# Temporal Ventriloquism: Sound Modulates the Flash-Lag Effect

Jean Vroomen and Beatrice de Gelder
Tilburg University

A sound presented in close temporal proximity to a visual stimulus can alter the perceived temporal dimensions of the visual stimulus (temporal ventriloquism). In this article, the authors demonstrate temporal ventriloquism in the flash-lag effect (FLE), a visual illusion in which a flash appears to lag relative to a moving object. In Experiment 1, the magnitude and the variability of the FLE were reduced, relative to a silent condition, when a noise burst was synchronized with the flash. In Experiment 2, the sound was presented before, at, or after the flash ($\pm \sim 100$ ms), and the size of the FLE varied linearly with the delay of the sound. These findings demonstrate that an isolated sound can sharpen the temporal boundaries of a flash and attract its temporal occurrence.

Time is a special dimension of the environment in that there are no specific sense organs for its perception. Yet time can be appreciated in vision, hearing, and touch. Despite the multimodal nature of time, however, it has been claimed that audition is relatively specialized for temporal processing, whereas vision is specialized for spatial processing (e.g., Kubovy, 1988; Welch & Warren, 1980). Empirical evidence for this claim comes mainly from studies comparing sensitivities between modalities and from studies showing dominance of one modality over the other when temporal or spatial information are in conflict.

Concerning visual dominance, numerous studies have indeed shown that vision has a strong impact on spatial information processing in audition. This impact is most clearly demonstrated in the *spatial ventriloquist effect*, in which the apparent location of a sound is shifted in the direction of a spatially discordant visual stimulus with which the sound is associated (for reviews, see Bertelson, 1999; Vroomen & de Gelder, in press). The spatial ventriloquist effect is a very robust phenomenon: It can be observed even when participants are explicitly trained to ignore the visual distractor (Vroomen, Bertelson, & de Gelder, 1998); when cognitive strategies of the participant can be excluded (Bertelson & Aschersleben, 1998); when the visual distractor is not attended, either endogenously (Bertelson, Vroomen, de Gelder, & Driver, 2000) or exogenously (Vroomen, Bertelson, & de Gelder, 2001); and even when the visual distractor itself remains unnoticed, as is the case when the visual distractor is presented in the affected visual field of hemineglect patients (Bertelson, Pavani, Ladavas, Vroomen, & de Gelder, 2000).

Audition affecting vision in the temporal dimension (i.e., *temporal ventriloquism*), however, has been much less investigated, and results are less consistent. The best-known example of temporal ventriloquism (also known as *auditory dominance*) is that of *flutter driving flicker* (Gebhard & Mowbray, 1959; Welch, Dut-

tonHurt, & Warren, 1986). When observers are asked to judge the rate at which a light is flickering and that light is presented together with a repeating (i.e., fluttering) sound, increasing or decreasing the flutter rate (>4 Hz) can cause the apparent flicker rate to increase or decrease in tandem. The reverse phenomenon—*flicker affecting flutter*—either has not been found (Gebhard & Mowbray, 1959) or has proven to be much smaller (Welch et al., 1986). Welch and Warren (1980) therefore proposed the *modality appropriateness hypothesis*, arguing that vision is specifically designed to process spatial information, whereas audition is designed to process temporal information.

More recently, Fendrich and Corballis (2001) reported a case of temporal ventriloquism with a task that did not rely on introspective reports of perceived flicker rate. They asked participants to judge when a flash occurred by reporting the clock position of a continuously rotating marker (rotating at 1.67 Hz). The flashes were seen earlier when they were preceded by audible clicks and later when they were followed by audible clicks, relative to a condition in which the flashes and clicks occurred simultaneously. Although the effect was somewhat smaller, Fendrich and Corballis also observed that clicks were heard earlier when they were preceded by flashes than when they were followed by flashes. Using a sensorimotor task, Repp and Penel (2002) also reported a case of temporal ventriloquism. Their participants were asked to synchronize finger taps with a sequence of audiovisual events presented at a rate of 2 Hz. Despite instructions to synchronize with the visual stimuli, the authors observed that simultaneous auditory tones controlled variability of the finger-tap asynchronies (see also Aschersleben & Bertelson, 2003, for related results). Morein-Zamir, Soto-Faraco, and Kingstone (2003) have also reported a case of temporal ventriloquism, using a visual–temporal order judgment task. They observed that sounds presented before and after two flashes improved visual–temporal order judgments, whereas sounds intervening between the two lights hindered performance, as if the sounds attracted the lights in the temporal dimension.

