Audiovisual tau effect

Kawabe, Takahiro
Faculty of Human-Environment Studies, Kyushu University

Miura, Kayo
Faculty of Human-Environment Studies, Kyushu University

Yamada, Yuki
Faculty of Human-Environment Studies, Kyushu University

https://hdl.handle.net/2324/12407

出版情報：Acta Psychologica. 128 (2), pp.249-254, 2008-06. Elsevier
バージョン：

Audiovisual tau effect

Takahiro Kawabe*, Kayo Miura and Yuki Yamada
Kyushu University

*Corresponding author
Faculty of Human-Environment Studies
Kyushu University
6-19-1 Hakozaki Higashi-ward
Fukuoka 8128581
Japan
E-mail: tkawabe@psycho.hes.kyushu-u.ac.jp
TEL & FAX: +81-92-642-2418

Keywords: audiovisual integration, tau effect, constant velocity assumption
PsychoINFO classification codes: 2320 (Sensory perception)

Abstract

This study investigated how spatial intervals between successive visual flashes are influenced by the temporal intervals between auditory pure tones presented concurrently with the flashes. Three successive visual flashes defined two spatial intervals with different extents as well as two equal temporal intervals. The onsets of the first and third tones were temporally aligned with those of the first and third flashes, while the onset of the second tone was temporally offset to that of the second visual flash, resulting in shorter or longer temporal intervals between pairs of tones. Observers judged which of the first or second spatial intervals between flashes was shorter. The results showed that the shorter temporal interval between tones caused underestimation of the spatial interval between flashes. On the other hand, stimuli without the first and third tones did not result in underestimation of spatial intervals between flashes. These results indicate an audiovisual tau effect which is triggered by a constant velocity assumption applied to moving objects defined by more than one modality.

Introduction

It is a well-known phenomenon that the perceived extent of spatial intervals between successive visual flashes increases with the duration of the temporal intervals between flashes. This phenomenon, the so-called 'tau effect', has been reported in studies investigating the perceived lengths of spatial intervals between successive visual flashes (Benussi, 1913), auditory pulses (Sarrazin, Giraudo, & Pittenger, 2007) and tactile stimulations (Helson, 1930; Helson & King, 1931).

The tau effect has been explained on the basis of the hypothetical idea that the visual system 'imputes uniform motion to discontinuous dynamic display' (Jones & Huang, 1982). Specifically, given that the three successive visual flashes defining two spatial intervals (S1 and S2) and two temporal intervals (T1 and T2), the visual system intuitively imputes motion at a given speed to the flashes, and tries to equalize the ratios S1/S2 and T1/T2; thus, it follows that S1/T1 = S2/T2 (Jones & Huang, 1982). In this way, the visual system attempts to equalize the velocity between the first and second flashes (S1/T1) and that between the second and third flashes (S2/T2). This assumption is called the constant velocity assumption. Jones and Huang (1982) also suggested that, for the judgment of spatial intervals between flashes, the visual system makes use of a weighted average of the physical spatial intervals and expected spatial intervals that would be traversed at a given velocity. Recently, employing the Bayesian observer model, Goldreich (2007) demonstrated that the brain automatically incorporates prior expectation for slow speed when judging spatiotemporal intervals defined by successive tactile stimulations, indicating that velocity assumption seems to underlie the perception of the extent of spatial intervals between successive sensory stimulations.

We were interested in the tau effect induced by dynamic stimuli that were audiovisually defined. Audition is often more dominant than vision in several tasks in which observers judge the temporal properties of stimuli. For example, a perceived visual flicker rate is strongly affected by a task-irrelevant auditory flutter rate, but the reverse is unlikely (Gebhard & Mowbray, 1959; Recanzone, 2003; Shipley, 1964; Vatakis & Spence, 2006; Wada, Kitagawa, & Noguchi, 2003). Similarly, the dominance of audition over vision is observed in judging the duration of stimuli defined by auditory tones as well as visual continuous light (Walker & Scott, 1981).

