



Universiteit Twente

Does Sound Affect Visual Motion Change Detection?

Bachelorthesis of Sabine Michaela Staufenbiel

University of Twente, The Netherlands

August 28th, 2009

First Supervisor: Dr. Durk Talsma

Second Supervisor: Dr. Rob van der Lubbe

Abstract

The present study shows that the “Pip and Pop” effect, which is an illusion of a visual target change popping out of an array with distractor stimuli when the target change is accompanied by a short auditory sound, can be applied to a search task regarding moving stimuli. Until now, this effect had only been demonstrated in recent research using a search task regarding the orientation of stimuli. A short, static sound with no information on the location or identity of the target, presented at the moment at which one moving stimulus among moving distractor stimuli changes direction makes the target pop out, resulting in a more efficient search in sound present trials in comparison to performance in sound absent trials. Experiment 1 showed that participants were able to observe a higher number of objects in the sound present condition than in the sound absent condition while still answering correctly. This conclusion was drawn by means of a tracking algorithm adjusting the dynamic number of stimuli displayed to the performance of each participant. In Experiment 2, participants’ individual mean of number of objects were computed in a dynamic test phase, using a tracking algorithm to adjust the number of objects to the individual performance. These numbers of objects were used in a static test phase, demonstrating that participants scored higher on accuracy and sensitivity in the sound present trials than in the sound absent trials.

Introduction

In our everyday lives we are surrounded by auditory, visual, olfactory, gustatory and tactile stimuli almost every time. The stimuli provided by any one sense in isolation often do not provide sufficient information about the environment, but in combination they give meaning to the different inputs one perceives. The crossmodal integration of the single senses and stimuli to one big picture of perception is a process that occupies researchers ever since.

Recent research (e.g. Talsma, Doty & Woldorff, 2007) has shown that cognitive processes, i.e. attention, interact with the integration of audiovisual processes. On the one hand, there is evidence that attention is guided by audiovisual integration in an exogenous manner (Van der Burg, Olivers, Bronkhorst & Theeuwes, 2008a), thus automatically, via bottom-up processing, which happens when the mental processing of a stimulus is guided by the features and elements of the stimulus. On the other hand, there is evidence that attention guides audiovisual integration in an endogenous manner (Mozolic, Hugenschmidt, Pfeiffer & Laurienti, 2008), in other words, via top-down processes. Top-down processing means that attentional control can be applied to the mental processing as that knowledge about which features of stimuli are relevant to the task determines which objects are selected. For example, as Sekuler and Ball (1977) demonstrated, the mental set, thus foreknowledge about the character of an upcoming stimulus, affects visibility directly. They showed in their experiment that moving targets are easier to detect when the participant knows which speed and direction are to expect.

Leber and Egeth (2004) provided evidence that top-down search strategies can interact with bottom-up signals by hindering attentional capture in a search paradigm and thus influencing the natural competition between the different salient objects.

Exploring the role of attention in audiovisual integration, Fujisaki, Koene, Arnold, Johnston and Nishida (2006) conducted an experiment regarding the question whether audiovisual synchrony is determined by a serial or a parallel search process. Their participants had to detect which visual stimuli out of a number of stimuli changed in synchrony with an auditory stimulus. The authors used differing set sizes of visual stimuli to investigate how search performance changed with the number of distractor stimuli. A serial search implies that the detection of audiovisual synchrony is determined by an attentive process. Contrarily, a parallel search implies that audiovisual synchrony detection is a pre-attentive process and the visual target would “pop out” in case of an audiovisual synchrony.

Does Sound Affect Visual Motion Change Detection?

Fujisaki's et al. findings contributed to evidence pointing towards the serial search, leading to the conclusion that human beings can only check one or a small number of signals at a time for possible audiovisual pairings. However, as Van der Burg, Olivers, Bronkhorst and Theeuwes (2008b) argued, this experiment had several shortcomings regarding the event of audiovisual integration. For instance, Fujisaki et al. designed an experiment in which the sound, the distractors and the target changed continuously and frequently, thus providing several sources of confusion for the participants. Van der Burg et al. (2008b) pointed out that these factors combined with the high range of the used modulation frequencies (up to 40 Hz) did not result in optimal circumstances for an occurrence of spatiotemporal audiovisual integration, as previous research has shown that multisensory integration becomes difficult at temporal frequencies higher than 4 Hz (see Fujisaki & Nishida, 2005; Lewald, Ehrenstein & Guski, 2001). In the experiment of Van der Burg et al. (2008b), in which the single auditory-visual events were temporally isolated, the pop out did take place, indicating a parallel search.

Similarly as Van der Burg et al. (2008b), Vroomen and de Gelder (2000) reported the discovery of the "freezing phenomenon". They stated that this phenomenon is an illusion that occurs when during a rapidly changing visual display a short sound is presented. This results in an illusory "freezing" of the visual display for a short moment.

Fujisaki et al. (2006) pressed that audiovisual synchrony is neither entirely generated by bottom-up or by top-down attentive processes but detected by a general-purpose mid-level perceptual mechanism that compares the salient features that are delivered by each modality-specific stream of information. In line with that, McDonald, Teder-Sälejärvi and Hillyard (2000) demonstrated that involuntary orienting of attention to the location of a nonpredictive sound can improve the perceptual quality of a subsequent visual stimulus at early perceptual levels as well as at later, decision-related levels.

Whatever the cause of this effect, the conclusion drawn is that a sound seems able to enhance the detectability of a visual target.

In many studies which reported this occurrence, the beneficial effects of audiovisual integration were found solely when the auditory stimulus came from the same location as the visual target. Spatial disparity between the auditory and visual stimulus could not only weaken the effect or result in a loss of the capacity to enhance the visual stimulus, but could even worsen performance (McDonald, Teder-Sälejärvi & Hillyard, 2000; Perrott, Saberi, Brown & Strybel, 1990; Perrott, Sadralodabai, Saberi & Strybel, 1991; Spence & Driver, 1997). Contrarily, Stein, London, Wilkinson and Prince (1996) demonstrated in an

Does Sound Affect Visual Motion Change Detection?

experiment that a short auditory stimulus can boost the perceived intensity of a concurrent flash of light with spatially correlated and spatial disparate stimuli. In line with that, Van der Burg et al. (2008b) demonstrated that a sound from a spatially disparate location can, despite the distance, enhance visual search detection performance.

