Spring 2015

The effect of auditory stimuli on visual time-to-contact perception

Chelsea L. Rugel
James Madison University

The Effect of Auditory Stimuli on Visual Time-to-Contact Perception

An Honors Program Project Presented to

the Faculty of the Undergraduate

College of Health and Behavioral Sciences

James Madison University

by Chelsea Rugel

May 2015

Accepted by the faculty of the Department of Psychology, James Madison University, in partial fulfillment of the requirements for the Honors Program.

FACULTY COMMITTEE:

Project Advisor: Jeffrey Andre, PhD.,
Professor, Psychology

Reader: Michael Hall, PhD.,
Professor Psychology

Reader: Jeffrey Dyche, PhD.,
Professor Psychlogy

HONORS PROGRAM APPROVAL:

Philip Frana, Ph.D.,
Interim Director, Honors Program

PUBLIC PRESENTATION

This work is accepted for presentation, in part or in full, at the Honors Symposium on 4/24/15

Table of Contents

Abstract

Previous research has demonstrated that auditory and visual stimuli have individual effects on the accuracy of a person's estimation of time-to-contact (TTC), the time at which two objects collide. Prior findings also suggest that there is cross-modal interference between vision and audition; however, this phenomenon has never been studied in a TTC situation. (Driver & Spence, 1998; Ichikawa & Masskura, 2006; Roseboom, Kawabe, & Nishida, 2013) In this study we attempted to fill in this research gap by examining the effect of auditory speed cues over visual speed cues in a two-dimensional TTC scenario, and by determining if an object's temporal presence influences accurate perception of TTC by using occlusion. Our results indicate that in the presence of auditory and visual speed disparity, participants rely more heavily on auditory cues, but when auditory and visual speeds are equivalent, or when there is no audition present, participants rely more on visual cues.

Keywords: Time-to-contact, TTC, audition, vision, occlusion, cross-modal interference

The purposes of this study were to examine the effect of auditory cues over visual cues in a two-dimensional time-to-contact (TTC) scenario and to determine if an object's temporal presence influences accurate perception of TTC.

Time-to-contact, or time-to-collision, is used to explain the spatial localization of two objects by relating time and distance. It is defined as the remaining time before contact between two or more objects. TTC is typically used to define how quickly an object travels in a three-dimensional plane toward an observer. In studies that examine this aspect of TTC, tau (τ), the inverse of the rate of expansion of an oncoming image on the retina, is the traditional unit of measure (Lee, 1976).

However, in studies such as this one that involve two objects colliding on a two-dimensional plane rather than a three-dimensional plane, τ would be an invalid measure. In experiments that focus on two objects coming together on a two-dimensional plane, TTC is best determined by employing the equation TTC = d/v (distance to collision/approach speed) (Bootsma & Oudejans, 1993). Little research has been conducted with TTC on a two-dimensional plane, and the research that has been done in this type of scenario has typically only examined the effect of one sensory modality. Since our study was concerned with how audition influences TTC perception, we investigated various visual and auditory effects on TTC as well as cross modal interferences.

A variety of visual effects have been studied in TTC settings. For instance, Calabro, Beardsley and Vaina (2011) examined the effect of the presence of a distractor object (an irrelevant horizontal movement) on TTC approximations. The irrelevant motion caused a significant decline in participants' performance with a 1.4% per cm/s decrease in TTC estimation. Alexander, Barham, and Black (2002) analyzed visual effects in an applied TTC research setting using a computer driving simulation. They found that if a leading car was red (as opposed to black, green, or yellow), a participant was better able to estimate TTC and prevent collision. Landwehr, Brendel, and Hecht, (2013) examined the visual aspects of luminance and object/background contrast in a three-dimensional TTC scenario. Despite previous findings that other animals are impaired by changes in contrast and environmental luminance, Landwehr, Brendel, and Hecht (2013) discovered no such decline in human TTC performance. Perhaps the most common visual aspect of TTC that has been studied however, is the perceived speed of colliding objects (Alexander, Barham, & Black, 2002; Bennett, Baures, Hecht, & Benguigui, 2010; Bootsma, & Oudejans, 1993; Calabro, Beardsley, and Vaina, 2011; Pundlik, Peli, & Luo, n.d.). The general finding observed from these experiments is that TTC estimation is less consistent in situations involving faster TTC (Bennett, Baures, Hecht & Benguigui, 2010).