The previously mentioned studies on temporal ventriloquism have been taken to demonstrate that basic temporal properties of a visual stimulus, such as its onset time and duration, are attracted toward those of a sound with which the visual stimulus is associated. So far, though, all studies on temporal ventriloquism have

Jean Vroomen and Beatrice de Gelder, Department of Psychology, Tilburg University, Tilburg, the Netherlands.

Correspondence concerning this article should be addressed to Jean Vroomen, Department of Psychology, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE, Tilburg, the Netherlands. E-mail: j.vroomen@uvt.nl

used sequences of audiovisual stimuli presented at a certain rate, rather than a single sound and a single flash presented in isolation. Compared with a single sound, sound sequences possess a unique temporal feature, namely interonset time, which is at the basis of rhythm. For this reason, it is unclear whether temporal ventriloquism relies on the rhythmic properties of a sound sequence rather than the temporal relation between an isolated sound and light. At this point, it is therefore unclear whether one can observe a temporal ventriloquist effect when a sound and a flash are presented in isolation.

In the present study, we avoided this potential confound of rhythm by presenting a single sound and a single flash. To obtain a measure of the perceived timing of the flash, we used the visual *flash-lag effect* (FLE). The standard FLE is a robust visual illusion wherein a flash and a moving object that are presented in the same location are perceived to be displaced from one another. When the flash and the moving stimulus are physically aligned, observers typically report the flash as lagging behind the moving stimulus. The FLE is not confined to vision, however; a motor (Nijhawan & Kirschfeld, 2003) and an auditory FLE (Alais & Burr, 2003) have also been observed. The effect has prompted a variety of explanations, including ones based on motion extrapolation (Nijhawan, 1994), differential latency (Whitney, Murakami, & Cavanagh, 2000), attention (Baldo & Klein, 1995), postdiction (Eagleman & Sejnowski, 2000), and temporal integration (Krekelberg & Lappe, 2000; for reviews, see Krekelberg & Lappe, 2001; Nijhawan, 2002; Whitney, 2002).

In this study, we explored whether a sound that accompanied a flash would alter the visual FLE. In Experiment 1, we compared the FLE when there was no sound with a condition in which a sound was presented at the same time as the flash. Here, we expected that a synchronized sound would sharpen the temporal boundaries of the flash. In Experiment 2, we presented the sound at various delays before, at, and after the flash to explore whether a sound would alter the FLE in such a way that the sound appeared to attract the temporal occurrence of the flash.

## Experiment 1

In Experiment 1, we tested whether a sound that was presented in synchrony with a flash had an effect on the FLE. On each trial, participants judged the position of a flash that was presented at various timings relative to a moving bar. The flash was either accompanied by a simultaneously presented sound or presented without the sound. If the accompanying sound indeed captures the occurrence of the flash, one would expect the sound to serve as a temporal anchor for judgments of when the flash occurred. We therefore expected the categorization function of the flash with sound to be less variable than the one of the flash without sound.

### Method

*Participants.* A total of 10 participants, all first-year psychology students from Tilburg University, Tilburg, the Netherlands, took part in the experiment. They all reported normal or corrected-to-normal vision and normal hearing. All were unaware of the purpose of the experiment.

