

Static sound timing alters sensitivity to low-level visual motion

Hulusi Kafaligonul

Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA



Gene R. Stoner

Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA



Visual motion processing is essential to survival in a dynamic world and is probably the best-studied facet of visual perception. It has been recently discovered that the timing of brief static sounds can bias visual motion perception, an effect attributed to “temporal ventriloquism” whereby the timing of the sounds “captures” the timing of the visual events. To determine whether this cross-modal interaction is dependent on the involvement of higher-order attentive tracking mechanisms, we used near-threshold motion stimuli that isolated low-level pre-attentive visual motion processing. We found that the timing of brief sounds altered sensitivity to these visual motion stimuli in a manner that paralleled changes in the timing of the visual stimuli. Our findings indicate that auditory timing impacts visual motion processing very early in the processing hierarchy and without the involvement of higher-order attentional and/or position tracking mechanisms.

Keywords: cross-modal interactions, temporal ventriloquism, motion perception, motion processing, multisensory

Citation: Kafaligonul, H., & Stoner, G. R. (2012). Static sound timing alters sensitivity to low-level visual motion. *Journal of Vision* 12(11):2, 1–9, <http://www.journalofvision.org/content/12/11/2>, doi:10.1167/12.11.2.

Introduction

Perception has traditionally been viewed as modular with different sensory modalities operating largely independently. New reports of various types of cross-modal sensory interactions have, however, overturned this dogma. Of particular interest here is a growing realization of the importance of crossmodal interactions in perception of the fundamental attributes of space and time. Perceptual interactions between auditory and visual stimuli have been well studied in this context: When put in conflict, visual stimuli can drive the perception of *where* a sound originates (*spatial ventriloquism*) (Bertelson & Aschersleben, 1998; Howard & Templeton, 1966) whereas auditory stimuli can drive the perception of *when* visual events occur (*temporal ventriloquism*) (Fendrich & Corballis, 2001; Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Recanzone, 2003). These interactions make adaptive sense given the auditory system’s superior temporal resolution and the visual system’s superior spatial resolution (Alais & Burr, 2004a; Alais & Burr, 2004b).

Several studies have reported that the timing of brief sounds can alter how visual apparent-motion (AM) stimuli are perceived, presumably by altering the perceived timing of the visual stimuli that yield the motion percept (temporal ventriloquism) (Freeman & Driver, 2008; Getzmann, 2007; Kafaligonul & Stoner,

2010; Shi, Chen, & Müller, 2010; Staal & Donderi, 1983). In these studies, the moving stimuli (e.g., rectangular bars) were localizable in each frame of the animation. The direction in which this type of stimulus is displaced can be extracted by both low-level motion energy mechanisms (Anstis, 1980; Braddick, 1980) and high-level attention-dependent positional tracking (Cavanagh, 1992; Lu & Sperling, 1995). Thus the mechanisms underlying these crossmodal temporal interactions, and the processing stage(s) at which they occur, have been unclear.

We had two goals in designing the experiments described here. First, we wanted to determine whether changes in auditory timing could mimic the impact that changes in visual timing have on motion discrimination. If so, this finding would suggest that auditory timing directly alters the neuronal processing of visual timing rather than acting at a decision stage after the processing of visual motion. Second, we wished to determine whether temporal ventriloquism could affect motion perception in the absence of higher-order attentional tracking. If so, this outcome would imply that auditory timing impacts visual processing at or before the level of low-level motion mechanisms, which are generally attributed to cortical areas V1 and/or MT.

To achieve these goals, we constructed random-dot AM stimuli that minimized attention-based position tracking since the moving object was not identifiable in

individual frames. Moreover, since the motion direction of these stimuli can only be discriminated for small interstimulus intervals (ISIs) between AM frames (unlike the bar stimuli used in previous related experiments for which the direction of motion was always easily discriminable), we were able to assay the influence that static sound timing had on motion discrimination. We found that the timing of brief sounds systematically changed motion discrimination thresholds for these low-level AM stimuli (either increasing or decreasing sensitivity). We also found that auditory timing changed sensitivity to *reverse-phi* motion (a visual motion illusion attributed to low-level motion energy mechanisms) (Anstis, 1970). Our findings demonstrate for the first time that the timing of brief sounds can impact the ability to discriminate motion direction and, moreover, that sounds can affect motion processing even when attentional tracking is ruled out. These findings suggest that auditory timing impacts visual motion processing very early in the visual motion processing stream.

Experiment 1: Random-dot apparent motion

In [Experiment 1](#), we examined the effect of auditory timing on the sensitivity of low-level visual motion by using near-threshold random-dot stimuli that discouraged high-level positional tracking. We determined the visual ISI (ISI_v) that yielded threshold direction-discrimination performance when these stimuli were accompanied by static sounds with different auditory ISIs. If auditory timing alters the processing of visual timing prior to the computation of low-level motion, then these thresholds should depend on the auditory ISI. Conversely, if positional tracking is necessary for sounds to impact motion perception, then these thresholds should be unaffected by auditory timing. More generally, if the auditory influence on AM is limited to a change in perceptual bias rather than a change in sensitivity, we should find that sound timing has no effect on direction discrimination.