In the above mentioned experiment conducted by Van der Burg et al. (2008b) on which the present study will be based primarily, the authors discovered a phenomenon that they named the “pip and pop effect”. They showed that an auditory signal can raise the saliency of a visual change, which results in the impression that the visual change pops out. Their findings are based on several experiments. The common task of the participants was to detect a horizontal or a vertical line segment (the target), among up to 48 distractor line segments of various orientations. All stimuli constantly changed color between red and green, which made the search task more difficult. When the target changed color, on average once every 900ms, it always did so alone, thus without simultaneously changing distractor stimuli.

On one half of the trials, there was no auditory signal, on the other half the target change was accompanied by a short auditory pip. The pip was unpredictable and uninformative about the orientation, location or color of the target, thus informed only about the moment of color change. The participants had to respond as fast as possible by pressing a button which indicated the orientation of the target. In the sound present conditions, the participants responded faster and the search was more efficient than in the sound absent condition.

The authors concluded that the observed benefits in visual search are due to successful binding of the auditory signal with the visual target event. By means of follow-up experiments they ruled out the possibilities that the benefits are due to the auditory signal serving as a mere cue or an alerting signal, showing that the pip and pop effect is not due to temporal top-down knowledge. Apparently, the integration of the synchronous auditory and visual stimuli occurred largely automatically, with the sound guiding attention towards the visual target location even when there was no or little strategic reason to do so.

Van der Burg, Olivers, Bronkhorst, Talsma and Theeuwes (2008) reported a follow-up experiment in the Object Perception, Attention, and Memory 2008 Conference Report (OPAM 2008 report), in which it was found that while the pip and pop effect is largely automatically evoked, there is still evidence that the effect can be influenced by top-down processes.

Aside from behavioral data, as Van der Burg et al. (2008b) summarized, there is also overwhelming evidence of early (Falchier, Clavagnier, Barone & Kennedy, 2002; Giard &

Does Sound Affect Visual Motion Change Detection?

Peronnet, 1999; Molholm et al., 2002; Talsma, Doty & Woldorff, 2007) and effortless (Vroomen & De Gelder, 2000) audiovisual integration in the neuronal system. Additionally, they summarized other studies which demonstrated that an auditory signal can enhance the saliency of a concurrently presented visual target by demonstrating multisensory convergence in low level sensory cortical structures (Schroeder & Foxe, 2005). They cite e.g. findings of Molholm et al. (2002) who stated that auditory activation can be observed in the extrastriate visual cortex. In line with that, Middlebrooks and Knudsen (1984) showed that in multimodal neurons, the auditory receptive fields are larger than visual receptive fields, therefore can excite neurons over a larger region than visual stimuli.

Frassinetti, Bolognini and Làdavas (2002) showed the existence of an integrated visuo-acoustic system in humans in demonstrating enhanced sensitivity in a visual detection task due to auditory stimuli. They found this enhancement in spatial correlated as in spatial disparate auditory and visual stimuli.

These experiments all support the idea that auditory stimuli can affect visual processing quickly and that visual processing is modified by auditory inputs before it is completed. As Molholm et al. (2002) showed, the auditory signal is rapidly passed to the early visual cortex, resulting in the possibility of interaction between auditory stimuli and synchronously processed visual stimuli (Giard & Peronnet, 1999). Hence, as Van der Burg et al. (2008b) summarize, the sound has a modulation function across the visual cortex in increasing visual signals that must be already present but are alone by themselves not strong enough to claim priority for selection.

Similarly, as cited in Calvert (2001), Stein and Meredith (1993) developed the *inverse effectiveness rule*, which states that when two unimodal stimuli each by themselves evoke no response, the combination of both results in an effective enhancement of the saliency and produces a powerful response.

Girelli and Luck demonstrated in 1997 that during the detection of visual search targets that are defined by motion, color and orientation, a common attentional mechanism is activated. The authors pointed out that this finding is consistent with the known anatomy and physiology of the visual system. They provided evidence that the process of selecting a target from an array of distractors activates the same attentional subsystem, independent of the feature that defines the target, as motion, color or orientation. This finding replicates the general view, namely, as Goodale and Milner (1992) state it, that “the properties of the individual objects are analyzed by a common set of structures within the ventral pathway, independent of the specific features being processed”. Nevertheless, even if there might be a

Does Sound Affect Visual Motion Change Detection?

common process of selecting a target across feature types being applied, Girelli and Luck (1997) further found that only motion singletons elicited an automatic attentional response. They suggested that the bottom-up control of attention might be different because of physical features of the stimulus.

Regarding neural structures, Vroomen en de Gelder (2000) reported on the neural peculiarities of multimodal convergence and integration. They highlighted the superior colliculus, a midbrain structure known to play an important role in attentive and orientive behavior, as one locus in the brain where the converging and integrating processes take place. In a review, Burr and Alais (2006) stated that the superior colliculus has strong reciprocal links with the middle-temporal (MT) cortical area (Standage & Benevento, 1983), which is an area specialized for the processing of visual movement and whose activity is strongly correlated with visual motion perception (Britten, Shadlen, Newsome & Movshon, 1992). Alais and Burr further cite several authors who showed that MT outputs project directly to area VIP where they combine with input from auditory areas to create bimodal cells with strong motion selectivity (Bremmer et al., 2001; Graziano, 2001).

Taking into account these audiovisual- and motion-related findings, the question arises whether the conclusions reported by Van der Burg. et al. (2008b) drawn from their experiment (i.e. that when accompanied by a spatially disparate, static auditory pip, singleton color changes in an otherwise static display are better perceived than without an auditory cue) can be applied to moving stimuli. This is an important question, as it provides further insight into the topic of how our attention system works in the complex dynamic of our everyday life. If, as Girelli and Luck (1997) stated, the same attentional mechanisms are applied in orientation, color and motion detection, and that there is even evidence pointing towards a neural connection of multimodal integration areas and motion perception areas (Standage & Benevento, 1983; Britten, Shadlen, Newsome & Movshon, 1992), then the results of Van der Burg et al. (2008b) should be replicable with motion change detection tasks.