*Stimuli and design.* All visual stimuli were displayed on a color monitor (17-in. [43.18-cm] VGA with a frame rate of 60 Hz) controlled by an IBM-type personal computer. A black fixation cross was visible against a gray background throughout the session. At the beginning of each trial,

a black bar (1.33° × 0.28°) appeared randomly at either the left or to the right side of the screen. The center of the bar was initially placed 3.05° above and 6.74° to the left or the right of the fixation cross. Immediately after its appearance, the bar moved laterally toward the opposite side of the screen (see Figure 1), reaching it in 1.34 s. The frame-to-frame spatial offset was 0.16°, which resulted in a smooth and constant speed of 9.98°/s. A small white disk was flashed 3.05° above the fixation cross for one video frame (~16.7 ms). The diameter of the disk was 0.28°, which was equal to the thickness of the bar. The disk was flashed at various timings (−4, −3, −2, −1, 0, 1, 2, 3, and 4 refresh cycles of the video, or from −66.7 ms to 66.7 ms) relative to the time the bar crossed the center of the screen so that it appeared at various positions relative to the bar. The sound, if present, was delivered through two loudspeakers placed next to the computer screen. The sound was a 16.7-ms white-noise burst played at 72 dB (with a 3.0-ms fade-in and fade-out) that was synchronized to the flash. To measure any timing errors between the sound and the flash, we connected both the input to the speaker and a diode on the screen at the position of the flash to a digital scope. No timing errors exceeding 1 ms were detected between sound and flash.

Participants viewed the monitor screen binocularly from a distance of 60 cm, with their heads positioned by a chin cup. Throughout a trial, they fixated the black fixation cross. After viewing the stimulus sequence, they reported where the flash had appeared relative to the bar by pressing an appropriate key (an unspeeded, two-alternative forced-choice task). If they perceived the flash directly on top of the bar, they were to make an arbitrary choice. The next trial began 1 s after the response.

For each of the nine different timings of the flash and two sound conditions (sound present or absent), we repeated 20 trials randomly, using the method of constant stimuli. The whole experiment thus consisted of 360 trials (lasting ~15 min). A practice block of 36 trials was given before testing started.

### Results and Discussion

Figure 2 shows the proportion of *lag* responses as a function of the timing of the flash, separately for when the sound was present and when it was absent. When the flash was on the bar, it was perceived most of the time as lagging (i.e., the basic FLE). More
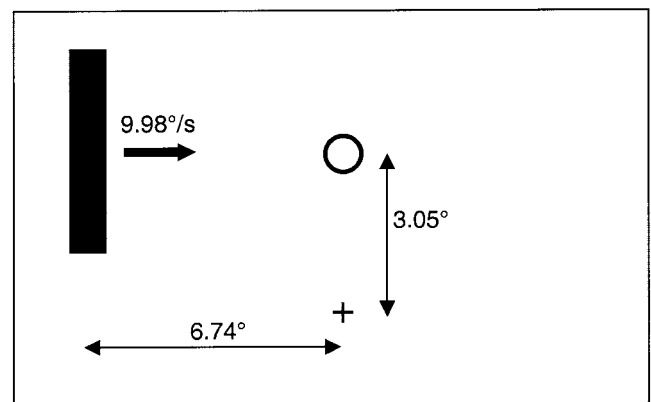


*Figure 1.* An example of the stimulus display. While participants fixated the cross, a black bar moved laterally at a constant speed against a gray background. A small white disk was flashed above the fixation cross for one frame (~16.7 ms) at various timings so that it was placed at various positions relative to the bar. The flash could be accompanied by a noise burst (not shown). The noise burst was either synchronous with the flash (Experiment 1) or presented at various delays before or after the flash (Experiment 2).
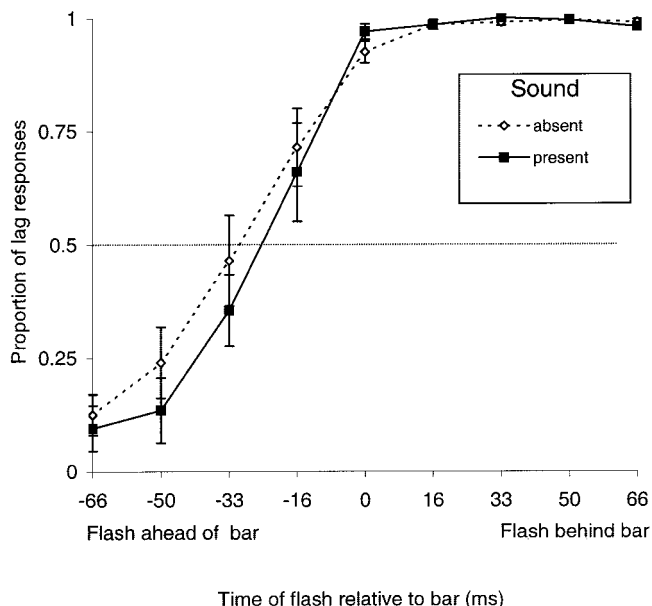
*Figure 2.* Results of Experiment 1: The proportion of lag responses as a function of the timing of the visual stimulus. The horizontal axis shows the time that the flash was presented relative to when the bar crossed the center. Time 0 depicts when flash and bar were presented at the same position. The 50% point (dotted line) on the vertical axis is estimated to measure the size of the flash-lag effect (FLE). The steepness of the categorization function at that position is a measure of the variability. The variability and the magnitude of the FLE were smaller when a sound was present. Error bars represent mean square errors.