Method

Participants

Six human observers completed this experiment with four being naïve to the purpose of the experiment. In all of the experiments, observers had normal hearing and normal or corrected-to-normal visual acuity. Participants gave informed consent, and all procedures were in accordance with international standards (Declaration of Helsinki) and NIH guidelines.

Apparatus

We used the CORTEX program (Laboratory of Neuropsychology, National Institute of Mental Health) for stimulus presentation and data acquisition. Visual stimuli were presented on a 19" CRT monitor (Sony Trinitron E500, 1024×768 pixel resolution and 100 Hz refresh rate) at a viewing distance of 57 cm. A PR701S photometer was used for luminance calibration and gamma correction of the display. Sounds were emitted by two speakers (ALTEC Lansing) positioned at the top of the visual display and amplitudes were measured with a sound-level meter. Timing of visual and auditory stimuli was confirmed with a digital (Tektronix TDS 1002) oscilloscope connected to the computer soundcard and a photodiode. Head movements were constrained by a chin rest. The same apparatus was used in all the experiments.

Stimuli and procedure

A small red circle (10.9 arc-min diameter) at the center of the display served as a fixation target. Visual stimuli consisted of three “flashed” (50 ms) random-dot (8.6×8.6 deg) frames presented with the inter-stimulus interval (ISI_v) chosen pseudorandomly from eight values: 20, 40, 60, 80, 100, 120, 160, and 200 ms. During the ISI_v , the entire display was a mean gray. Each dot (4.36×4.36 arc-min) in the random-dot field was either lighter (45.5 cd/m^2) or darker (0.09 cd/m^2) than the gray background (22.8 cd/m^2). As indicated by the green outline (not present in the actual stimulus) in [Figure 1a](#), a square region (4.3×4.3 deg) of the random-dot field was displaced 13.08 arc-min (leftward or rightward) in each of the two consecutive presentations. There was nothing in the individual frames that indicated the spatial boundaries of the moving object. Moreover, the background dots (i.e., the dots that were not displaced) were randomly updated with each displacement so that all positions in the dot field were equally likely to be replaced with a dot of the opposite polarity. Accordingly, there were no temporal cues for the position or the motion direction of the moving dots ([Figure 1b](#)). Auditory stimuli were three 10 ms clicks. Each click was comprised of a rectangular windowed 480 Hz sine-wave carrier, sampled at 22 kHz with 8-bit quantization.

As shown in [Figure 2a](#), each experimental session had a balanced mixture of two *auditory* stimulus conditions: *shorter* (auditory ISIs were 20 ms) and *longer* (auditory ISIs were 70 ms longer than visual ISIs), and one *visual-only* (no clicks) stimulus condition. Every stimulus was presented eight times per session. Accordingly, there were 192 trials (3 auditory conditions \times 8 ISI_v conditions \times 8 trials per stimulus) in each experimental session.