This question will be addressed by means of a computer task, during which the participants have to attend to an array of in random directions moving stimuli on a computer display in two conditions, i.e. a sound present condition and a sound absent condition. The participants are instructed to focus on the visual stimuli and to neglect the auditory stimuli if and when they occur. Their task is to detect one stimulus that at a given time changes the direction of movement. The number of stimuli changes due to a tracking algorithm. Hereby, the number of displayed stimuli increases or decreases depending on the performance of each participant to ensure a stable accuracy of 80 percent for each participant, thus providing the

Does Sound Affect Visual Motion Change Detection?

participants with a differing amount of visual load until the maximal number of objects for each participant is computed.

Based on the findings of the cited experiments, it is expected that in the sound present condition, the participants will be confronted with a higher number of objects, as with being more accurate the number of objects on the screen increases, than in the sound absent condition.

Thus, in other words, it is hypothesized that in the sound present condition, participants are capable to still answer correctly with a stronger competition of stimuli than in the sound absent condition.

Methods Experiment 1

Participants. Fourteen participants, eleven females and three males, took part in Experiment 1. The mean age was 19.2 with a range from 18 – 24 years. Subjects were students at the University of Twente and enrolled via sona-systems, the sampling pool for participants of this university. They received one credit point each for their participation. Inclusion criteria for the participants were normal hearing functions, normal or corrected-to-normal vision and not having color blindness. All gave informed consent and were naïve as to the purpose of the experiment.

Procedure/Task. The experiment was conducted in a university laboratory in a closed room with regular artificial lighting. The participants were seated in front of a table on which the apparatus was placed. On the screen of a computer, in random directions moving red and green dots, with an even number regarding color, were presented on a black background. One of the dots changed direction at an angle of 90° during the trial. In fifty percent of the trials, short sounds were presented exactly at the moment of change. In the sound absent condition, participants were instructed to search for the one dot changing direction and to respond as accurately as possible to its color. In the sound present condition, the task was the same, but the change was accompanied by a short auditory pip. Notably, this sound provided no information about the location, the color, or the direction of the dot, only about the moment of direction change. The participants had to press certain keys for either color to indicate of which color the changed dot was. While they had to press ‘G’ (for “groen”, the Dutch

Does Sound Affect Visual Motion Change Detection?

equivalent for “green”) for a green dot, they had to press ‘R’ (for “rood”, the Dutch equivalent for “red”) when it was a red dot. The number of dots presented was dynamic. Via a tracking algorithm, the number of presented dots was adjusted to the capacity of each participant during the experiment, such that each participant would be able to maintain an accuracy of 80%. Participants were constantly monitored via closed-circuit camera installed behind the participants’ back.

Participants were instructed to press the correct associated key for an answer when, as after each trial, they were asked on the screen of which color the dot was that changed direction. Immediately after the releasing of the key, the following trial was presented.

It was not possible to predict which dot would change direction. After each trial, the participants received feedback in Dutch about their answer and the number of remaining trials. Then the participants could continue by pressing a random key.

Practice Phase. At the beginning of the practice phase, participants received instruction in Dutch by the researcher and via the computer screen. The practice phase consisted of four blocks, two with sound and two without sound. The order of blocks was distributed randomly. Each block contained twenty trials. The first trials in each condition contained ten dots. Due to the tracking algorithm, whenever a wrong answer was given, the following trial consisted of two stimuli less and whenever five correct answers in a row were given, six objects were added at the following trial. The number of present dots per block was transferred to the following block that belonged to the same condition, i.e. sound present or sound absent.

After each block, the participants had a short pause and could self determine when to continue with the experiment.

Test Phase. At the beginning of the test phase, participants received instruction via the computer screen. The task to be conducted and the procedure were the same as in the practice phase, except for the number of blocks. The test phase consisted of 16 blocks, of which eight were sound present and eight were sound absent.

Apparatus and stimuli. Stimulus presentation, timing, and data collection were achieved by using the E-prime© 1.1 experimental software package on a standard Pentium© IV class PC. Stimuli were presented on a 17 Inch Philips 107T5 display running at 800 by 600 pixel resolution in 32 bit color, and refreshing at a rate of 60 Hz. The viewing distance was

Does Sound Affect Visual Motion Change Detection?

approximately 60 cm, but not strictly controlled. The experiment was run in a special secured computer mode to ensure a precise response presentation and stimulus registration. Input was given by means of a standard keyboard.

Each trial consisted of at least 120 frames, each lasting 16,7 ms, before the direction change could take place. The number of frames in which the change could take place was variable. There was a likelihood of 16% in each frame that a direction change could take place. After each change, there were another 120 frames before the trial was finished.

Whenever a dot moved out of the window, it reappeared on the adverse side of the screen. This was to ensure that always a constant number of stimuli remained in the presented field. Whenever a wrong answer was given, the following trial consisted of two dots less. Whenever five correct answers were given in a row, six dots were added at the following trial.

Each dot had a radius of three pixels. The red stimuli were completely red, whereas the green stimuli were completely green. They moved with a velocity of one pixel and 0.036° per frame.

The sound used in this experiment was of 1000 Hz and lasted 15 ms, including a fade-in of 5 ms and a fade-out of 5 ms. It was presented via loudspeaker in front of the participant.

Data Analysis. The data of this experiment were merged with E-Prime E-Merge and exported with E-DataAid©. The data were then analyzed by use of SPSS 16.0 for Windows. Univariate repeated measures analyses were run on the data. The factor “sound presence” with the levels “sound present” and “sound absent” served as dependent variable, whereas “accuracy” and “number of objects” served as independent variables.

Results Experiment 1

The One-Sample Kolmogorov-Smirnov Test proved that all data were distributed normally. Thus, application of parametric multivariate measures for the analyses of the data was legitimated. Two univariate repeated measures analyses (ANOVAs) with the factor “sound presence” and the levels “sound present” and “sound absent” were run on the data, one with accuracy (in percentage of correct responses) and one with number of objects (with two decimals) as dependent variables, for each the practice phase and the test phase.

Does Sound Affect Visual Motion Change Detection?