important, the categorization function of the flash when the sound was present was sharper than the one when the sound was absent. The magnitude of the FLE was also smaller when the sound was present rather than absent (i.e., the categorization function of the flash with sound was shifted to the right).

To estimate these effects, we fitted a cumulative normal distribution for each participant (logit transformation; Finney, 1964), obtaining an estimate of the boundary (the 50% crossover point) and the slope at the boundary. The slope indicates how sharply the two categories were distinguished from one another (i.e., the variability), and the boundary indicates where the disk had to be flashed so that the flash and the bar were perceived to be at the same location (i.e., the magnitude of the FLE). The slope was 53.7° for flashes without sound and 59.3° for flashes with sound, $t(9) = 3.26$, $p < .005$ (one-tailed, because there was a clear prediction). The average point of subjective equality at which participants were equally likely to say that the flash was "ahead of" or "behind" the vertical bar was 0.33° (or 33.8 ms) for flashes without sound, and 0.28° (or 28.5 ms) for flashes with sound, $t(9) = 3.77$, $p < .004$ (two-tailed).

The presence of a sound thus made the identification function steeper by ~5.60°, and it made the flash appear earlier by ~5.3 ms. As explained above, the sharpening of the categorization function can be attributed to the more clearly marked temporal boundaries of the flash—it is as if the sound offered a temporal anchor. The decrease of the FLE indicates that the presence of the sound made the flash appear earlier. This finding is in line with

several other studies that have shown that a sound can enhance processing of a visual stimulus. For example, we have shown that an abrupt sound can increase, at a perceptual level, the brightness of a briefly flashed visual stimulus with which it is synchronized (the freezing phenomenon; Vroomen & de Gelder, 2000). Moreover, a sound can speed up processing of a visual stimulus via a spatial attention mechanism, as is shown in the line-motion illusion (Shimojo, Miyauchi, & Hikosaka, 1997). Hence, we attribute the decrease in the FLE to a processing enhancement of the flash by the sound.

## Experiment 2

In Experiment 2, we introduced various relatively small delays between the onset of the flash and the sound (±~100 ms). If the sound indeed attracts the temporal occurrence of the flash, one would expect that a sound presented before the flash should reduce the FLE, whereas a sound presented after the flash should increase the FLE.

### Method

*Participants.* A total of 16 new participants, whose general characteristics were the same as those in Experiment 1, took part in this experiment.

*Stimuli and design.* The stimuli and design were as in Experiment 1, except that on each trial a sound was presented the timing of which, relative to the flash, varied randomly between −6 and 6 refresh cycles of the video in steps of 2 (i.e., ~−100.0 ms, ~−66.0 ms, ~−33.0 ms, ~0 ms, ~33.0 ms, ~66.0 ms, and ~100.0 ms; negative values indicate that the sound was presented first). For each of the nine different timings of the flash and seven delays of the sound, we repeated 12 trials. The whole experiment thus consisted of 756 trials (lasting ~40 min, including a short break).

