophysics, 56,* 669-680.

Gibson, B. S., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics, 55,* 323-339.

Greenberg, G., & Larkin, W. (1968). Frequency response characteristics of auditory observers detecting signals of a single frequency in noise: The probe signal method. *Journal of the Acoustical Society of America, 44*, 1513-1523.

Hillstrom, A. P., & Yantis, S. (1994). Visual motion and attentional capture. *Perception & Psychophysics, 55,* 399-411.

Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, *36*, 138–202.

Jones, M. R. (2001). Temporal expectancies, capture, and timing in auditory sequences. In C. L. Folk and B. S. Gibson (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture. Advances in psychology, 133.* (pp. 191-229). New York, NY: Elsevier Science.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 187-203). Hillsdale, NJ: Erlbaum.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43,* 346-354.

Kurylo, D. D., Reeves, A., Scharf, B. (1996). Expectancy of line segment orientation. *Spatial Vision, 10(2),* 149-162.

Miniussi, C., Wilding, E. L., Coull, J. T., & Nobre, A. C. (1999). Orienting attention in time: modulation of brain potentials. *Brain, 122,* 1507-1518.

Mondor, T. A. (1999). Predictability of the cue-target relation and the time-course of auditory inhibition of return. *Perception & Psychophysics, 61(8),* 1501-1509.

Mondor, T. A. (1998). A transient processing deficit following selection of an auditory target. *Psychonomic Bulletin and Review, 5(2)*, 305-311.

Mondor, T. A., & Amirault, K. J. (1998). Effect of same- and different-modality spatial cues on auditory and visual target identification. *Journal of Experimental Psychology: Human Perception and Performance, 24(3)*, 745-755.

Mondor, T. A., & Breau, L. M. (1999). Facilitation and inhibitory effects of location and frequency cues: Evidence of a modulation in perceptual sensitivity. *Perception & Psychophysics, 61(3),* 438-444.

Mondor, T. A., Breau, L. M., & Milliken, B. (1998). Inhibitory processes in auditory selective attention: Evidence of location-based and frequency-based inhibition of return. *Perception & Psychophysics, 60(2),* 296-302.

Mondor, T. A., & Bregman, A. S. (1994). Allocating attention to frequency regions. *Perception & Psychophysics, 56,* 268-276.

Mondor, T. A., & Finley, G. A. (2003). The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate. *Canadian Journal of Anesthesia, 50(3)*, 221-228.

Mondor, T. A., Terrio, N. A., & Hurlburt, J. (2000). On the role of eye movements and saccade preparation in generating location-based auditory inhibition of return. *Canadian Journal of Experimental Psychology, 54,* 326-338.

Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 66-79.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56-60.

Näätänen, R., Porkka, R., Merisalo, A., & Ahtola, S. (1980). Location versus frequency of pure tones as a basis for fast discrimination. *Acta Psychologica, 44,* 31-40.

Pashler, H. E. (1998). *The Psychology of Attention.* Cambridge, MA: MIT Press.

Pashler, H. E. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics, 43*, 307-318.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32(1),* 3-25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of Language Processes* (pp 531-556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: the role of set for spatial location. In H. L. Pick & E. J. Saltzman (Eds.), *Model of Perceiving and Processing Information.* Hillsdale, NJ: Erlbaum.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. Rabbitt & S. Dornic (Eds.), *Attention and performance. Vol. 5* (pp.669-682). San Diego, CA: Academic Press.

Psychology Software Tools, Inc. (1999). E-Prime Software System. Pittsburgh, PA.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: Attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18 (3),* 849-860.

Reuter-Lorenz, P. A., Jha, A., & Rosenquist, J. N. (1996). What is inhibited in inhibition of return? *Journal of Experimental Psychology: Human Perception and Performance, 22,* 367-378.

Ruz, M., & Lupianez, J. (2002). A review of attentional capture: On its automaticity and sensitivity to endogenous control. *Psicologica, 23(2),* 283-309.

Scharf, B. (1998). Auditory attention: The psychoacoustical approach. In H. Pashler (ed.) *Attention.* Hove, United Kingdom: Psychology Press.

Schmidt, W. C. (1996). Inhibition of return without visual input. *Neuropsychologia, 34,* 943-952.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to visual pattern information produces the attentional blink in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 357–371.

Shen, D., & Mondor, T. (2005). Effect of distractors on the auditory attentional blink. *(submitted).*

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20,*555-574.

Syntrillium Software Corporation. (2000). Cool Edit Software System. Scottsdale, AZ.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics, 50,* 184-193.

Theeuwes, J. (1992). Perceptual selectivity for colour and form. *Perception & Psychophysics, 51*, 599-606.

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 478-499.

Treisman, A. M. (1964). Effect of irrelevant material on the efficiency of selective listening. *American Journal of Psychology, 77(4),* 533-546.

Underwood, G., & Moray, N. (1971). Shadowing and monitoring for selective attention. *Quarterly Journal of Experimental Psychology, 23(3),* 284-295.

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, *30*, 79–109.

Woods, D. L., Alain, C., Diaz, R., Rhodes, D., & Ogawa, K. H. (2001). Location and frequency cues in auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 65-74.

Woods, D. L., Alain, C., & Ogawa, K. H. (1998). Conjoining auditory and visual features during high-rate serial presentation: Processing and conjoining features can be faster than processing one. *Perception & Psychophysics, 60,* 239-249.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 676-681.

Yantis, S. (1996). Attentional capture in vision. In A. Kramer, M. Coles, & G. Logan (Eds.), *Converging operations in the study of selective visual attention* (pp. 45-76). Washington, DC: American Psychological Association.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 95-107.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601-621.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 121-134.

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt visual onsets: New perceptual objects or visual masking? *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1505-1513.