Practice Phase. The analysis showed a mean number of 7.69 objects ($SD = 2.30$) in the sound present condition and a mean number of 8.26 objects ($SD = 2.60$) in the sound absent condition, but the difference was not significant ($F(1, 13) = 0.529, p = 0.480$).

The analysis revealed a mean accuracy of 77.3% ($SD = 0.04$) in the sound present condition and a mean accuracy of 74.6% ($SD = 0.07$) in the sound absent condition, but this effect did not reach significance ($F(1, 13) = 3.911, p = 0.07$).

Test Phase. The analysis revealed a mean number of 8.90 objects ($SD = 1.60$) in the sound present condition and a mean number of 8.22 objects ($SD = 1.67$) in the sound absent condition. The difference was not significant ($F(1, 13) = 2.896, p = 0.113$).

A mean accuracy of 76.6% ($SD = 0.02$) was found in the sound present condition and a mean accuracy of 76.8% ($SD = 0.03$) in the sound absent condition. This difference did not reach significance ($F(1, 13) = 0.042, p = 0.840$).

When only analyzing the second half of the test phase, different results emerged. While with a mean of 76.6% ($SD = 0.02$) in the sound present condition and a mean of 78.1% ($SD = 0.03$) in the sound absent condition, the difference in accuracy did not reach significance ($F(1, 13) = 2.095, p = 0.171$), the number of objects differed significantly, with a mean of 9.48 objects ($SD = 2.85$) in the sound present condition and a mean of 8.00 objects ($SD = 1.65$) in the sound absent condition ($F(1, 13) = 4.829, p = 0.047$). Pairwise comparisons revealed that participants could monitor 0.25 to 2.93 more objects in the sound present than in the sound absent condition.

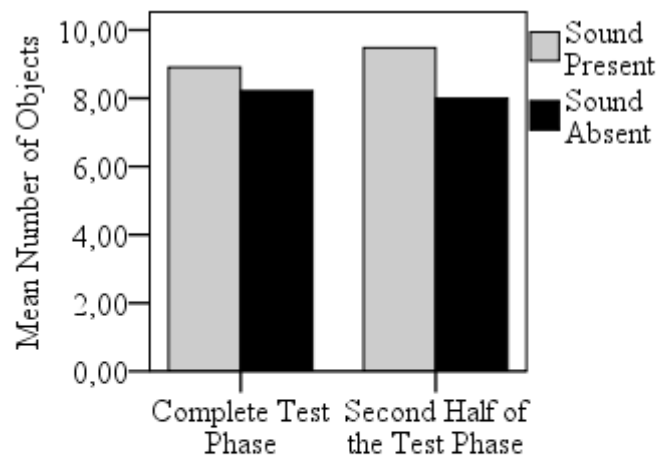


Figure 1. In Experiment 1, the mean number of objects for the sound present condition and for the sound absent condition, for the complete test phase and for only the second half of the test phase.

Discussion Experiment 1

Does Sound Affect Visual Motion Change Detection?

The results of Experiment 1 seem on the first occasion somewhat arbitrary. In the practice phase, although not significant, a trend in the reverse pattern of the expected results emerged. Participants performed more accurately and with a higher number of objects in the sound absent than in the sound present condition. This could be accounted for by a possible mental overload due to the abrupt begin of the experiment and a startling and distracting function of the sound. In the test phase, the results show a trend toward the expected direction, with a higher number of objects in the sound present condition than in the sound absent condition, but the difference was not sufficient for being significant. Notably, when taking into account only the second half of the test phase, the difference regarding the number of objects between the conditions turned significant, with 0.25 to 2.93 more objects being displayed in the sound present condition. This difference is shown in Figure 1. It thus seems that participants need more practice, probably to get familiar with the experimental procedure, before the integration can occur in an efficient way. One underlying source for the required prolonging of the practice phase may be due to the use of different colors. Taking the findings from Girelli and Luck (1997) into consideration, i.e. that the same attentional mechanisms are used to detect visual search targets that are defined by color, orientation and motion, it seems possible that it required too much mental effort to split up the one attentional resource, because the participants had to attend to motion and color.

A further explanation for the distinct trends in the different phases may lie in the application of the tracking algorithm. As the tracking algorithm is subject to random changes and initially tends to overcorrection, it needs time to stabilize the number of objects each participant can track.

The same explanation further applies to the weak correlation between the accuracy measures and the number of objects, as the accuracy can change with every trial and the number of objects can do so only in one direction (subtraction of two stimuli) and not in the other one, as five correct answers in a row had to be given before six objects were added.

These shortcomings led to the development of a follow-up experiment, in which these considerations were taken into account.

Experiment 2

Does Sound Affect Visual Motion Change Detection?

The most important correction was the modification of the color use and of the required response. Due to the possible interferences of color detection and motion detection on the attentional level, the stimuli in Experiment 2 were of only one color, i.e. white. On the response level, participants now had to indicate whether a motion change was present in the actual trial or whether it was not, as only fifty percent of all trials contained a dot changing direction, as opposed to Experiment 1 in which during every trial a direction change took place.

Additionally, a deviant static dot was used as fixation dot in the center of the screen to fixate the perceptual window, as visual search tends to be more selective and efficient when fixations were long lasting (Hooge & Erkelens, 1999). Additionally, as Neville and Lawson (1987) summarized, the peripheral retina has greater sensitivity to movement than the central area of the retina. Further on, as Leber and Egeth (2004) provided evidence that top-down search strategies can override attentional capture in a search paradigm as was used in this experiment, the fixation dot was used to prevent active and arbitrary search strategies.

Another modification was conducted on the tracking algorithm. Because in Experiment 2 no even number of colored dots had to be presented, the algorithm could be refined, with the distraction of one dot after every incorrect answer and an addition of three dots after five correct answers were given in a row.

As participants reported that the reappearance of disappearing dots at the adverse side of the display was requiring attention as to whether this is remarkable because it is a new dot or because it popped out due to a change in direction, this objection was used to eliminate this source of confusion. Objects could not move out of the display any longer but bounced back in the same speed that was present during the whole session.