### Results

Figure 3 shows the proportion of lag responses as a function of the timing of the flash for each of the seven delays of the sound. For clarity, Figure 4 shows the shift of the boundary as a function of the auditory delay. As predicted, sounds presented before the flash decreased the FLE, and sounds presented after the flash increased the FLE. The boundaries and slopes for each of the seven categorization functions were estimated as in Experiment 1. There were no differences between the slopes ($Fs < 1$), which is as expected, because a sound was present in all conditions. However, in an analysis of variance on the boundaries, there was a highly significant effect of the timing of the sound, $F(6, 90) = 8.87$, $p < .0001$. The average boundaries when sounds were presented at ~−100.0 ms, ~−66.0 ms, ~−33.0 ms, ~0 ms, ~33.0 ms, ~66.0 ms, and ~100.0 ms were 0.22° (or 22.6 ms), 0.23° (or 23.4 ms), 0.26° (or 26.4 ms), 0.27° (or 27.2 ms), 0.29° (or 29.6 ms), 0.29° (or 29.3 ms), and 0.31° (or 33.7 ms), respectively. Separate $t$ tests comparing the synchronized sound condition (delay = ~0 ms) with the other conditions showed that sounds played ~−66.0 ms or more before the flash decreased the magnitude of the FLE— $t(15) = −2.75$, $p < .008$; $t(15) = −3.15$, $p < .004$; and $t(15) = −0.56$, $p = .29$, for sounds played at ~−100.0 ms, ~−66.0 ms, and ~−33.0 ms, respectively—whereas sounds played at ~100.0 ms after the flash increased the FLE—$t(15) = 1.74$, $p = .06$; $t(15) = 1.32$, $p = .10$; and $t(15) = 2.81$, $p < .007$, for sounds played at ~33.0 ms, ~66.0 ms, and ~100.0 ms, respectively.
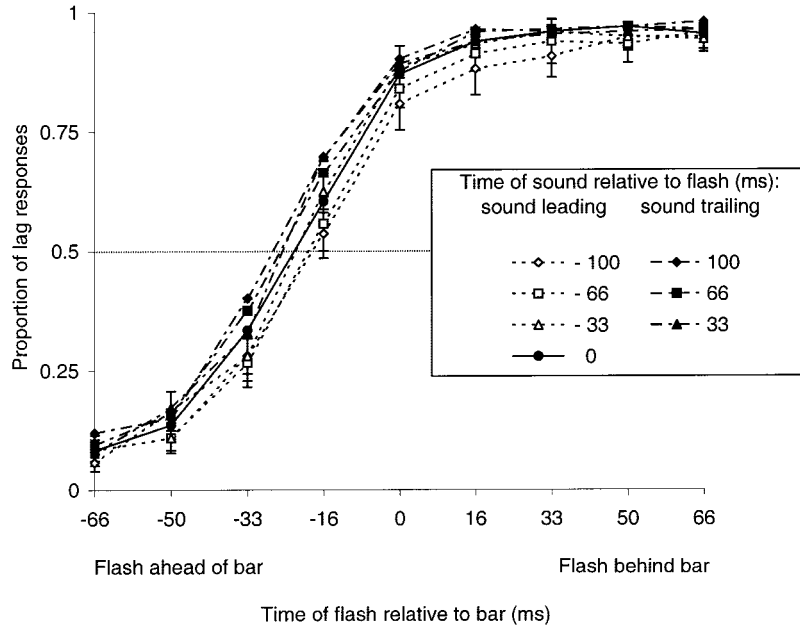
*Figure 3.* Results of Experiment 2: The proportion of lag responses as a function of the timing of the visual stimulus. The horizontal axis shows the time that the flash was presented relative to when the bar crossed the center. Time 0 depicts when flash and bar were presented at the same position. The 50% point (dotted line) on the vertical axis is estimated to measure the size of the flash-lag effect (FLE). The steepness of the categorization function at that position is a measure of the variability. The magnitude of the FLE varied systematically as a function of the delay between sound and flash (see also Figure 4). Error bars represent mean square errors.
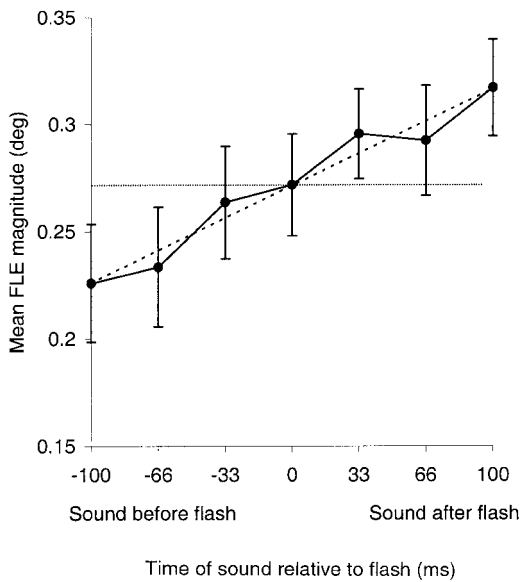
When the sum of squares was partitioned into linear, quadratic, and higher order components, there was a significant linear component, $F(1, 15) = 20.88$, $p < .0001$, but no quadratic or higher order components. The effect of the audiovisual temporal misalignment was, within the time window of $\pm \sim 100.0$ ms, approximately linear ($\sim 5.2\%$).