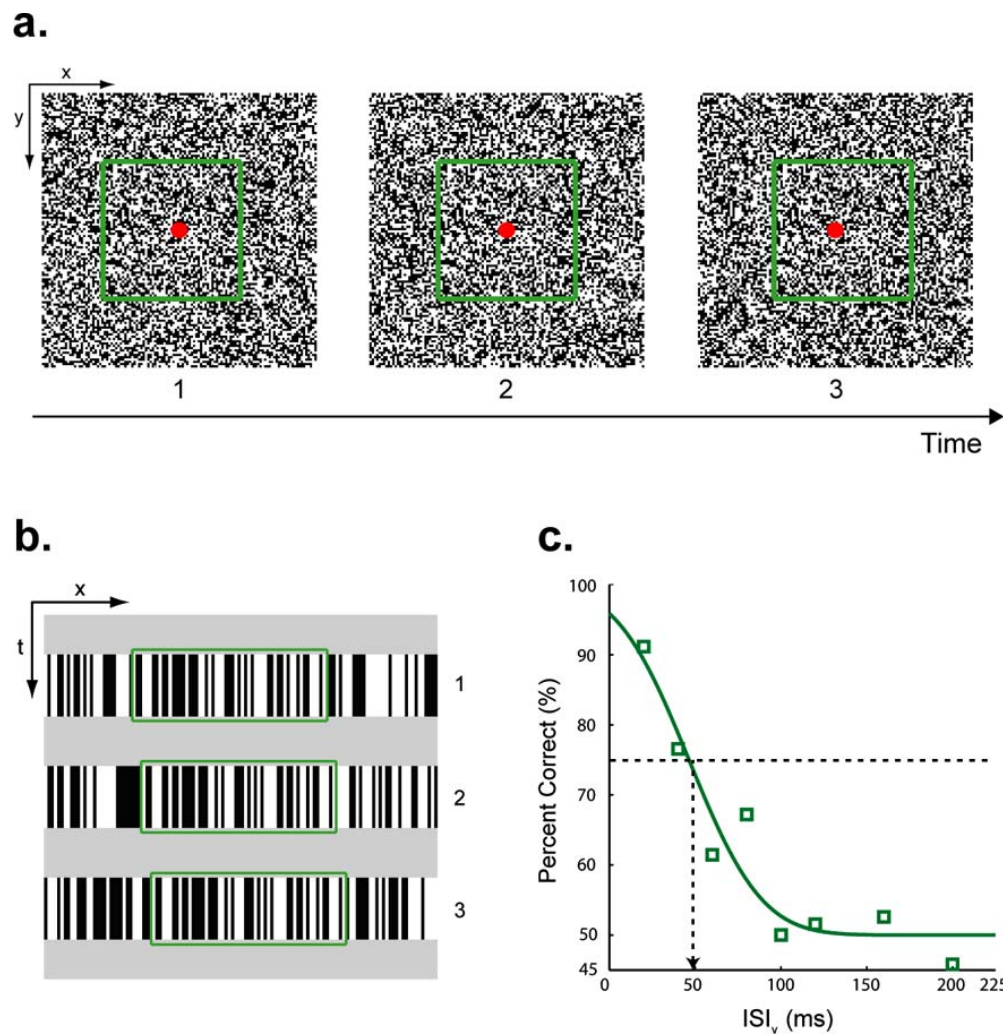


Figure 1. (a) Stimuli used in [Experiment 1](#). Apparent motion (AM) stimuli consisted of three random-dot frames. A central square region (4.3×4.3 deg) of each random-dot frame (green outline, not present in actual stimulus) was displaced sequentially either from left-to-right or right-to-left. The background dots (i.e., dots outside the square region) were updated randomly at each displacement. Auditory clicks were emitted by speakers positioned above display. (b) Space-time plots of the three random-dot frames. Each frame is shown as a single slice along the vertical spatial axis. The horizontal and vertical axes in the space-time plots correspond to the horizontal spatial axis and time, respectively. In the example shown here, the center dots, as shown by the green outline, moved left-to-right. The background dots were randomly updated at each spatial displacement. The entire display was a mean gray between AM frames. (c) Group averaged data ($n = 6$; four were naïve observers) for visual-only condition. Performance improves as the ISI between visual frames (ISI_v) decreases. A cumulative Gaussian function was fitted to the data. The ISI_v corresponding to 75% correct (vertical arrow) was our estimate of motion discrimination threshold.

Subjects engaged in a left-right motion direction discrimination task (two-alternative forced-choice). They sat in a dark room and fixated a red circle at the center of the display. They were told that there would be visual motion at the center of the display. They were instructed neither to attend dots (individual or groups) nor to track their displacement. They were told that some visual stimuli would be accompanied by clicks but to base their responses solely on the visual stimuli. New random-dot stimuli were generated for each trial to rule out tracking of previously identified dot patterns. Subjects started each trial by pressing a

key after which experimental stimuli were presented. At the end of each trial, observers indicated, by pressing left or right arrow keys, whether the center of the random-dot field moved leftward or rightward. Each subject completed four experimental sessions (for a total of 768 trials). Prior to these experimental sessions, each participant was shown examples of visual-only AM stimuli followed by two practice sessions of 192 visual-only trials without feedback. If a participant could not achieve threshold performance during the practice sessions, he or she was not included in the main experimental sessions.