Based on the suggestions made by previous studies, we inferred that perceived temporal intervals between auditory stimulations would strongly distort the perceived temporal intervals between visual flashes, and consequently, that the distorted temporal intervals between visual flashes would distort the perception of spatial intervals between them if the visual system imputes uniform motion to the stimuli. Recent studies have examined the effects of audition on vision in the perception of space. For example, an auditory transient signal emitted at the coincidence of two visual moving objects sometimes induces the percept of bouncing in a streaming-bouncing motion display (Sekuler, Sekuler, & Lau, 1997). By contrast, continuous auditory signals speeded the responses for the percept of streaming (Sanabria, Lupianez, & Spence, 2007). Furthermore, auditory motion affects the detection of visual motion (Meyer & Wuerger, 2001; Meyer, Wuerger, Rohrbein, & Zetzsche, 2005; Wuerger, Hofbauer, & Meyer, 2003), and vice versa (Soto-Faraco, Kingstone, & Spence, 2003; Soto-Faraco, Spence, & Kingstone, 2005). The direction of a pitch change alters the percept of motion direction in an ambiguous motion signal (Maeda, Kanai, & Shimojo, 2004). Finally, the location of auditory signals modulates shifts in visual attention (Arnott & Goodale, 2006; Shimojo, Miyauchi, & Hikosaka, 1997). However, none of the previous studies could predict how visual temporal intervals distorted by auditory temporal intervals would lead to the distortion of visual spatial intervals.

Imagine audiovisual stimuli consisting of three successive auditory stimulations as well as three successive visual flashes. When three successive auditory stimulations produce two temporal intervals of different magnitudes ($aT1 < aT2$, where aT means auditory temporal interval and the following number refers to the presentation order of the interval) and three successive visual stimulations produce two temporal intervals with the same magnitude ($vT1 = vT2$, where vT means visual temporal interval), the perceived duration of vT will be $vT1 < vT2$ because aT is strongly expected to modulate the perceived duration of vT (Walker & Scott, 1981). Moreover, given visual spatial intervals with equal durations ($vS1 = vS2$, where vS represents visual spatial intervals and the following number refers to the presentation order of the interval), the distorted visual temporal intervals ($vT1 < vT2$) will distort the visual spatial intervals as $vS1 < vS2$, since the visual system tries to equalize the ratios $vS1/vT1$ and $vS2/vT1$ (Jones & Huang, 1982)

In this study we first examined whether temporal intervals between successive brief pure tones distorted the spatial intervals between successive disks in accordance with the constant velocity assumption. Moreover, we confirmed which of the perceptual differences in duration between auditory and visual temporal intervals or the perceived temporal offset between pure tone and disk was important. Based on the results, we discuss the relationships in audiovisual integration, the perception of spatiotemporal intervals, and the constant velocity assumption.

Method

Observers

Five and four people, including the first author (TK), served as observers in Experiments 1 and 2, respectively. The four observers participating in Experiment 2 also participated in Experiment 1. Except for the author, observers were unaware of the specific purpose of this experiment. They reported that they had normal or corrected-to-normal visual acuity as well as normal hearing ability.

Apparatus

A Macintosh computer (MacPro, Apple) controlled stimulus presentation and data collection. Visual stimuli were presented on a 19-inch CRT monitor with 1024 x 768 pixels. Loudspeakers (EMC2.0-USB, Diamond Audio Technology, Inc) located on both sides of the CRT monitor were employed to present auditory stimuli. The physical simultaneity of visual and auditory onsets was ensured by using an oscilloscope (V-302B, 30MHz, HITACHI, Japan) with a phototransistor (TPS603A, Toshiba, Japan) and sound-level meter (LA-5110, Ono Sokki, Japan) with an accuracy of less than 5 ms. Chin and head rests (TKK930A, Takei, Japan) were used to stabilize the observers' visual fields.