## General Discussion

The present study clearly shows that a single sound presented in close temporal proximity to a flash attracts the temporal dimensions of the flash. The results can be summarized as follows: When a sound was presented in synchrony with a flash, it sharpened the temporal boundaries of the flash and made the flash appear earlier ($\sim 5.0$ ms). Moreover, when the sound was presented before or after the flash within a range of $\pm \sim 100.0$ ms, the sound attracted the temporal occurrence of the flash by approximately 5.2%. These findings are important because they demonstrate that not only rhythmic sequences of sounds but also a sound presented in isolation can affect temporal dimensions of visual information processing in a task in which visual space, not time or rate per se, is the relevant dimension. Furthermore, the results have implications for the interpretation of FLE itself in that they broaden the scope of enquiry.



*Figure 4.* Results of Experiment 2: The magnitude of the flash-lag effect (FLE) as a function of the delay of the sound. The vertical axis shows the physical distance (in degrees [deg] of visual angle) where the disk had to be flashed so that the flash and bar were perceived to be at the same location. The magnitude of the FLE increased linearly when the sound was presented later relative to the flash. The dotted line represents the linear fit. Error bars represent mean square errors.

### Implications for Temporal Ventriloquism

There are several examples in the literature that show that sounds may affect the temporal dimension of visual events. The best-known example is that a change in the rate of an auditory

flutter can modulate the perception of a constant visual flicker (Gebhard & Mowbray, 1959; Welch et al., 1986). However, the usefulness of this example is limited in that, to the best of our knowledge, all studies demonstrating it have used stimulus sequences rather than a single sound (see also Aschersleben & Bertelson, 2003; Fendrich & Corballis, 2001; Repp & Penel, 2002). This creates a potential confound, because auditory sequences typically have rhythm in which interonset time is salient. Hence, these studies may have been investigating cross-modal frequency capture rather than phase capture per se. The present study avoided this problem by presenting a single sound with a single flash.

In addition, in many studies, participants have been asked to make direct comparisons between flicker and flutter. Such direct comparisons, however, may show simple response biases rather than changes in perception per se. In the present study, we avoided this problem of response bias by asking participants to judge visual spatial attributes rather than temporal ones. The present data therefore constitute a particularly clear example of cross-modal phase capture.

### Implications for Current Theories of the FLE

Although the present study was not set up as a critical test to falsify one of the various interpretations of the FLE, our findings nevertheless have implications for the interpretation of the effect. One of the more prominent proposals is that the visual system is predictive, accounting for neural delays by extrapolating the trajectory of a moving stimulus into the future (Nijhawan, 1994). On this account, motion extrapolation is implemented in a low-level visual motion correction process, possibly implemented in the retinal ganglion cells (Nijhawan, 2002). Previously, it has been questioned whether this version of the motion-extrapolation model is viable as the sole mechanism underlying the FLE. Contradicting with this notion are, among others, the findings that (a) the FLE depends on the motion velocity after the flash rather than before the flash (Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000), (b) retinal motion can be yoked by head and body movements (Schlag, Cai, Dorfman, Mohempour, & Schlag-Rey, 2000), and (c) an FLE can be observed in the auditory modality (Alais & Burr, 2003). Our results, like the previous ones, contradict the retinal-motion extrapolation as the sole account of the FLE because they imply that the FLE cannot rely entirely on a low-level motion-correction process. Rather, the fact that the FLE is altered by a sound suggests that the FLE is modulated, if not originated, in higher level brain regions where interactions between modalities are more common.