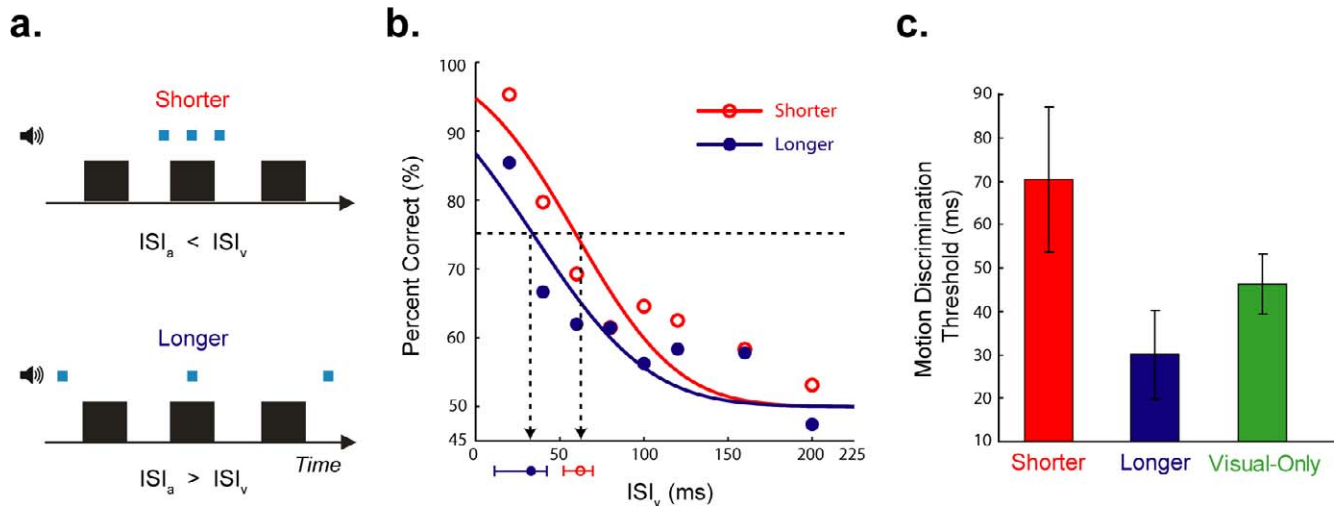


Figure 2. (a) Timing diagram for the two auditory conditions of Experiment 1. The visual frames and clicks are indicated by black rectangles and blue squares, respectively. Relative durations of visual and audio events are indicated by thickness of squares and rectangles (height of these icons distinguishes stimulus modality and is otherwise irrelevant). When the ISI_v was 20 ms, the ISI_a was equal to this value for the shorter auditory timing condition. (b) Results of Experiment 1. Group averaged data ($n = 6$; four were naïve observers) for different auditory timing conditions. The plot shows performance as a function of ISI_v . The open and filled symbols represent shorter and longer auditory conditions, respectively. The red and blue curves indicate corresponding psychometric fits for these auditory conditions. The ISI_v corresponding to 75% correct (vertical arrow) was our estimate of motion discrimination threshold. The error bars near the horizontal axis are the 95% confidence intervals for these estimates. (c) The averaged motion discrimination thresholds of all observers for different auditory conditions. Error bars correspond to $\pm SEM$.

Threshold estimation

We calculated the average performance across observers for the different visual ISIs (ISI_v) and auditory (including visual-only) conditions. To estimate motion discrimination threshold for each auditory condition, a cumulative Gaussian function was fitted to the group-averaged data by using *psignifit* (version 2.5.6), a software package that implements the maximum likelihood method described by Wichmann and Hill (2001a). The ISI_v that yielded 75% correct level was defined as the motion discrimination threshold for each auditory condition. The 95% confidence intervals of the threshold estimates were obtained by the BCa bootstrap method implemented in *psignifit*, based on 1999 simulations (Wichmann & Hill, 2001b). A threshold difference between two auditory timing conditions was considered significant if the 95% confidence intervals did not overlap. In addition to these analyses of the group-averaged data, for each observer we used the same procedure to estimate individual motion discrimination thresholds for the shorter, longer, and visual-only conditions. These individual thresholds were then compared by paired samples *t*-tests.

Results

As shown in Figure 1c, motion discrimination was significantly (one way repeated measures ANOVA:

$F[7, 35] = 18.121, p < 0.001$) dependent on ISI_v : As ISI_v increased, average performance (i.e., percent correct) decreased, approaching chance level (50%) for the largest ISI_v s. If auditory timing impacts visual processing before motion direction is computed, then we would expect to see different discrimination thresholds for the two auditory timing conditions. Alternatively, if auditory timing affects visual processing at a later decision stage, then we would not expect any difference in the discrimination thresholds found under the two conditions. In agreement with the early-impact account, the ISI_v that yield threshold performance was significantly larger for the shorter auditory condition than for the longer auditory condition. In other words, the shorter auditory ISIs resulted in a shortening of the ISI_v relative to the longer auditory ISIs, thus yielding better direction discrimination. This statistical dependency of motion threshold on auditory timing was found in both the group-averaged threshold estimates (Figure 2b, $p < 0.05$) and in the threshold estimates from individual observers (Figure 2c, paired-samples two-tailed *t* test: $t[5] = 2.653, p < 0.05$). Motion discrimination thresholds for the visual-only condition fell between these two auditory conditions. On the other hand, based on both group-averaged and individual observers data, the slope difference between two auditory conditions was not significant (paired-samples two-tailed *t* test: $t[5] = -1.003, p = 0.362$).

The three-frame AM design used in these experiments evolved from earlier two-frame AM experiments. Rather than using two auditory timing conditions, in those initial experiments we varied auditory ISIs from 20 to 380 ms (see the [Supplemental Material](#) for additional description). Subjects found these two-frame stimuli hard to discriminate (many subjects were unable to reach threshold performance) and this led to our development of the three-frame stimuli. While data from that earlier experiment were noisy, we did find that the ISI_v s that yielded threshold performance were significantly larger for the smallest auditory ISI (20 ms) than for the largest auditory ISI (380 ms) ([Figure S1](#) in the [Supplemental Material](#)). That finding thus supports the data presented here, which provide the first evidence that changes in the auditory ISI can alter sensitivity to visual motion in a manner that mimics changes in visual ISI. These findings suggest that auditory timing exerts its effect on visual processing before the computation of visual motion.