Stimuli

A schematic illustration of the stimulus presentation is shown in Figure 1. Stimuli were generated and presented using MATLAB and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Experiment 1 At the initiation of each trial, a fixation cross (vertical and horizontal lines each 0.19 deg x 1.89 deg) was presented at the center of display for 300 ms. Two hundred milliseconds after the disappearance of the fixation cross, three disks each having a radius of 1.4 deg, a luminance of 3 cd/m^2, and a duration of 16.7 ms were successively flashed. Three pure tones each having 1000 Hz frequency, 10 ms presentation duration, and 60 dB sound pressure level accompanied the presentation of the disks. The successive presentation of disks defined two temporal intervals with equal durations of 167 ms. The vertical positions of the disks were always set at the vertical center of the display. The horizontal apparent motion of the disks was initiated at 4.73 deg left or right, and was terminated at 4.73 deg right or left from the horizontal center of the display. The direction of apparent motion was randomly determined in each trial. The horizontal position of the second visual flash was also controlled: the disk was located with spatial offsets of -2.37, -1.18, 0, 1.18, and 2.37 deg from the horizontal center of the display (here, negative and positive values indicate spatial offsets to the left or right from the center of the display, respectively). Thus, one spatial interval was a distance of 2.37, 3.55, 4.73, 5.20 or 7.10 deg, and the other interval was 7.10, 5.20, 4.73, 3.55 or 2.37 deg. In Experiment 1, three successive pure tones were presented to define two temporal intervals: either the first temporal interval was 67 ms and the second was 267 ms, or the first temporal interval was 267 ms and the second was 67 ms. Hereafter, we designate the first and second temporal interval conditions for

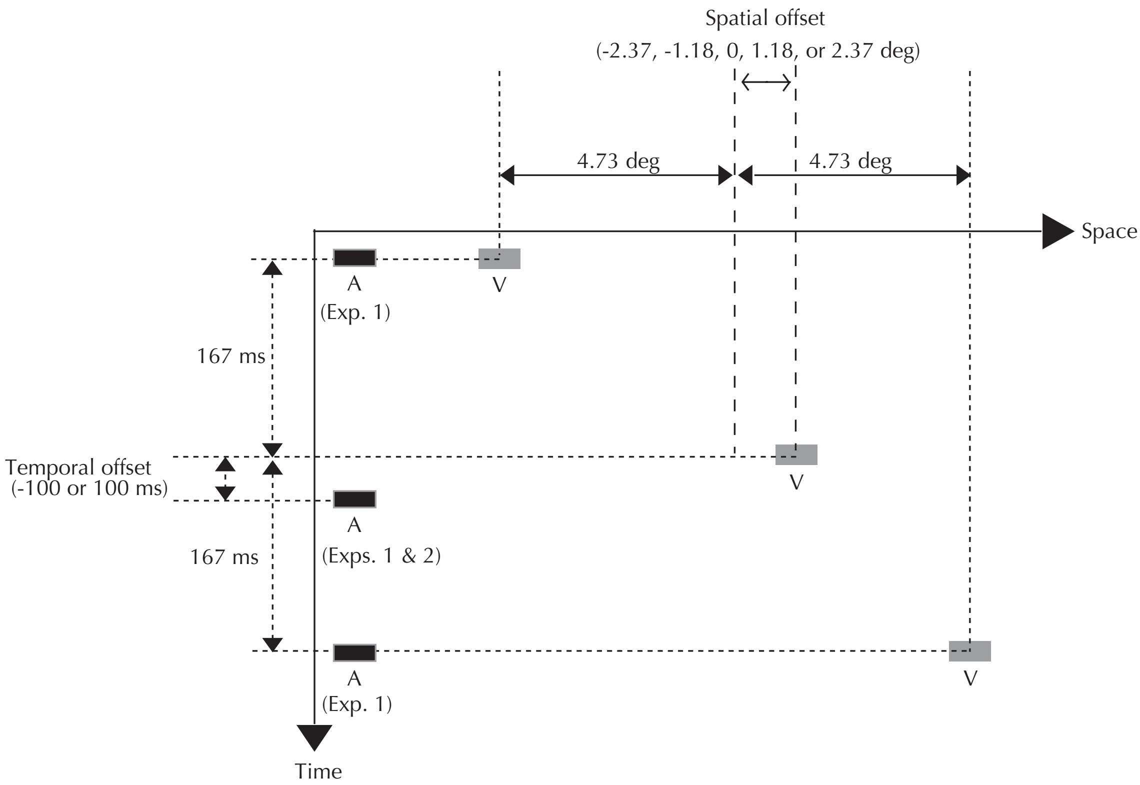

Figure 1. A spatiotemporal diagram of the stimulus presentation used in Experiments 1 and 2.