Yet another account of the FLE that does not incorporate the present results is that of Eagleman and Sejnowski (2000). They proposed that the position of moving objects is not determined instantaneously but, rather, averaged over some time interval: When a flash is presented, it resets motion integration of the moving object, which then causes the visual system to integrate position signals from the postflash positions of the moving object. On this account, visual awareness is postdictive, and the percept attributed to an event is entirely dependent on what happened in the ~80-ms time window following the flash. Eagleman and Sejnowski's most important demonstration was that the FLE was completely independent of the preflash trajectory of the moving object but, rather, was entirely dependent on the movement after the flash. In seeming contradiction of this proposal, however, the present findings show that the preflash period is of importance, because a sound presented before the flash reduced the FLE. Of course, one could probably adapt the model so that a sound can reset motion integration as well, but this is something that was not incorporated in the present version.

In conclusion, this study demonstrated that a sound can alter visual perception in the temporal domain. A sound sharpens and, when in temporal conflict with it, attracts the occurrence of a flash. Future studies will be needed to further unravel the mechanism by which temporal integration occurs. The relevant questions are, among others, whether temporal ventriloquism occurs in other modalities (e.g., vision–touch) and whether the effects depend on spatial proximity of the attractor.

### References

Alais, D., & Burr, D. (2003). The "flash-lag" effects occurs in audition and cross-modally. *Current Biology, 13,* 59–63.

Aschersleben, G., & Bertelson, P. (2003). Temporal ventriloquism: Cross-modal interaction on the time dimension: 2. Evidence from sensorimotor synchronization. *International Journal of Psychophysiology, 50,* 157–163.

Baldo, M. V. C., & Klein, S. A. (1995, December 7). Extrapolation or attention shift? *Nature, 378,* 565–566.

Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–363). Amsterdam: Elsevier.

Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of auditory location. *Psychonomic Bulletin & Review, 5,* 482–489.

Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & de Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia, 38,* 1634–1642.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics, 62,* 321–332.

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Research, 40,* 1645–1648.

Eagleman, D. M., & Sejnowski, T. J. (2000, March 17). Motion integration and postdiction in visual awareness. *Science, 287,* 2036–2038.

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

Finney, D. J. (1964). *Probit analysis.* Cambridge, England: Cambridge University Press.

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

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research, 40,* 201–215.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends in Neurosciences, 24,* 335–339.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance, 14,* 318–320.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research, 17,* 154–163.

Nijhawan, R. (1994, July 28). Motion extrapolation in catching. *Nature, 370,* 256–257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences, 6,* 387–393.

Nijhawan, R., & Kirschfeld, K. (2003). Analogous mechanisms compensate for neural delays in the sensory and the motor pathways: Evidence from motor flash-lag. *Current Biology, 13,* 749–753.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1085–1099.

Schlag, J., Cai, R. H., Dorfman, A., Mohempour, A., & Schlag-Rey, M. (2000, January 6). Extrapolating movement without retinal motion. *Nature, 403,* 38–39.

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

Vroomen, J., Bertelson, P., & de Gelder, B. (1998). A visual influence in the discrimination of auditory location. In D. Burnham, J. Robert-Ribes, & E. Vatikiotis-Bateson (Eds.), *Proceedings of the International Conference on Auditory–Visual Speech Processing (AVSP'98*; pp. 131–135). Adelaide, South Australia, Australia: Causal Productions.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics, 63,* 651–659.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1583–1590.

Vroomen, J., & de Gelder, B. (in press). Perceptual effects of cross-modal stimulation: The cases of ventriloquism and the freezing phenomenon. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes.* Cambridge, MA: MIT Press.

Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contribution of audition and vision to temporal rate perception. *Perception & Psychophysics, 39,* 294–300.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88,* 638–667.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends in Cognitive Sciences, 6,* 211–216.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40,* 137–149.