When numerous visual manipulations of object motion are present, estimations of TTC appear to be more greatly impaired than when only one visual manipulation is present. Examples of visual manipulations studied in various combinations include: relative size of object, depth cues, motion parallax, occlusion, and height of object in the field (DeLucia, Kaiser, Bush, Meyer, & Sweet, 2003), as well as viewing time and monocular/binocular depth cues (López-Moliner, Supèr, & Keil, 2013). DeLucia et al., (2003) attributed the greater error of TTC estimation in these scenarios to the integration of visual information. From their experiments they hypothesized that when only a single visual cue is available, people utilize a selection process, but when multiple visual cues are available, people use an integration process that requires greater neural activity. This greater neural activation leads to a loss of accuracy in each of the visual tasks that are integrated, thus resulting in a loss of accuracy in the overall TTC task.

Since we were interested in determining if the amount of time that an object is present during a trail impacts accurate perception of TTC, we decided to use occlusions in our experiment as well as manipulations of visual and auditory speed. The interaction between various object speeds and occlusion were studied extensively by DeLucia et al., (2003). The researchers examined TTC on a three-dimensional plane with an object moving toward an observer. They found that when occlusion was present, the observer tended to estimate TTC significantly earlier than when contact would have actually occurred. In addition, the average difference in TTC estimates between observers was greater in the presence of occlusion.

However, we did not want to confine our experiment only to visual effects because visual and auditory events often occur together, especially in real-world applications of TTC such as driving. Therefore, we decided to examine the effects of audition on TTC as well. However, there has been much less research conducted on how different auditory stimuli affect TTC, and most of the research that has implemented audition variables has been examined in a three-dimensional plane. For instance, in a driving study, it was found that if a collision warning signal (e.g. a constant intensity sound or car horn) was able to redirect the driver’s attention to the leading vehicle, the signal greatly decreased the odds that a rear-end collision occurred (Gray, 2010). In a second part of the same study, Gray’s results also suggested that TTC can be signaled by a looming warning (a warning of growing intensity) which could increase or decrease a driver’s reaction to a potential collision. In another driving related study, experimental results indicated that having more than one auditory warning origin location, or placing sounds symmetrically behind and in front of a driver, increases the time needed for a driver to detect a visual target located in the same place as the sound (Wallace & Fisher, 1998).

Although understanding the effects of auditory and visual stimuli independently was useful, since we were interested in the influence of auditory stimuli on visual TTC, we also examined previous findings regarding the impact of cross-modal interference. The overall finding of auditory and visual cross-modal research was that visual stimuli typically dominate over auditory stimuli in bimodal spatial perceptions, and auditory stimuli usually dominate over visual stimuli in bimodal temporal perception (Driver & Spence, 1998; Ortega, Guzman-Martinez, Grabowecky & Suzuki, 2014). In a review article compiling research findings from previous cross-modal attention studies, Driver and Spence, (1998) explained how sound can dominate over sight. From the results of their study on cross-modal interference in visual direction discrimination, Driver and Spence (1998) claimed that a sudden auditory event at a specific location does not cause visual attention by stimulating a specific retinal location but rather, the sound stimulates a representation of the area of the retina that corresponds to the external location of the sound at the time of the auditory event. Since then, other claims regarding how auditory stimuli affect visual processing have been proposed.

Ichikawa and Masskura (2006) hypothesized that auditory stimuli alters the visual processing of motion by controlling spatiotemporal resolution and the amount of displacement. They tested this hypothesis by observing visual and auditory stimulus onset asynchrony and the resulting displacement of the visual stimulus. Their findings revealed an interaction between auditory stimulus frequency, visual stimulus frequency, and displacement. Ichikawa and Masskura (2006) claimed the reason for this displacement was likely due to the fact that, although vision is used more effectively than audition when determining the spatial location of stimuli, because there are also temporal factors involved in motion processing, audition can influence vision. If the visual system used for a specific task is associated with a temporal aspect of an object (e.g. motion processing), auditory stimuli could impact not only the processing of the object's temporal features, but the processing of spatio-temporal aspects as a whole (Ichikawa & Masakura, 2006). Research has also shown that even spatially uninformative auditory signals can modulate the direction of the apparent visual motion of an object (Roseboom, Kawabe, & Nishida, 2013). Roseboom, Kawabe, and Nishida, (2013), suggested from these results that the organization of auditory stimuli facilitates the segmentation of visual event streams.