Further, another allocation of experimental phases was applied. The practice phase now was of short duration and included only to let the participants get familiar with the procedure. The subsequent phase was the dynamic test phase in which due to the applied tracking algorithm the mean number of objects each participant was confronted with could be analyzed for each participant. With the knowledge of the exact number, these numbers of objects could be entered and used in the third and last phase, the static test phase. In this phase, the number of objects was static and did not change. The precise accuracy rate for the sound present condition and for the sound absent condition could then be used for computation.

Based on the mentioned adjustments, in Experiment 2 it was hypothesized that during the static test phase, participants in the sound present condition would be more accurate than

Does Sound Affect Visual Motion Change Detection?

in the sound absent condition, which would be reflected in a significant difference in the accuracy measures between the two conditions. Further, it was expected that in the sound present condition, participants show an enhanced sensitivity as compared to the sound absent condition, which would result in a significant difference in sensitivity measures between the sound present and the sound absent condition.

Methods Experiment 2

Participants. Thirteen participants, twelve females and one male, took part in Experiment 2. The mean age was 22.7 with a range from 17 – 24 years. All subjects were students at the University of Twente and enrolled via sona-systems, the sampling pool for participants of this university. They received one credit point each for their participation. No subject who enrolled in Experiment 1 participated in Experiment 2. Selection criteria for the participants were normal hearing functions, normal or corrected-to-normal vision and no indication of color blindness. All participants gave their informed consent and were naïve as to the purpose of the experiment. Due to poor data, one participant had to be excluded from the analysis.

Procedure/Tasks. The experiment was conducted in a university laboratory in a closed room with regular artificial lighting. The participants were seated in front of a table on which the apparatus was placed.

The experiment consisted of three parts; one practice phase, one dynamic test phase and one static test phase. On a computer screen, in random directions moving white dots and one static red fixation dot were presented on a black background. In 50% of the trials, one of the dots changed direction during the trial and in the other 50% they did not. Furthermore, in 50% of all trials, thus in 25% of the trials with and in 25% of the trials without direction change, short auditory pips were presented.

In the sound absent condition, participants were instructed to detect if there was one dot changing direction and to respond as accurately as possible if there was one changing dot among the other dots. In the sound present condition, the task was the same, but the change was accompanied by a short auditory pip. This sound provided no information about the location or the direction of the dot, only about the moment of change of the dot in the change present condition. In the change absent condition, the sound was without any information.

Does Sound Affect Visual Motion Change Detection?

The participants had to press certain keys for either answer to indicate if there was a direction change or if there no direction change. While they had to press ‘J’ (for “ja”, the Dutch equivalent for “yes”) for the presence of a direction change during the trial, they had to press ‘N’ (for “nee”, the Dutch equivalent for “no”) when there was no direction change during the trial. Via a tracking algorithm, the number of dots was tailored for the capacity of each participant during the experiment. This tailoring was achieved by use of the dynamic test phase, in which the number of dots changed and adjusted. In the last experimental phase, the static test phase, the participants could be confronted with a fixed number of dots, the number being the average number of dots each participant could monitor while still responding correctly. The participants were constantly monitored via a closed-circuit camera installed behind their backs.

Participants were told to keep their eyes fixated at the fixation dot during the trial and to press the correct associated key for an answer when, as after each trial, they were asked on the screen if there was a dot that changed direction. Immediately after the depressing of the key, the following trial was presented.

It was not possible to predict if on the following trial there would be a direction change. After each trial, the participants got feedback in Dutch about their answer and the number of remaining trials. Then the participants could continue by pressing a random key.

Practice Phase. At the beginning of the practice phase, the participants received instruction in Dutch by the researcher and via the computer screen. The practice phase consisted of two blocks, one with sound and one without sound. Each block contained twenty trials. Each trial contained twelve dots. The order of blocks was distributed randomly.

After each block, the participants had a short pause and could self determine when to continue with the experiment.

Dynamic Test Phase. The task to be conducted was the same as in the practice phase. The dynamic test phase consisted of eight blocks, of which four were sound-present and four were sound-absent. Each block contained twenty trials. The order of blocks was distributed randomly. The first trial of each condition contained twelve dots. The number of present dots per block was transferred to the following block that belonged to the same condition, i.e. sound present or sound absent. Due to the tracking algorithm, the number of dots was tailored to the capacity of each participant during the dynamic test phase. Every wrong answer

Does Sound Affect Visual Motion Change Detection?

resulted in the distraction of one dot, whereas after a row of five correct answers, three dots were added.

After each block, the participants had a short pause and could self determine when to continue with the experiment. After the dynamic test phase, a short analysis using E-Prime DataAid© was conducted to retrieve the average number of objects that was used in this phase.

Static Test Phase. At the beginning of the static test phase, the participants received instruction in Dutch via the computer screen. The task was the same as in the practice phase and the dynamic test phase, with the only exception that the number of dots displayed on the screen was tailored to each participant's performance and was static during the whole phase, so that during the static test phase, the participants worked at the computed maximum of their capacities.

Apparatus and stimuli. Stimulus presentation, timing, and data collection were achieved by use of the E-prime© 2.0 experimental software package on a standard Pentium© IV class PC. Stimuli were presented on a 17 Inch Philips 107T5 display running at 800 by 600 pixel resolution in 32 bit color and a refreshing at 60 Hz. The viewing distance was approximately 60 cm, but not strictly controlled. The experiment was run in a special secured computer mode to ensure a precise response presentation and stimulus registration.

Each trial consisted of at least 121 frames, each in which lasted 16,7 ms, before the direction change could take place. The number of frames in which the change could take place was variable. There was a likelihood of 16% in each frame that a direction change could take place. After each change, there were another 120 frames before the trial was finished. In trials without change, the procedure identical, but no direction change did take place.

An adjustment of Experiment 2 concerned the freedom of movement of the dots. Dots could not move out of the window any longer. Instead, when reaching the edge of the screen, they bounced back with the angle they reached the edge and moved back over the screen. The spontaneous direction change could only take place within a window of 1/8 of the total display size distant from the edges, so that the possibility that the change took place within the bouncing back of the dot could be ruled out.