Experiment 2: Reverse-phi motion

The design of [Experiment 1](#) (as well as the instructions to our subjects) was intended to isolate low-level mechanisms. Hence the results of that experiment suggest that auditory timing alters the neuronal signals that feed into low-level motion processing rather than into a feature-tracking mechanism. Nevertheless, subjects could have conceivably identified random patterns in individual frames and then noted positional changes of these features. The results of [Experiment 1](#) thus leave open the possibility that auditory timing impacts visual timing before the involvement of tracking mechanisms but after the computation of low-level motion. To definitively rule out a tracking strategy and to isolate low-level (motion-energy) motion mechanisms, we used reverse-phi motion in which the luminance contrast of successive frames is reversed (Anstis, 1970). Unlike in phi (i.e., regular) motion, the perceived direction of motion in these displays is in the direction *opposite* of the displaced features, thus inconsistent with the tracking of those features. This perceptual illusion is instead consistent with low-level motion energy mechanisms (Adelson & Bergen, 1985), which are believed to be implemented within early levels of motion processing (Emerson, Citron, Vaughn, & Klein, 1987; Krekelberg & Albright, 2005). In [Experiment 2](#), we asked whether the strength of the reverse-phi percept was affected by auditory timing. If the ability of sound timing to affect motion processing were dependent on tracking mechanisms, the reverse-phi percept should be unaffected by auditory timing.

Method

Participants

Five observers completed this experiment with four being naïve to the purpose of the experiment. Four of these observers took part in [Experiment 1](#).

Stimuli and procedure

Unlike in [Experiment 1](#), in which our dependent variable was the ISI_v that yielded threshold performance, in [Experiment 2](#) our dependent variable was the percent of reverse-phi reports for a *fixed* ISI_v . We asked whether auditory timing altered the probability of those perceptual reports. Our preliminary experiments revealed that detecting motion in reverse-phi displays was more difficult than in our phi displays. To facilitate detection, we changed the design somewhat from [Experiment 1](#). We eliminated the background dots, thus removing their masking effect as well as uncertainty as to the location of the moving stimulus. The removal of the background dots necessitated moving the dots within a fixed virtual aperture (2.9×2.9 deg) so that motion direction could not be determined by tracking the boundary of the dot field. Pilot data confirmed that these changes yielded substantially improved performance relative to the old design. The design of [Experiment 2](#) was otherwise like that of [Experiment 1](#). Three random-dot frames moved either rightward or leftward. The ISI_v was held constant at 30 ms ([Figure 3a](#)). As in [Experiment 1](#), each session had a balanced mixture of two auditory timing conditions: *shorter* (auditory ISIs were 20 ms) and *longer* (auditory ISIs were 100 ms), and one *visual-only* (no clicks) condition. On half of the trials phi (i.e., not contrast reversed) stimuli were presented. For the other half (randomly intermixed), we presented reverse-phi stimuli in which the second frame of the three-frame sequence had reversed contrast relative to the other two frames. The mixing of the two motion types discouraged subjects from adopting a different strategy to recover the motion of reverse-phi stimuli. Subjects were not told that there were two types of motion stimuli and, upon debriefing, none of the subjects reported that there were different types of stimuli.

A single experimental session consisted of 480 trials: 2 motion types (i.e., phi or reverse phi) \times 3 auditory conditions \times 80 trials per condition. The spatial displacement (i.e., center-to-center separation between each random-dot frame) was constant during each experimental session and chosen from three values: 4.36, 8.72, and 13.08 arc-min. Accordingly, for each subject, there were three experimental sessions corresponding to the three different spatial displacements (SDs). The order of these sessions was randomized for each subject. Prior to these experimental sessions, subjects were shown examples of random-dot AM