pure tones as -100 and 100 Stimulus-Onset-Asynchrony (SOA) conditions, respectively, since the onset of the second pure tone appeared 100 ms earlier or later than the onset of the second disk. We have chosen a 100-ms temporal offset between onsets of pure tone and disk, since a 100-ms offset was shown to be the temporal boundary condition for audiovisual integration (Fendrich & Corballis, 2001). The onsets of the first and third pure tones were always temporally matched to those of the first and third disks. We also used a no-tone condition, in which no pure tone was presented.

Experiment 2 The stimuli were identical to those in Experiment 1 except for the following: the first and third pure tones were withdrawn from the stimuli. Thus, three disks were successively presented, while pure tones were presented with temporal asynchrony of -100 or 100 ms from the onset of the second disk, as in Experiment 1.

Procedure The experiment was conducted in a darkened room. Observers sat 40 cm from a CRT display. Each trial started by pressing the spacebar on the keyboard of the computer. Observers reported which of the first or second spatial intervals between disk presentations was the shorter. Observers responded by pressing assigned keys. In each experiment, an observer received 300 trials consisting of 5 (spatial offsets of the second visual flash) x 3 (2 SOAs conditions of auditory pulses plus 1 no-tone condition) x 20 replications. The trial order was pseudo-randomized across observers. For each experiment, it took 30 min to complete all the trials, including several short breaks.

Results

Experiment 1

The averaged proportions of trials in which the first spatial interval was perceived to be shorter are shown in Figure 2a. For each condition and each observer, we calculated the point of subjective equality (PSE) by fitting a psychometric function (cumulative Gaussian distribution) with probit analysis (Finney, 1971) implemented in the statistical software (SPSS). The PSE was the spatial offset corresponding to 50% in psychometric function. We assessed the goodness of fit by calculating Pearson's Chi square value, and confirmed that the cumulative Gaussian function was well fitted ($p > 0.05$). Group PSE data is shown in Figure 2b. We analyzed group PSE data using a one-way ANOVA with pure tones condition as a factor. The main effect of the pure tone conditions was significant (F (2, 8) = 12.650, p < 0.004). Multiple comparison tests showed that the PSEs in the 100 ms SOA condition were significantly greater than those in the no tone and -100 ms SOA conditions (*ts* (4) = 2.874 and 5.012, respectively, p < 0.05), and a marginally significant difference was observed between the no-tone and 100 ms SOA conditions (*ts* (4) = 2.138, p = 0.065). We also compared the sample effect sizes among conditions by calculating Cohen's d. The sample effect sizes in the different conditions are shown in Table 1. The d values were generally large in comparison with the effect size index proposed by Cohen (1992) where d values of 0.2, 0.5 and 0.8 are suggested indices for small, medium and large effect sizes, respectively.