Each dot had a radius of three pixels. The white dots were completely white, whereas the red fixation dot was completely red. They moved with a velocity of one pixel and 0.036°

Does Sound Affect Visual Motion Change Detection?

per frame. The direction of the dots was limited in that they could not move horizontally or vertically but only diagonally, more specifically in a range from 30° – 60° from all axes, respectively. The direction change again had an angle of 90° .

The sound used in this experiment was of 1000 Hz and lasted 15 ms, including a fade-in of 5 ms and a fade-out of 5 ms. It was presented via loudspeaker in front of the participant.

Data Analysis. After the first experimental phase, a short analysis using E-Prime DataAid© was conducted to retrieve the average number of objects that resulted from the tracking algorithm in this phase for each participant.

For the subsequent analysis, the data of this experiment were merged with E-Prime E-Merge and exported with E-DataAid©. The data were then analyzed by use of SPSS 16.0 for Windows. Univariate repeated measures analyses were run on the data. The factor “sound presence” with the levels “sound present” and “sound absent” served as dependent variable, whereas “accuracy” and “number of objects” served as independent variables.

Additionally, signal detection analyses using the sensitivity index (d') and the likelihood ratio (β) as a bias measure were performed. A *hit* was defined as a correctly reported answer that a motion change did take place. A *false alarm* was defined as a reported motion change when there was none in the trial, a *miss* was defined as an actual motion change that was not reported and a *correct rejection* was defined as a reported “no change” when there in fact was no motion change. Paired-Samples T-tests were used for comparison of the d' - and β -values across the two conditions.

Results Experiment 2

As in the first experiment, the One-Sample Kolmogorov-Smirnov Test proved that all data were distributed normally. Thus, application of parametric multivariate measures for the analyses of the data was indicated appropriate. Univariate repeated measures analyses (ANOVAs) with the factor “sound presence” and the levels “sound present” and “sound absent” were run on the data. For the practice phase and the static test phase, only the accuracy rates were analyzed, as the number of objects was constant in those phases. For the dynamic test phase, both the accuracy and the number of objects served as dependent variables.

Does Sound Affect Visual Motion Change Detection?

Practice Phase. For the practice phase, the analysis showed a mean accuracy of 53.2% ($SD = 0.20$) in the sound present condition and a mean accuracy of 59.6% ($SD = 0.10$) in the sound absent condition, but this difference did not reach significance ($F(1, 12) = 1.008, p = 0.335$).

Dynamic Test Phase. The analysis showed a mean accuracy of 79.6% ($SD = 0.02$) in the sound present condition and a mean accuracy of 78.4% ($SD = 0.02$) in the sound absent condition, but this difference was not significant ($F(1, 12) = 2.356, p = 0.151$).

The analysis yielded a mean number of 8.08 objects ($SD = 2.65$) in the sound present condition and a mean number of 6.29 objects ($SD = 1.94$) in the sound absent condition. This difference did reach significance ($F(1, 12) = 19.958, p = 0.001$). Pairwise comparisons revealed that participants could monitor 0.92 to 2.67 more objects when the sound was present than when the sound was absent.

Static Test Phase. The analysis revealed a mean accuracy of 80.9% ($SD = 0.08$) in the sound present condition and a mean accuracy of 73.8% ($SD = 0.08$) in the sound absent condition, with this difference reaching significance ($F(1, 12) = 8.539, p = 0.013$). Pairwise comparisons showed that in the trials with sound, participants were 1.8% to 12.4% more accurate than in the trials without sound.

Sensitivity Measures. Further, measures of sensitivity were calculated for each condition. If the simultaneously presented sound facilitated visual perceptual processes, an increased sensitivity, indicated by a larger d' , is expected for the sound present condition. Conversely, if the sound affected postperceptual decision processes, then participants might be tempted by the mere presence of the sound to indicate that a motion change did take place. This response bias would be reflected in a higher value of the decision criterion parameter β in the sound present condition.

Because the participants had to indicate whether they had seen a motion change or whether they did not the “Yes/No” paradigm was applied for this analysis.

Dynamic test phase. In the dynamic test phase, the sound present condition yielded a mean d' of 1.76 ($SD = 0.43$), whereas the sound absent condition resulted in a mean d' of 1.72 ($SD = 0.28$). The difference was not significant ($T(12) = 0.249, p = 0.808$). The mean values of β

Does Sound Affect Visual Motion Change Detection?

were 2.17 ($SD = 1.90$) and 1.87 ($SD = 0.99$) respectively, but the difference did not reach significance ($T(12) = 0.485$, $p = 0.637$).

Static test phase. In the static test phase, mean d' in the sound present condition was 1.97 ($SD = 0.57$), whereas it had a mean value of 1.46 ($SD = 0.49$) in the sound absent condition. This difference was significant, with $T(12) = 3.351$, $p = 0.006$. Computing β , the analysis revealed a mean value of 2.69 ($SD = 3.53$) in the sound present condition and a mean value of 1.62 ($SD = 9.57$) in the sound absent condition, without resulting in a significant difference ($T(12) = 1.079$, $p = 0.302$).

Discussion Experiment 2

The practice phase in Experiment 2 showed the same trend as the practice phase in Experiment 1, i.e. that at first, although not significant, performance in the sound absent condition is more accurate. This duplicated finding supports the idea that initially, the sound has a startling influence on performance. However, in Experiment 2 the trend was reversed faster than in Experiment 1, as in the dynamic test phase, participants performed better in the sound present condition, as well as in the accuracy rates (not significant) and in the number of objects (significant), with 0.9 to 2.7 more objects being presented in the sound present condition than in the sound absent condition.

It seems that the use of only one color improves performance, as the expected trends in the results emerged considerably earlier than in Experiment 1.

In Experiment 2, the stated hypothesis was confirmed. When, as done in the static test phase, each participant was presented with the individual number of objects one could track and still respond correctly, participants were significantly more accurate in the sound present than in the sound absent condition. Their performance in the trials with sound was 1.8% to 12.4 percent more accurate than in trials without sound.