Findings such as those by Roseboom, Kawabe, and Nishida, (2013), Driver and Spence (1998), and Ichikawa and Masskura (2006) that claimed auditory information can alter perception of visual information served as the preliminary basis for the outline of our study. In order to assess the influence of audition over vision in a TTC scenario, we decided to examine the effects of TTC on a two-dimensional plane since most of the prior TTC research was conducted on a three-dimensional plane. Because there has been no previous research focusing on TTC and sensory interference, it is unknown if audition has a greater effect than vision in a TTC situation. In order to test this idea we examined the interaction between TTC perception, auditory speed, and object speed. We hypothesized that disparities between the speed of auditory and visual stimuli would effect a participant's estimation of TTC.

The secondary purpose of our experiment was to look at occlusion time and TTC estimates to determine if the amount of time that an object is visible on the screen effects TTC perception. We hypothesized that the later the occlusion occurred, the more accurate TTC perception would be.

Method

Subjects

Our experiment was conducted with 17 undergraduate students from James Madison University (15 females and 2 males) ages 18-25 (average of 21 years). Students were recruited for the study by word of mouth. All participants had unimpaired or corrected sight and hearing.

Apparatus

In all of the trials for this experiment, a black square was located at one of three positions: at the left of the screen, in between the middle of the screen and the left of the screen, and in the middle of the screen. This square moved toward a stationary black square on the right hand side of the screen. As soon as the square on the left started moving, one of three auditory stimuli speeds (slow, medium, or fast) began. An occlusion appeared at one of two times before the moving black square collided with the stationary square (Figure 1).

Short video files for this experiment were created using Microsoft PowerPoint. Each trial was composed of an animated PowerPoint slide with an audio file imbedded into it. The PowerPoint slides were then converted into WAV files and imported into a DirectRT Empirisoft program. We chose a visual TTC of 2.2s for our trials because this value was within the range of TTC times used in previous research (1-3s: Landwehr, Brendel, & Hecht, 2013; 0.4-1.3s: Bennett, Baures, Hecht, & Benguigui, 2010; 2.5s: Bootsma & Oudejans, 1993). Although a TTC value of 2.2s was slower than the majority of TTC values in previous research, we wanted to make sure that our results were due to participants' TTC perception, not their reaction time abilities.

Visual stimuli. The visual component of each trial was created entirely in Microsoft PowerPoint. The visual element of the trials consisted of two 0.5in x 0.5in black squares on a white background. For all trials one square was located on the left-hand side of the screen, and the other square was located on the right-hand side of the screen. During each trial, the square on the left-hand side of the screen was animated so that it moved linearly toward the stationary square on the right-hand side of the screen. Since the two squares collided at the same time in all trials (2.2s), the starting position of the animated square in relation to the center of the screen (-8in, -4in, 0in) determined object speed. For trials including an auditory stimulus, the sound commenced simultaneously with the movement of the square. A white 1in x 2in rectangle appeared at one of two times (1.69s or 1.96s), obstructing the view of the animated black square before it could collide with the square on the right-hand side of the screen (Figure 1).