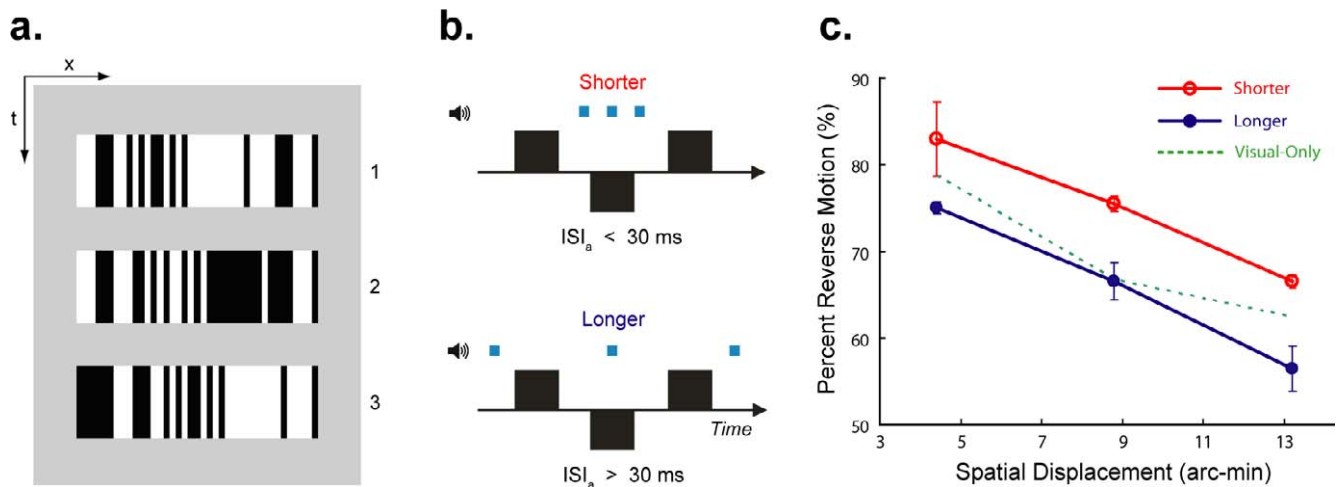


Figure 3. (a) Space-time plots for reverse-phi motion. The horizontal and vertical axes correspond to space and time, respectively. Dots moved behind a virtual stationary aperture. In this example, dots moved from left-to-right. As a result, consecutive frames introduced “new” dots at the left border of the aperture whereas “old” dots disappeared at the right border. Contrast reversal in the second frame led to a right-to-left motion percept, which is opposite to the left-to-right direction yielded by dot tracking (compare dot positions in frames 1 and 3). The entire display was a mean gray between apparent-motion frames. (b) Timing diagram for two auditory conditions of Experiment 2. The middle black rectangle indicates contrast reversal for the second random-dot frame. The ISI_v was held constant at 30 ms. Other conventions are the same as those in Figure 2a. (c) Results of Experiment 2 ($n = 5$; four were naïve observers). The plot indicates proportion of perceived direction opposite to the dot displacement direction as a function of spatial offset. The open and filled symbols represent shorter and longer auditory conditions, respectively. Error bars indicate $\pm SEM$. The dashed line indicates values for visual-only condition.

stimuli to confirm that they understood the task. Other stimulus parameters and procedures were the same as those in Experiment 1. For each spatial displacement and auditory timing, the perception of reverse-phi motion was quantified by dividing the number of responses corresponding to the opposite direction of dot displacement by the total number of reverse-phi trials (80 trials for each condition).

Results

For reverse-phi stimuli, subjects consistently reported the motion to be in direction opposite to that in which the dots were displaced. The average percentage of reverse-phi reports (across five observers) is shown in Figure 3c. In line with previous studies (e.g., Bours, Kroes, & Lankheet, 2009; Sato, 1989; Wehrhahn, 2006), the likelihood of these reports increased significantly as the spatial displacement between visual frames decreased (two-way repeated measures ANOVA, spatial displacement and auditory timing as factors: $F [2, 8] = 22.655$, $p < 0.01$). Most important, reverse-phi was significantly enhanced for the shorter relative to the longer auditory timing conditions ($F [1, 4] = 18.325$, $p < 0.05$). For phi stimuli, subjects consistently reported motion to be in the same direction as the dot displacement. The performance for these

stimuli was well above threshold for all conditions. While the effect of auditory timing on performance was not significant for these easy-to-discriminate stimuli ($F [1, 4] = 16.875$, $p = 0.221$), performance was better for the shorter than for the longer auditory condition for all three spatial displacements (Figure S2 in the Supplemental Material).

Discussion

We have found that the appropriate timing of brief static sounds can alter visual motion discrimination thresholds. Earlier studies had reported that the timing of brief sounds can render motion more salient (Freeman & Driver, 2008; Getzmann, 2007; Shi et al., 2010; Staal & Donderi, 1983) and/or alter perceived speed (Kafaligonul & Stoner, 2010). Our findings extend these earlier demonstrations that auditory timing can alter *how* motion is perceived by demonstrating that sound timing can determine *whether* motion is, in fact, perceived at all. Our findings, and those earlier findings, are consistent with growing evidence of *temporal ventriloquism* whereby sounds capture the timing of visual events (Fendrich & Corballis, 2001; Morein-Zamir et al., 2003; Recanzone, 2003). Unlike the stimuli used in previous studies of

temporal ventriloquism and visual motion perception, the visual stimuli used in the current study did not allow attentional position tracking and hence isolated low-level pre-attentive visual motion processing. Our findings thus imply that the neuronal mechanisms underlying temporal ventriloquism are to be found early in the visual processing stream.

Alternative explanations of our results

Since factors other than the auditory timing can impact motion perception, it is worthwhile considering whether explanations other than those based on temporal ventriloquism might account for our findings. It is conceivable, for example, that the shorter auditory timing condition served to better alert subjects to the onset of motion than did the longer auditory timing condition. This explanation seems very unlikely, however, given that the longer condition provided a sound before the onset of visual motion whereas the first sound of the short condition occurred after motion onset. Moreover, it is unclear how this explanation would account for the finding, in [Experiments 1](#) and [2](#), that performance on the visual-only condition was intermediate to the two timing conditions. One would have to propose that sounds can either alert or distract subjects, depending on their precise timing. Finally, an explanation in terms of differential alerting fails to account for previous findings that sound timing alters the perceived speed (Kafaligonul & Stoner, 2010) and perceived visual ISI for AM (Freeman & Driver, 2008). We conclude that the most parsimonious explanation for our current findings is temporal ventriloquism.