Sensitivity measures have the advantage over ordinary accuracy measures that are conducted with SPSS in that they control for the response bias each participant can exert. The analysis in terms of d' and β is thus to favor because it is more precise and allows for more valid conclusions. Regarding sensitivity, the dynamic test phase showed an enhanced sensitivity and a response bias in the sound present condition as opposed to the sound absent

Does Sound Affect Visual Motion Change Detection?

condition. In the static test phase, these trends were even more pronounced. As the difference in response bias, β , was not significant, it is ruled out that participants had a negative bias, which would have been reflected in the tendency to indicate a motion change in the sound present condition more often due to the mere presence of the sound.

It is thus left to conclude that the indicated advantage of the sound is due to successful audiovisual integration and not due to the extrinsic difference of the two conditions.

General Discussion

The present study supports the idea that the findings reported by Van der Burg et al. (2008b) can be applied to moving stimuli. By means of two experiments, it was shown that a static, spatial disparate and uninformative pip at the moment of target change improves the performance as opposed in a condition without such a sound.

For finer analysis, there are several implications. As Van der Burg et al. (2008b) stated, the pip and pop effect benefits from control over eye movements, achieved by use of an electro-oculogram (EOG). They found that without such controls, the effect may be underestimated. It may thus be that with eye-movement controls the effect regarding the motion detection gets even more pronounced.

Further, although the possibilities of the sound serving as an alerting signal or cue was ruled out in the experiment of Van der Burg et al. (2008b), it is not proven that this might have influenced the results of the present study. The need of further follow-up experiments is acknowledged. This is in line with a statement of Freeman and Driver (2008) who concluded that changing just the timing of static auditory sounds can substantially affect processing of visual motion. Thus, follow-up studies with different stimulus onset asynchronies (SOAs) are required to control for possible side effects of the mere timing.

Another factor to be considered regards the distribution of attention on the different modalities and the point of one's focus. In the present study participants were instructed to only concentrate on the visual stimuli and to neglect the auditory stimulus. Although only one modality was attended, the expected effect was sound. This finding seems in contradiction with findings of Talsma et al. (2007) who stated that integration effects depend on both the visual and the auditory objects being fully attended. However, there is still ongoing research conducted to explore the influence of attention and to investigate when attention to both

Does Sound Affect Visual Motion Change Detection?

modalities is required for audiovisual integration and when it is not. It seems that in case of an incidental and thus salient stimulus, attention to one modality is sufficient, as can be interpreted from the findings of Vroomen and de Gelder (2000), i.e. that the afore mentioned freezing phenomenon was observed when the used sound was abrupt and could easily be separated from a sequence of sounds, but not when the same sound was less abrupt or belonged to a melody. This would be in line with the present study and the present results. Still, the effect of the focus of attention has to be considered in a follow-up experiment to control for effects of attentional and distributional processes.

Other than that, when regarding attention, Beer and Röder (2004a) found that attention to motion enhances the processing of both visual and auditory stimuli, although they did not observe any reliable crossmodal effects of attention to motion. The difference to other studies that found these crossmodal effects (Beer & Röder, 2004b) is that in this experiment by Beer and Röder (2004a) no response to stimuli of the irrelevant modality was required. Due to the similarity of instruction in this experiment and the present study, the need for experiments controlling for the factor of unimodal focused attention grows.

Another consideration is the source of the sound and the influence on audiovisual integration. Whereas in the experiment by Van der Burg et al. (2008b) the sounds were presented via headphones, thus lateral; in the present study the sounds were presented via the loudspeakers in front of the participants, thus medial. In both experiments the audiovisual facilitation was present. This is in contradiction with the results of Frassinetti et al. (2002), who reported that enhanced sensitivity by sounds only were found when the sounds were presented from a lateral location, thus as in Van der Burg's et al. (2008b) experiment, rather than a nasal location, as in the present study. An explanation for these conflicting results has yet to be found.

A finding from another area of research also calls for attention, namely the inversed set size effect. Typically, performance decreases as the number of distractors increases, as indicated in deteriorated accuracy rates or increased reaction time (Palmer, 1995). However, when the observers are without knowledge about the target except for the fact that it is different from the distractors, a task known as oddity search (Schoonveld, Shimozaki & Eckstein, 2007), in some instances the set-size effects are weakened or even reversed, so that performance improves with increasing set size, as Bacon and Egeth (1991) showed. An explanation for this inverted effect is distractor grouping i.e. that, as similar stimuli get physically closer to each other, they tend to group together, resulting in a reduced effective set size. By varying grouping, they further showed that grouping accounts for the inversed set

Does Sound Affect Visual Motion Change Detection?

size effect. When taking this effect into consideration, it seems a plausible matter for research regarding the topic of the present study. For a pop-out to take place, it has to be salient. When the effective set size is reduced, the salience will be more pronounced. However, the stimuli used in this study differed not in physiognomy but only during one moment in their ability to change direction. Further research is needed, considering the possibility that with training, the scenario as used in the practice phase of Experiment 2, thus the question whether there is one target among too many distractor stimuli to track them all, performance will increase, and then to explore the effect of a sound in these circumstances.

Finally, the sample size is open to discussion. With a considerably small group of participants, significant effects were found. At first appearance this seems convincing, however, this also bears possible limitations regarding the generalizability of the results. It has to be investigated if with other social groups or a higher sample size the results of the present study can be replicated.

Conclusion

The present study provides evidence that the detection of a direction change of one moving stimuli across moving distractor stimuli can be facilitated by use of a static and uninformative auditory pip.

Although further research needs to be conducted to control for the limitations of this study, the present findings contribute to the research of audiovisual integration and the influence of attention.