Auditory stimuli. The specific sound of the auditory stimuli was created using Ableton Live9 programming (https://www.ableton.com/en/live/new-in-9/) and a Camel Audio plug in, Alchemy v.1.55.0 (http://www.camelaudio.com/faqs/Alchemy/Alchemy_Release_Notes). The attack and release of the AHDSR component of the sound were altered so that there would not be a delayed onset or termination of the auditory stimuli. The attack was changed to a linear on-ramp of 20ms and the release was changed to a linear off-ramp of 20ms. The fundamental frequency of the sound was also altered in this plug in so that the auditory stimuli would not be irritating to the participants. The original auditory stimulus was a 220 Hz saw tooth frequency. A low pass 2 pole biquad LPZ-BQ filter and a 2498 Hz cutoff filter were applied to the original stimulus. Another plug in, Wavearts Panorama_5 (http://wavearts.com/products/plugins/panorama/) was used to define the spatial orientation of the auditory stimuli. We decided to use headphone mode HRTF (head related transfer function) based on MIT Kemar data to define our spatial parameters. The distance from the center of the participant's head to the space out in front of him/her where the auditory stimuli sounded was four feet. This distance was greater than the distance between the participant and the computer screen so that the change in sound localization during the trials would be easier to perceive. The horizontal distance of the auditory stimuli spanned -2.42ft to +2.42 ft with 0ft representing the center of the participant's face. Three conditions were created: -2.42 ft to +2.42 ft, -1.21 ft to +2.42 ft, and 0ft to +2.42 ft. The auditory stimuli were extended horizontally beyond the screen distance so that the auditory traveling distance between the left and right side of the headphones was slightly exaggerated. Therefore, the auditory stimulus from -2.42 ft to +2.42 ft was comparable to the visual stimulus with a start position of -8 in, -1.21 ft to +2.42 ft to -4 in, and 0ft to +2.42 ft to 0in. Final changes to the auditory stimuli were made in Adobe Audition v.3 (http://www.adobe.com/support/downloads/product.jsp?platform =Windows&product=92). The silence at the end of the file was trimmed off at 2.7s and the file was compressed to a duration of 2.66s. The auditory stimuli were intentionally created to last slightly longer than the TTC of the two black squares so that participants could not use the end of the sound to accurately estimate TTC.

Design

Independent variables for this experiment included: object speed (slow: object start position 0in, medium: object start position -4in, fast: object start position -8in), auditory speed (slow: sound movement 0ft to +2.42 ft, medium: sound movement -1.21 ft to +2.42 ft, fast: sound movement -2.42 ft to +2.42 ft, and none: no auditory stimulus), and occlusion time (later: appears at 1.96s, and sooner: appears at 1.69s). Each combination of these variables was represented within 18 different trials. Each trial within the set of 18 was randomized for each participant using DirectRT. The set of 18 trials was then run three times for each participant. Six control trials, without an auditory stimulus, of every combination of object speed and occlusion time were randomly run at the beginning and end of the experiment. The control trials were not interspersed with experimental trials so that the participants would not be disrupted by the lack of auditory stimuli. Three practice trials (control type trials in which response time was not recorded) were provided for the participant before any other trials in order to minimize practice effects. Overall, there were 3 practice trials, 12 control trials, and 54 experimental trials.

The dependent variable studied was estimated TTC. It was quantified as the time at which the participant pressed the spacebar, his/her best approximation of the moment of collision between the two squares. The time at which the participant pressed the spacebar was recorded in milliseconds using DirectRT.

Procedure

This experiment was held in a vison studies laboratory with the lights on. Once a participant entered the room, he/she was provided basic information about the study. Participants signed an informed consent form and then filled out demographics paper. These two forms were stored separately. The participants sat down in a chair and rested their heads on a chin rest 15in from a 16in x 12in computer monitor. They then placed a set of headphones over their ears. The volume on the computer was set at 35. Instructions for the experiment were provided on the computer screen. The participants were asked to read the instructions and then press the spacebar on the keyboard below to begin the practice trials. At the end of the three practice trials, the participants were instructed to press the spacebar to begin the experiment. After each trial, a screen appeared informing the participant to press the spacebar when they were ready to start the next trial. Separate instruction screens were provided for the participant before a set of trials with an auditory stimulus began and when these trials ended and control trials resumed. Once the experiment was completed, the participant was debriefed.

Results

When running analyses, object speed and auditory speed were condensed into one variable comparing auditory speed to object speed (auditory speed faster, auditory speed equal, auditory speed slower, no auditory stimulus) because without condensing the two variables, the design would not be orthogonal (Figure 2). A 4x2 within subjects analysis of variance (ANOVA) was used to determine the results of this study. The variables we examined were auditory speed compared to object speed (levels: auditory faster, auditory equal, auditory slower, and no auditory) and occlusion time (levels: sooner, later). Our results indicated that there was a main effect of auditory time compared to object speed [$F(3,14, MSE= 46.70) = 33.91$, $p<0.001$, partial $\eta^2= 0.879$]. Our data did not reveal a main effect of occlusion time [$F(1,16, MSE= 46.16) =1.621$, $p=0.221$, partial $\eta^2= 0.092$], and there was a marginally significant interaction between auditory time compared to object speed and occlusion time [$F (3,14, MSE=48.21) = 2.99$, $p= 0.067$, partial $\eta^2= 0.390$] (Figure 3).