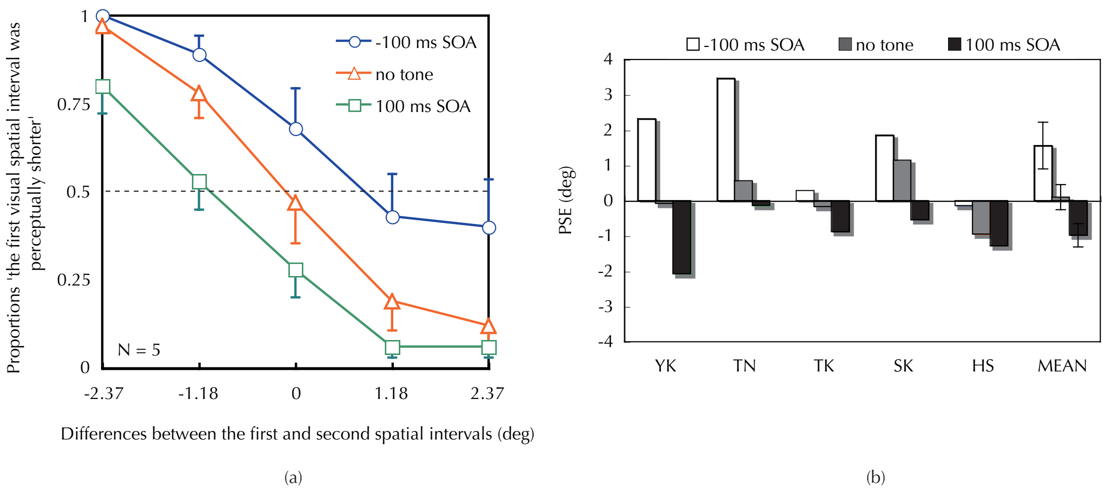

Figure 2. The results of Experiment 1. (a) The proportions of trials in which the first visual spatial interval was perceived to be shorter than the second spatial interval. The error bars denote standard errors of the means. (b) Individual and averaged data on the PSEs for visual spatial intervals in each auditory pulse condition. The error bars denote standard errors of the means.

Experiment 2

As in Experiment 1, we fitted a psychometric function to the proportions of trials in which the first spatial interval (Figure 3a) was perceived to be shorter than the second spatial interval, and calculated the PSE for each condition (Figure 3b). We analyzed group PSEs using a one-way ANOVA with pure tones condition as a factor. The main effect of the pure tone conditions was not significant (F (2, 6) = 1.505, p > 0.2). As in Experiment 1, we also calculated Cohen's *d*. The sample effect sizes among conditions are shown in Table 1. The *d* values were generally low in comparison with the effect size index proposed by Cohen (1992), as described above.

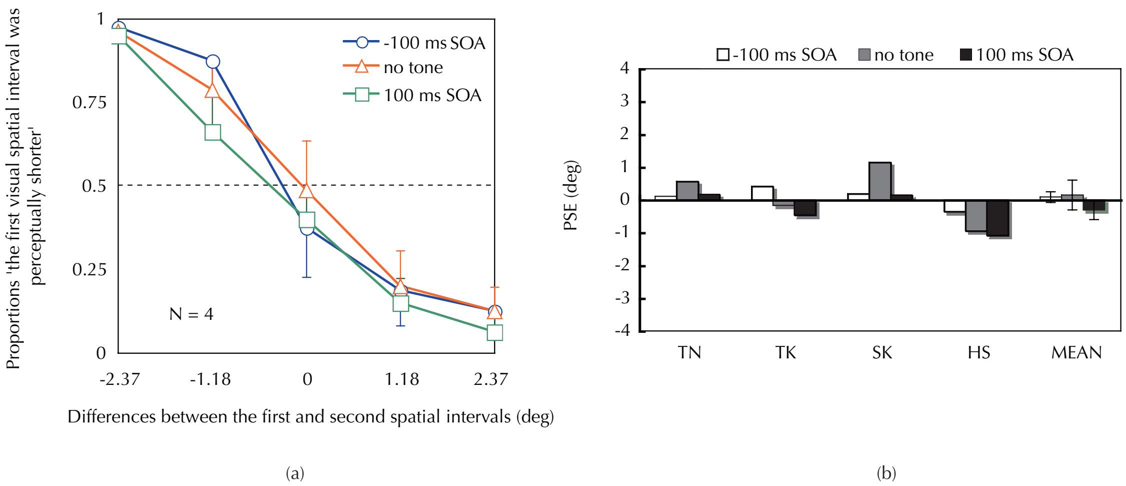

Figure 3. The results of Experiment 2. (a) The proportions of trials in which the first visual spatial interval was perceived to be shorter than the second spatial interval. The error bars denote standard errors of the means. (b) Individual and averaged data on the PSEs for visual spatial intervals in each auditory pulse condition. The error bar denotes standard errors of the means.