References

- Bacon, W.F. & Egeth, H.E. (1991). Local Processes in Preattentive Feature Detection. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1), 77-90.
- Beer, A.L. & Röder, B. (2004a). Attention to Motion Enhances Processing of Both Visual and Auditory Stimuli: An Event-Related Potential Study. *Cognitive Brain Research*, 18, 205-225.
- Beer, A.L. & Röder, B. (2004b). Unimodal and Crossmodal Effects of Endogenous Attention to Visual and Auditory Motion. *Cognitive, Affective, & Behavioral Neuroscience*, 4(2), 230-240.
- Bremmer, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., & Fink, G.R. (2001). Polymodal Motion Processing in Posterior Parietal and Premotor Cortex: A Human fMRI Study Strongly Implies Equivalencies Between Humans and Monkeys. *Neuron*, 29(1), 287-296.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., & Movshon, J.A. (1992). The Analysis of Visual Motion: A Comparison of Neuronal and Psychophysical Performance. *Journal of Neuroscience*, 12(12), 4745-4765.
- Burr, A. & Alais, D. (2006). Combining Visual and Auditory Information. *Progress in Brain Research*, 155, 243-258.
- Calvert, G.A. (2001). Crossmodal Processing in the Human Brain: Insights from Functional Neuroimaging Studies. *Cerebral Cortex*, 11(12), 1110-1123.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, 22(13), 5749-5759.
- Frassinetti, E., Bolognini, N., & Làdavas, E. (2002). Enhancement of Visual Perception by Crossmodal Visuo-Auditory Interaction. *Experimental Brain Research*, 147, 332-343.
- Freeman, E. & Driver, J. (2008). Direction of Visual Apparent Motion Driven Solely by Timing of a Static Sound. *Current Biology*, 18, 1262-1266.
- Fujisaki, W., & Nishida, S. (2005). Temporal Frequency Characteristics of Synchrony-Asynchrony Discrimination of Audio-Visual Analysis. *Experimental Brain Research*, 116, 455-464.

Does Sound Affect Visual Motion Change Detection?

- Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual Search for a Target Changing in Synchrony with an Auditory Signal. *Proceedings of the Royal Society B: Biological Sciences*, 273, 865-874.
- Giard, M.H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysical Study. *Journal of Cognitive Neuroscience*, 11(5), 473-490.
- Girelli, M. & Luck, S.J. (1997). Are the Same Attentional Mechanisms Used to Detect Visual Search Targets Defined by Color, Orientation, and Motion?. *Journal of Cognitive Neuroscience*, 9(2), 238-253.
- Goodale, M.A., & Milner, D.A. (1992). Separate Visual Pathways for Perception and Action. *Trends in Neuroscience*, 15(1), 20-25.
- Graziano, M.S.A. (2001). A System of Multimodal Areas in the Primate Brain. *Neuron*, 29(1), 4-6.
- Hooge, I.T.C. & Erkelens, C.J. (1999). Peripheral Vision and Oculomotor Control during Visual Search. *Vision Research*, 39, 1567-1575.
- Leber, A.B. & Egeth, H.E. (2006). It's under control: Top-Down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, 13(1), 132-138.
- Lewald, J., Ehrenstein, W.H., & Guski, R. (2001). Spatio-Temporal Constraints for Auditory-Visual Integration. *Behavioral Brain Research*, 121, 69-79.
- McDonald, J., Teder-Sälejärvi, W.A., & Hillyard, S.A. (2000). Involuntary Orienting to Sound Improves Visual Perception. *Nature*, 407, 906-908.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., & Foxe, J.J. (2002). Multisensory Auditory-Visual Interactions During Early Sensory Processing In Humans: A High-Density Electrical mapping Study. *Brain Research: Cognitive Brain Research*, 14(1), 115-128.
- Mozolic, J.L., Hugenschmidt, C.E., Pfeiffer, A.M., & Laurienti, P.J. (2008). Modality-Specific Selective Attention Attenuates Multisensory Integration. *Experimental Brain Research*, 184, 39-52.
- Neville, H.J. & Lawson, D. (1987). Attention to Central and Peripheral Visual Space in a Movement Detection Task: An Event-Related Potential and Behavioral Study. I. Normal Hearing Adults. *Brain Research*, 405, 253-267.
- Palmer, J. (1995). Attention in Visual Search: Distinguishing Four Causes of Set-Size Effects. *Current Directions in Psychological Science*, 4, 118-123.

Does Sound Affect Visual Motion Change Detection?

- Perrott, D.R., Saberi, K., Brown, K., & Strybel, T.Z. (1990). Auditory Psychomotor Coordination and Visual Search Performance. *Perception & Psychophysics*, *48*(3), 214-226.
- Perrott, D.R., Sadralodabai, T., Saberi, K., & Strybel, T.Z. (1991). Aurally Aided Visual Search in the Central Visual Field: Effects of Visual Load and Visual Enhancement of the Target. *Human Factors*, *33*(4), 389-500.
- Schoonveld, W., Shimozaki, S.S. & Eckstein, M.P. (2007). Optimal Observer Model of Single-Fixation Oddity Search Task Predicts a Shallow Set-Size Function. *Journal of Vision*, *7*(10), 1-16.
- Schroeder, C.E., & Foxe, J.J. (2005). Multisensory Contributions To Low-Level, 'Unisensory' Processing. *Current Opinions in Neurobiology*, *15*(4), 454-458.
- Sekuler, R., & Ball, K. (1977). Mental Set Alters Visibility of Moving Targets. *Science*, *198*(4312), 60-62.
- Spence, C., & Driver, J. (1997). Audiovisual Links in Exogenous Covert Spatial Orienting. *Perception & Psychophysics*, *59*(1), 1-22.
- Standage, G.P., & Benevento, L.A. (1983). The Organization of Connections between the Pulvinar and Visual Area MT in the Macaque Monkey. *Brain Research*, *262*(2), 288-294.
- Stein, B.E., London, N., Wilkinson, L.K., & Price, D.D. (1996). Enhancement of Perceived Visual Intensity by Auditory Stimuli: A Psychophysical Analysis. *Journal of Cognitive Neuroscience*, *8*(6), 497-506.
- Talsma, D., Doty, T.J., & Woldorff, M.G. (2007). Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration?. *Cerebral Cortex*, *17*, 679-690.
- Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., & Theeuwes, J. (2008a). Audiovisual Events Capture Attention: Evidence from Temporal Order Judgments. *Journal of Vision*, *8*(5), 1-10.
- Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., & Theeuwes, J. (2008b). Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053-1065.
- Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., Talsma, D. & Theeuwes, J. (2008). Multisensory Synchrony Guides Attention in Dynamic Cluttered Environments. *OPAM 2008 Report*, *16*(8), 1140-1143.

Does Sound Affect Visual Motion Change Detection?

Vroomen, J., & de Gelder, B. (2000). Sound Enhances Visual Perception: Cross-Modal Effects of Auditory Organization on Vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1583-1590.