Discussion

The results of our experiment support our hypothesis that disparities between the speed of auditory and visual stimuli effect a participant’s estimation of TTC. Overall participants perceived TTC more accurately (response time closer to 2,200ms) when the speed of the auditory stimuli and visual stimuli matched and when the auditory stimuli was faster than the visual stimuli (Figure 3). Participants were less accurate at perceiving TTC when the auditory stimuli was slower than the visual stimuli, and were least accurate when there was no auditory stimulus (Figure 3). Because participants were most inaccurate in their perceptions of TTC when no sound was present, the auditory stimuli may have facilitated more accurate perceptions of TTC. Since our experiment was a temporal task, these results correspond with previous findings that auditory stimuli usually dominate over visual stimuli in bimodal temporal perception (Driver & Spence, 1998; Ichikawa & Masskura, 2006; Roseboom, Kawabe, & Nishida, 2013).

Occlusion time was marginally significant; participants were most accurate in their estimations of TTC when the auditory and visual stimuli speeds matched and occlusion occurred later (Figure 3). This pattern in our occlusion data also supports the claim that auditory stimuli dominate over visual stimuli in temporal situations (Driver & Spence, 1998; Ichikawa & Masskura, 2006; Roseboom, Kawabe, & Nishida, 2013). Because the auditory stimuli in our trials displayed more of an impact when the auditory and visual stimuli were the same speed, or when there was no auditory stimuli, the visual cue of occlusion could have had a greater effect in those specific situations since there was less competition for sensory dominance. However, our occlusion findings did not match previous research that had specifically examined occlusion. DeLucia et al., (2003) found that when occlusion was present, participants tended to underestimate TTC. In all of our conditions, participants overestimated TTC (average error of TTC response time > 0.0 ms; Figure 3). There were a few potential problems in our study however that could account for this disparity between our findings and previous research that should be addressed in future studies.

One of the major limitations of our study was that since we were only able to create three different auditory speeds, there could be confounds within the collapsed variables of faster auditory speed and slower auditory speed (Figure 2). The auditory stimuli could also have created a ceiling effect of the participants' response times. Despite instructions before the experimental trials informing participants not to respond at the end of the auditory stimulus (2.66s), participants still could have done so. In addition, the time difference between our two occlusions (0.27s) might not have been significant enough to alter participants' TTC perception. Because of these potential problems in our study, future research should examine TTC using a greater variety of auditory stimuli, providing auditory conditions faster, equal to, and slower than each object speed, as well as increasing the time difference between occlusions.

Overall, our study suggests that audition is useful in accurately perceiving TTC. When auditory stimuli and visual stimuli compete for temporal dominance, auditory stimuli appear to be more influential. However, if auditory stimuli is not competing with visual stimuli (e.g. in the auditory equal and no auditory conditions), later occlusion time allows for more accurate TTC perception.

Acknowledgements

Special thanks to the Psychology department at James Madison University, Dr. Jeffrey T. Andre, Dr. Michael D. Hall, Dr. Jeffrey S. Dyche, and the James Madison University Honors Program for their assistance in this study.

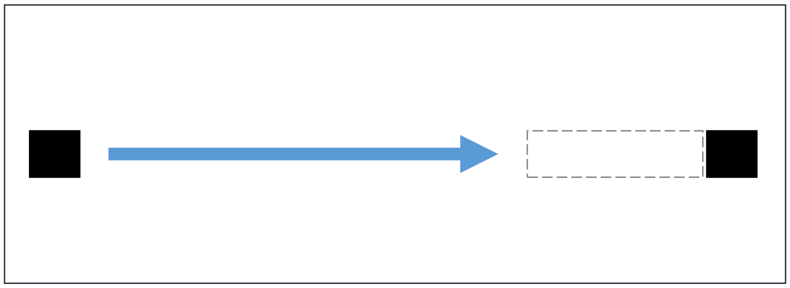

Figure 1. A general diagram of a TTC trial (fast object speed, later occlusion). The arrow represents object movement and the dashed rectangle represents occlusion.