Table 1. Sample effect sizes (Cohen's *d* values) between tone conditions in Experiments 1 and 2.

	Pairs of tone conditions		
	-100 ms SOA and no tone	No tone and 100 ms SOA	-100 ms SOA and 100 ms SOA
Experiment 1	1.22	1.41	2.17
Experiment 2	0.09	0.59	0.7

Discussion

This study examined whether temporal intervals defined by three successive pure tones affected the perception of spatial intervals defined by three successive disks eliciting apparent motion. In Experiment 1, we found that shorter temporal intervals between pure tones caused the perception of shorter spatial intervals between disks. However, in Experiment 2, employing stimuli in which the first and third pure tones were withdrawn, we showed that the temporal offset between the onsets of the second pure tone and second disk was not a sufficient condition to induce the illusory perception of spatial interval observed in Experiment 1.

Our results are consistent with the idea that the constant velocity hypothesis applies to moving objects defined by more than one sensory stimulation. As described in the Introduction, the visual system imputes motion with constant velocity to an apparent motion display with visual flashes defining different temporal intervals (Jones & Huang, 1982). Thus, the perceived distance traveled by visual flashes needs to be smaller when the perceived time interval between flashes is smaller. In our audiovisual stimulation, the temporal intervals between tones and disk presentations may be independently estimated, and later combined. Consequently, the temporal interval between the disks seemed to be strongly modulated by that between the pure tones (Walker & Scott, 1981). The constant velocity assumption predicts that the visual system would estimate that the spatial interval between disks would be shorter when the temporal interval between disks was underestimated by the concurrent presentation of short temporal intervals between pure tones.

Our results are consistent with those of Mamassian (2005) in that the estimation of the speed of moving objects is performed after auditory and visual information are combined. Mamassian showed that the perceived speed of a ‘stationary drifting’ Gabor patch was modulated by the concurrent presentation of auditory flutters. More specifically, the faster (or slower) an auditory flutter, the faster (or slower) was the perceived speed. In the present study, we also suggest that the constant velocity assumption applies to moving objects defined by auditory and visual stimulation, indicating that the computation of speed is performed after auditory and visual information are combined. However, since it was unclear whether the perceived extents of spatial intervals (i.e. phase shifts in stationary drifting Gabor patches) were also modulated by auditory flutter rate in the study by Mamassian, more work is needed to clarify the mechanism underlying the modulation of visual motion speed by auditory temporal stimulations.

The experiments in the present study might have been performed with a high uncertainty of disks’ locations. The standard deviation of the fitted cumulative Gaussian function in the no tone condition was 1.5 deg, and this value indicates that the task in our experiment was difficult for our observers. In general, the initial, current and final positions of moving objects are perceived with many sorts of localization bias, such as the attentional repulsion effect (Suzuki & Cavanagh, 1997), the Fröhlich effect (Fröhlich, 1923), spatiotemporal position integration (Roulston, Self, & Zeki, 2006), and representational momentum (Freyd & Finke, 1984). Thus, we surmise that judging visual spatial intervals might suffer from complex interactions among visual factors, causing high uncertainty of disks’ location, resulting in difficulty with the task.

Interestingly, a recent study (Flach & Haggard, 2006) provided evidence that the cutaneous rabbit illusion in the tactile dimension is not affected by concurrent auditory stimuli. Specifically, the perceived duration of spatial intervals between tactile events was not altered by the duration of auditory temporal intervals. The results of Flach and Haggard's study are inconsistent with ours, and further research is necessary to resolve this discrepancy. Flach and Haggard did not measure pure temporal modulations of tactile events by auditory events; thus, it was unclear whether the distortion of perceptual temporal intervals, which is necessary for the distortion of spatial temporal intervals, could occur. Moreover, since there is a possibility that the effectiveness and reliability of each unimodal stimulus are different between Flach and Haggard (tactile and auditory stimuli) and the present study (visual and auditory stimuli), it is difficult to directly compare the results of these studies. In order to draw conclusions about the differences between the studies, we would need to equate the unimodal stimuli in terms of their effectiveness and reliability for the particular task.