Auditory contributions to the perception of visual motion

While our study is the first to demonstrate that *static* sounds can impact visual motion sensitivity, there are several previous studies that have provided evidence that *moving* sounds can impact visual motion sensitivity (for a review see Soto-Faraco, Kingstone, & Spence, 2003). Although initial studies suggested auditory motion and visual motion interactions were restricted to the decisional level (Alais & Burr, 2004a; Meyer & Wuerger, 2001), recent findings have found that auditory motion can affect visual motion at the perceptual level (Kim, Peters, & Shams, 2011) and can alter sensitivity to visual motion (Alink et al., 2012; Meyer, Wuerger, Röhrbein, & Zetzsche, 2005; Sanabria, Spence, & Soto-Faraco, 2007). Moreover, it has been recently shown that auditory motion can even induce a perception of visual motion from stationary visual flashes (Hidaka et al., 2009; Hidaka et al., 2011,

Teramoto et al., 2010). The auditory stimuli in these previous crossmodal studies provided explicit direction of motion information. Conversely, in our study, auditory stimuli only provided *timing* information and thus the impact on directional discrimination was necessarily a consequence of a more primary impact on temporal processing within the visual system. Moreover, unlike our design, the stimuli used in these previous experiments did not rule out a contribution from a high-level position tracking mechanisms (see below) and hence the observed effects may have been mediated within higher-order cortical areas.

Low-level motion energy versus high-level position tracking

Two distinct motion systems have been proposed to underlie visual motion processing.¹ The first one is a low-level pre-attentive “motion-energy” system (Adelson & Bergen, 1985; Anstis, 1980; Braddick, 1980) and the other is a high-level motion system that requires attentional tracking of the position of salient features (Cavanagh, 1992; Lu & Sperling, 1995). Low-level motion is believed to be computed within lower visual areas, such as V1 and/or MT, whereas attentional tracking is believed to be mediated by higher-order cortical areas (Claeys, Lindsey, De Schutter, & Orban, 2003; Ho & Giaschi, 2009). We designed our experiments to remove the contributions from the high-level motion system. In [Experiment 1](#), we used random-dot stimuli in which the individual movie frames did not offer boundaries that could be tracked. We found that the visual ISI that yielded threshold performance depended on the ISI between accompanying brief sounds: smaller auditory ISIs yielded better discrimination (i.e., larger visual ISI-thresholds) than larger auditory ISIs. Those findings are consistent with a temporal-ventriloquism account in which auditory timing alters the temporal processing of the visual signals that feed into the motion detection mechanism.

In [Experiment 1](#), subjects could have conceivably identified patterns of dots and tracked them. While we think this situation is unlikely, we designed [Experiment 2](#) to definitively rule out feature tracking and hence to isolate the low-level motion system. We took advantage of the reverse-phi illusion in which the perceived direction of motion is in the direction opposite to that of trackable features. This illusion is instead consistent with motion-energy models (Adelson & Bergen, 1985; Chubb & Sperling, 1989). We found that the strength of this illusion depended on the ISI between accompanying sounds: smaller auditory ISIs, like smaller visual ISIs, yielded a greater incidence of reverse-phi reports.

Conclusions

Taken together, our findings demonstrate that auditory timing alters the temporal processing of the visual signals that feed into low-level motion detectors. The neuronal mechanisms underlying this temporal ventriloquism effect are unknown. It has been proposed that “cross-modal stochastic resonance,” based on stochastic resonance (Moss, Ward, & Sannita, 2004), underlie some types of multisensory integration (e.g., Lugo, Doti, & Faubert, 2008, 2012; Lugo, Doti, Wittich, & Faubert, 2008). Another mechanism with some experimental support is that of “cross-modal phase resetting” whereby detection of a sensory target is modulated by phase reset of ongoing neural oscillations (e.g., Kayser, Petkov, & Logothetis, 2008; Lakatos, Chen, O’Connell, Mills, & Schroeder, 2007; Thorne, De Vos, Viola, & Debene, 2011). Determining whether these mechanisms underlie the temporal effects documented here likely awaits neurophysiological experimentation.

Acknowledgments

We thank T. T. Tian for superb technical assistance. We also thank T. D. Albright and A. Holcombe for discussions on this work and comments on the manuscript. This research was supported by NEI grant 521852.

Commercial relationships: none.

Corresponding author: Hulusi Kafaligonul.

Email: hulusi.kafaligonul@gmail.com.

Address: Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA.

Footnote

¹ Lu and Sperling (1995) proposed an additional system referred to as the *second-order motion system*. This system extracts the motion of features that are not defined by differences in luminance.