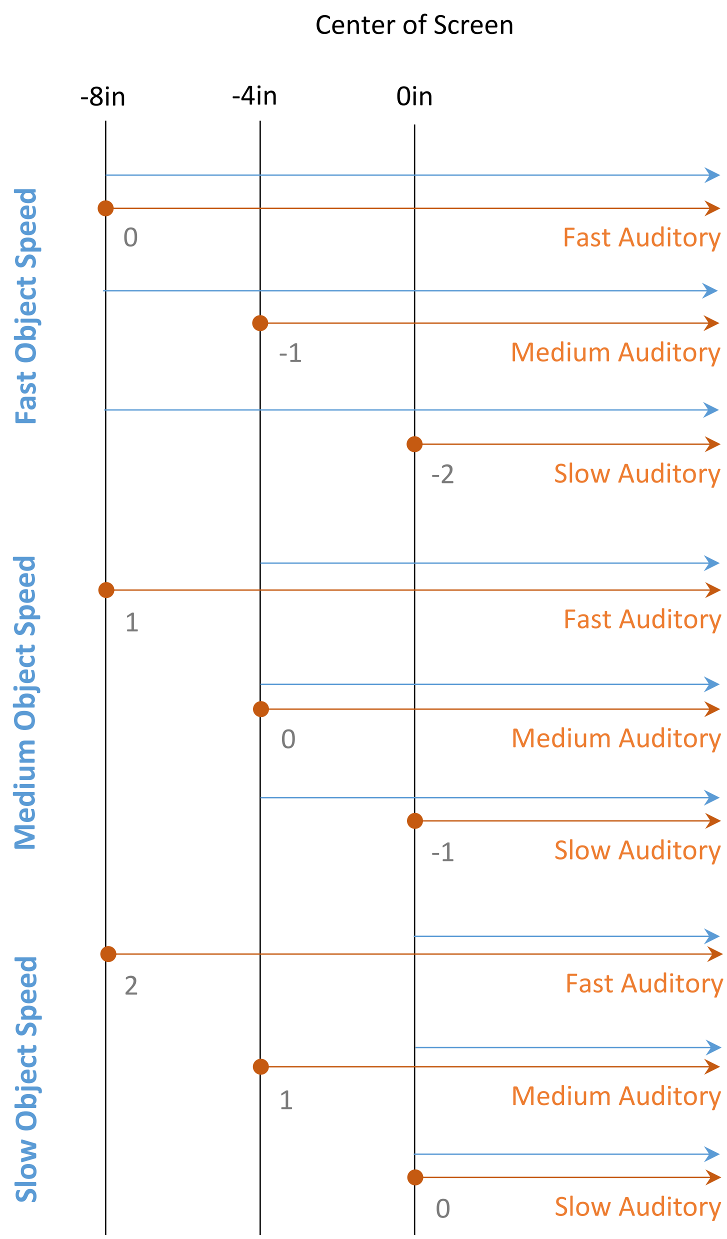

Figure 2. Difference between auditory start position (orange) and visual start position (blue) relating to speed. Numbers represent comparison between auditory and visual speeds (2= auditory twice as fast, 1= auditory one fourth faster, 0= auditory equal, 1= auditory one fourth slower, -2= auditory twice as slow). The 2s and 1s were grouped to form a general faster auditory condition, and the -2s and -1s were grouped to create a general slower auditory condition.