References

Arnott, S. R., & Goodale, M. A. (2006). Distorting visual space with sound. *Vision Research*, 46, 1553-1558.

Benussi, V. (1913). Versuche zur Analyse taktil erweckter Scheinbewegungen (Contribution to the analysis of tactile processing of apparent motions). *Archiv für die Gesamte Psychologie*, 36, 58-1315.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433-436.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics*, 63, 719-725.

Finney, D. J. (1971). *Probit analysis*. Cambridge, UK: Cambridge University Press.

Flach, R., & Haggard, P. (2006). The cutaneous rabbit revisited. *Journal of Experimental Psychology: Human Perception and Performance,* 32, 717-732.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 126-132.

Fröhlich, F. W. (1923). Uber die Messung der Empfindungszeit [Measuring the time of sensation]. *Zeitschrift für Sinnesphysiologie*, 54, 58–78.

Gebhard, J. W., & Mowbray, G. H. (1959). On discriminating the rate of visual flicker and auditory flutter. *American Journal of Experimental Psychology*, 72, 521-528.

Goldreich, D. (2007). A Bayesian perceptual model replicates the cutaneous rabbit and other tactile spatiotemporal illusions. *PLoS ONE*, 2, e333. doi:10.1371/journal.pone.0000333

Helson, H. (1930). The tau effect - an example of psychological relativity. *Science*, 71, 536-537.

Helson, H., & King, S. M. (1931). The tau effect: an example of psychological relativity. *Journal of Experimental Psychology*, 14, 202–217.

Jones, B., & Huang, Y. L. (1982). Space-time dependencies in psychological judgment of extent and duration: algebraic models of the tau and kappa effects. *Psychological Bulletin*, 91, 128–142.

Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, 14, R990-991.

Mamassian, P. (2005). Auditory tones influence perceived speed in apparent motion [Abstract]. *Journal of Vision*, 5, 877, http://journalofvision.org/5/8/877/, doi:10.1167/5.8.877.

Meyer, G. F., & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport*, 12, 2557-2560.

Meyer, G. F., Wuerger, S. M., Rohrbein, F., & Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, 166, 538-547.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437-442.

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology*, 89, 1078-1093.

Roulston, B. W., Self, M. W., & Zeki, S. (2006). Perceptual compression of space through position integration. *Proceedings of the Royal Society of London, B, Biological Sciences*, 273, 2507-2512.

Sanabria, D., Lupianez, J., & Spence, C. (2007). Auditory motion affects visual motion perception in a speeded discrimination task. *Experimental Brain Research*, 178, 415-421.

Sarrazin, J., Giraudo, M., & Pittenger, J. B. (2007). Tau and Kappa effects in physical space: the case of audition. *Psychological Research*, 71, 201-218.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385, 308.

Shimojo, S., Miyauchi, S., & Hikosaka, O. (1997). Visual motion sensation yielded by non-visually driven attention. *Vision Research*, 37, 1575-1580.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science*, 145, 1328-1330.

Soto-Faraco, S., Kingstone, A., & Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, 41, 1847-1862.

Soto-Faraco, S., Spence, C., & Kingstone, A. (2005). Assessing automaticity in the audiovisual integration of motion. *Acta Psychologica*, 118, 71-92.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: an attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 443-463.

Vatakis, A., & Spence, C. (2006). Temporal order judgments for audiovisual targets embedded in unimodal and bimodal distracter streams. *Neuroscience Letters*, 408, 5-9.

Wada, Y., Kitagawa, N., & Noguchi, K. (2003). Audio-visual integration in temporal perception. *International Journal of Psychophysiology*, 50, 117-124.

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 1327-1339.

Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics*, 65, 1188-1196.