References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of Optical Society of America A*, 2, 284–299.
- Alais, D., & Burr, D. (2004a). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research*, 19, 185–194.
- Alais, D., & Burr, D. (2004b). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257–262.
- Alink, A., Euler, F., Galeano, E., Krugliak, A., Singer, W., & Kohler, A. (2012). Auditory motion capturing ambiguous visual motion. *Frontiers in Psychology*, 2, 391.
- Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, 10, 1411–1430.
- Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of Royal Society London B: Biological Sciences*, 290, 153–168.
- Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin and Review*, 5, 482–489.
- Bours, R. J., Kroes, M. C., & Lankheet, M. J. (2009). Sensitivity for reverse-phi motion. *Vision Research*, 49, 1–9.
- Braddick, O. J. (1980). Low-level and high level processes in apparent motion. *Philosophical Transactions of Royal Society London B: Biological Sciences*, 290, 137–151.
- Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257, 1563–1565.
- Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of National Academy of Sciences USA*, 86, 2985–2989.
- Claeys, K. G., Lindsey, D. T., De Schutter, E., & Orban, G. A. (2003). A higher order motion region in human inferior parietal lobule: Evidence from fMRI. *Neuron*, 40, 451–452.
- Emerson, R. C., Citron, M. C., Vaughn, W. J., & Klein, S. A. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology*, 58, 33–65.
- Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception and Psychophysics*, 63, 719–725.
- Freeman, E., & Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Current Biology*, 18, 1262–1266.
- Getzmann, S. (2007). The effect of brief auditory stimuli on visual apparent motion. *Perception*, 36, 1089–1103.
- Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., et al. (2009). Alternation

- of sound location induces visual motion perception of a static object. *PLoS ONE*, 4(12), e8188.
- Hidaka, S., Teramoto, W., Sugita, Y., Manaka, Y., Sakamoto, S., & Suzuki, Y. (2011). Auditory motion information drives visual motion perception. *PLoS ONE*, 6(3), e17499.
- Ho, C. S., & Giaschi, D. E. (2009). Low- and high-level first-order random-dot kinematograms: Evidence from fMRI. *Vision Research*, 49, 1814–1824.
- Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. London: Wiley.
- Kafaligonul, H., & Stoner, G. R. (2010). Auditory modulation of visual apparent motion with short spatial and temporal intervals. *Journal of Vision*, 10(12):31, 1–13, <http://www.journalofvision.org/content/10/12/31>, doi:10.1167/10.12.31. [PubMed] [Article]
- Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, 18, 1560–1574.
- Kim, R., Peters, M. A. K., & Shams, L. (2011). $0 + 1 > 1$: How adding noninformative sound improves performance on a visual task. *Psychological Science*, 23, 6–12.
- Krekelberg, B., & Albright, T. D. (2005). Motion mechanisms in macaque MT. *Journal of Neurophysiology*, 93, 2908–2921.
- Lakatos, P., Chen, C. M., O’Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53, 279–292.
- Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, 35, 2697–2722.
- Lugo, J. E., Doti, R., & Faubert, J. (2008). Ubiquitous crossmodal stochastic resonance in humans: auditory noise facilitates tactile, visual and proprioceptive sensations. *PLoS ONE*, 3(8), e2860.
- Lugo, J. E., Doti, R., & Faubert, J. (2012). Effective tactile noise facilitates visual perception. *Seeing and Perceiving*, 25, 29–44.
- Lugo, J. E., Doti, R., Wittich, W., & Faubert, J. (2008). Multisensory integration: central processing modifies peripheral systems. *Psychological Science*, 19, 989–997.
- 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., Röhrbein, F., & Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, 166, 538–547.
- Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research*, 17, 154–163.
- Moss, F., Ward, L., & Sannita, W. (2004). Stochastic resonance and sensory information processing: a tutorial and review of application. *Clinical Neurophysiology*, 115, 267–281.
- Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology*, 89, 1078–1093.
- Sanabria, D., Spence, C., & Soto-Faraco, S. (2007). Perceptual and decisional contributions to audio-visual interactions in the perception of apparent motion: A signal detection study. *Cognition*, 102, 299–310.
- Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research*, 29, 1749–1758.
- Shi, Z., Chen, L., & Müller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: the influence of time interval. *Experimental Brain Research*, 203, 723–735.
- Soto-Faraco, S., Kingstone, A., & Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, 41, 1847–1862.
- Staal, H. E., & Donderi, D. C. (1983). The effect of sound on visual apparent movement. *American Journal of Psychology*, 96, 95–105.
- Teramoto, W., Manaka, Y., Hidaka, S., Sugita, Y., Miyauchi, R., Sakamoto, S., et al (2010). Visual motion perception induced by sounds in vertical plane. *Neuroscience Letters*, 479, 221–225.
- Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *Journal of Neuroscience*, 31, 3853–3861.
- Wehrhahn, C. (2006). Reversed phi revisited. *Journal of Vision*, 6(10):2, 1018–1025, <http://www.journalofvision.org/content/6/10/2>, doi:10.1167/6.10.2. [PubMed] [Article]
- Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and Psychophysics*, 63, 1293–1313.
- Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics*, 63, 1314–1329.