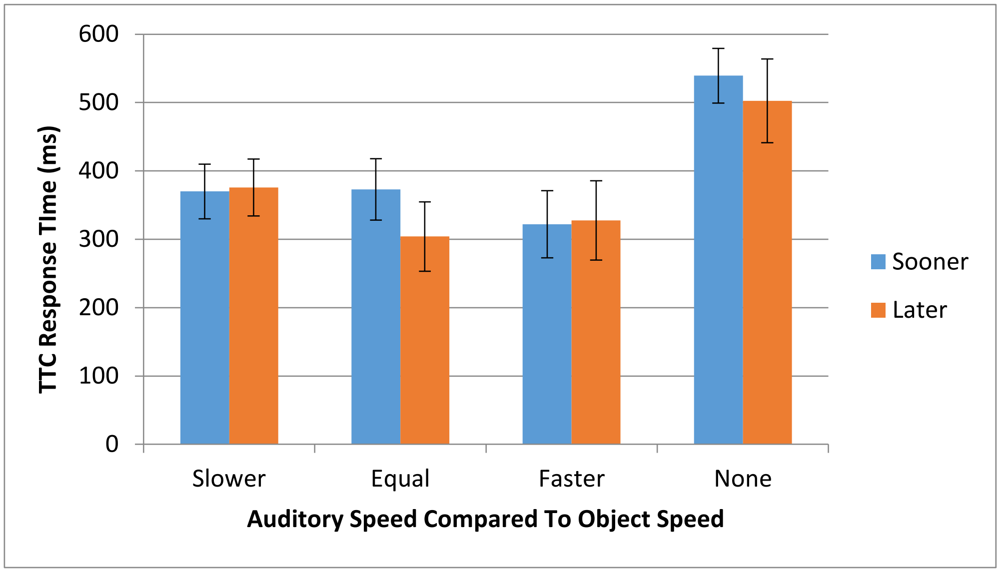

Figure 3. Average error of TTC response time (in ms) as a function of auditory speed compared to object speed and occlusion time. Accurate TTC was 2200ms. Average SEM +/- 48.21.

Works Cited

AbletonLive (Version 9.2) [Software]. Available from https://www.ableton.com/en/live/new-in9/

Adobe: Audition (Version 3.0) [Software]. Available from http://www.adobe.com/support/downloads/product.jsp?platform%20=Windows&product=92

Alexander, J., Barham, P., & Black, I. (2002). Factors influencing the probability of an incident at a junction: results from an interactive driving simulator. *Accident Analysis and Prevention, 34,* 779-792.

Bennett, S.J., Baures, R., Hecht, H., & Benguigui, N. (2010). Eye movements influence estimation of time-to-contact in prediction motion. *Experimental Brain Research, 206(4),* 399-407.

Bootsma, R.J., & Oudejans, R.R. (1993). Visual information about time-to-collision between two objects. *Journal of Experimental Psychology: Human Perception and Performance, 19(5),* 1041-1052.

Calabro, F.J, Beardsley, S.A., & Vaina, L.M. (2011). Difference motion cues are used to estimate time-to-arrival for frontoparallel and looming trajectories. *Vision Research, 51,* 2378-2385.

Camel Audio: Alchemy (Version 1.55.0) [Software]. Available from http://www.camelaudio.com/faqs/Alchemy/Alchemy_Release_Notes

DeLucia, P., Kaiser, M., Bush, J., Meyer L., & Sweet, B. (2003). Information integration in judgments of time to contact. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 56(7),* 1165-1189.

Driver, J. & Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology, 8,* 245-253.

Gray, R. (2010). Looming auditory collision warnings for driving. *Human Factors, 53(1),* 63-74.

Gray, R., Macuga, K., & Regan, D. (2004). Long range interactions between object-motion and self-motion in the perception of movement in depth. *Vision Research, 44(2),* 179-195.

Ichikawa, M., & Masakura, Y. (2006). Auditory stimulation affects apparent motion. *Japanese Psychological Research, 48(2),* 91-101.

Landwehr, K., Brendel, E., & Hecht, H. (2013). Luminance and contrast in visual perception of time to collision. *Vision Research, 89,* 18-23.

Lee, D.N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception, 5(4),* 437-459.

López-Moliner, J., Supèr, H., & Keil, M.S. (2013). The time course of estimating time-to-contact: Switching between sources of information. *Vision Research, 92,* 53-58.

Ortega, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2014). Audition dominates vision in duration perception irrespective of salience, attention, and temporal discriminability. *Attention, Perception, & Psychophysics, 76(5),* 1485-1502.

Pundlik, S., Peli, E., & Luo, G. (n.d.). Time to collision and collision risk estimation from local scale and motion. Schepens Eye Research Institute. Harvard Medical School.

Roseboom, W., Kawabe, T., Nishida, S. (2013). Direction of visual apparent motion driven by perceptual organization of cross-modal signals. *Journal of Vision, 13(1),* 1-13.

Wallace, J.S., & Fisher, D.L. (1998). Sound localization: Information theory analysis. *Human Factors, 40(1)*, 50-68.

Wavearts: Panorama (Version 5) [Software]. Available from http://wavearts.com/products /plugins/panorama/