Adaptation minimizes distance-related audiovisual delays

James Heron Department of Optometry, University of Bradford, Bradford, UK

David Whitaker Department of Optometry, University of Bradford, Bradford, UK

Paul V. McGraw Visual Neuroscience Group, School of Psychology, University of Nottingham, University Park, Nottingham, UK

Kirill V. Horoshenkov School of Engineering, Design and Technology, University of Bradford, Bradford, UK

A controversial hypothesis within the domain of sensory research is that observers are able to use visual and auditory distance cues to maintain perceptual synchrony—despite the differential velocities of light and sound. Here we show that observers are categorically unable to utilize such distance cues. Nevertheless, given a period of adaptation to the naturally occurring audiovisual asynchrony associated with each viewing distance, a temporal recalibration mechanism helps to perceptually compensate for the effects of distance-induced auditory delays. These effects demonstrate a novel functionality of temporal recalibration with clear ecological benefits.

Introduction

Multisensory research has traditionally focused on how proximal sensory signals are processed by the human and the animal nervous system. It is only recently that researchers have turned their attention to how signals arising from more distal events are integrated. When audiovisual events occur proximally (e.g., within arms reach), the relatively sluggish velocity of sound (relative to that of light) has a negligible effect on the physical time taken for each stimulus to arrive at their respective receptor sites. In this situation, the relatively swift transduction of sound waves into electrical signals is thought to give audition a neural "head start" over visual signals, whose biochemical transduction process takes approximately 50 ms longer (Fain, 2003). Evidence for a resultant *perceptual* lead of audition over vision remains controversial, as evidenced by observed dissociations between reaction time and temporal order judgments (TOJs) (Jaskowski, 1996; Rutschmann & Link, 1964; Stone et al., 2001). Once audiovisual events become more distal, this theoretical neural advantage is quickly abolished as physical auditory arrival times are linearly delayed with distance (e.g., the asynchronous flash of lightning and clap of thunder). Under such conditions, physical auditory delays might be expected to produce concomitant perceptual delays (vision leads audition). Despite this, recent studies have found that observers are able to maintain the perception of synchrony by utilizing visual (Sugita & Suzuki, 2003), auditory (Alais & Carlile, 2005), or audiovisual (Engel & Dougherty, 1971; Kopinska & Harris, 2004) distance information to modify the perceived auditory arrival times. Specifically, sound is perceived as having arrived at our ears simultaneously with light at our retinae (King, 2005; Spence & Squire, 2003). Quite how the nervous system achieves this is currently unclear and the effects themselves remain highly controversial, with other studies finding no evidence for such temporal recalibration (Arnold, Johnston, & Nishida, 2005; Lewald & Guski, 2004; Stone et al., 2001).

Controversy aside, it is known that humans possess at least one other mechanism for manipulating perceived audiovisual timing. Two recent studies have demonstrated rapid recalibration of perceived timing following a relatively brief period of exposure to a fixed level of audiovisual asynchrony (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Vroomen, Keetels, de Gelder, & Bertelson, 2004). The consequence of such adaptation is to pull the point of subjective simultaneity (PSS—the physical asynchrony that produces perceptual synchrony) away from veridicality and toward the asynchrony level presented during the adaptation period.

In the current study, we employed a TOJ task to investigate how perceived audiovisual timing varies as a function of observer–source distance (OSD). In our first experiment, we find that perceived audiovisual timing is indeed predicted by the differential velocities of light and sound in air and that observers show no evidence of active maintenance of perceived synchrony across OSD. However, in our second experiment, we then progress to combine the effects of distance with that of adaptation by allowing observers to adapt to the "naturally occurring" audiovisual asynchrony (vision leads sound) arising from changes in OSD. Following adaptation to this temporal audiovisual lag, observers shift their adapted PSS values dramatically closer to physical synchrony. These effects demonstrate a novel mechanism used by the nervous system in minimizing the consequences of distance-related delays.

Experiment 1—Methods

Stimuli

The visual stimulus consisted of a white circular disc (140 cdm^2, 98.5% Weber contrast) subtending 0.17° at a viewing distance of 50 cm. The visual stimulus was generated using the macro capabilities of the public domain software NIH Image 1.61 (developed at the US National Institutes of Health and available for free download from http://rsb.info.nih.gov/nih-image/download.html). The disc was presented on a gamma corrected 2100 ForMac color monitor (frame rate of 75 Hz) for a single frame (13 ms). The auditory stimulus consisted of a 13 ms square wave windowed white noise "click" generated using Amadeus II v3.8 (available for free download from http://www.hitsquad.com/smm/programs/Amadeus/), stored as an apple alert sound and delivered binaurally at equal intensity (75 dB(A) peak SPL—measured using a CEL 383 integrating impulse sound level meter) either via Sennheiser HD650 linear headphones ("headphone condition" in Experiment 1) or via a small loudspeaker (Sony SRS Z30) placed adjacent to the computer monitor (all other conditions).

Procedures

For Experiments 1 and 2, the experimental setup was situated at one end of a long, narrow corridor under ambient lighting conditions with abundant auditory and visual distance cues. The host computer (PowerMac G4) called the auditory stimulus either synchronously with the display of the visual stimulus, or 120, 80, and 40 ms prior to (−ve) or following (+ve) visual stimulus presentation. Visual stimulus onset was synchronized with monitor

frame rate, allowing a whole number of monitor frames to elapse before/after delivery of the auditory stimulus (9, 6, or 3 monitor frames corresponding to 120, 80, or 40 ms asynchronous conditions, respectively). The temporal accuracy of the delivery of both stimuli—relative to one another—was verified via simultaneous storage of both signals on a dual storage oscilloscope. Each of the seven levels of asynchrony was randomly interleaved within a method of constant stimuli and presented 30 times.

Three observers (authors J.H. and D.W. plus naive subject F.K.) made binary forced choice decisions as to "which stimulus was presented first, sound or vision?" Observers responded via the keyboard while maintaining fixation at the center of the monitor screen at all times. Each of the OSDs—1, 5, 10, 20, 30, and 40 m and the headphone condition—were blocked and randomized. The resulting psychometric functions were fitted with a logistic function of the form

$$y = \frac{100}{1 + e^{\frac{(x-\mu)}{\theta}}} \tag{1}$$

where μ is the level of asynchrony corresponding to the 50% level on the psychometric function (the PSS) and θ provides an estimate of asynchrony detection threshold (half the offset between the 27% and the 73% response levels on the psychometric function approximately).

Experiment 1—Results

Figure 1 shows PSS values for each distance averaged across observers (solid green line—diamonds). Also shown in Figure 1 is the average PSS value from the headphone condition (unfilled square). The black dashed line projecting horizontally from this point represents the PSS values expected if observers were able to actively maintain the perception of synchrony across different OSDs.

It is notable that our headphone PSS value (Figure 1—unfilled square) is slightly negative (−8 ms on average), indicating that perceptual synchrony is attained with a small physical lead of sound over vision. This finding is in agreement with several other studies (Arnold et al., 2005; Engel & Dougherty, 1971; Fujisaki et al., 2004; Neumann, Koch, Niepel, & Tappe, 1992; Smith, 1933; Vroomen et al., 2004) yet is perhaps surprising given the aforementioned neural "head start" of audition over vision (see Introduction). Although differential visual and auditory reaction times often match neural latencies (e.g., Bell, Meredith, Van Opstal, & Munoz, 2006), when compared to PSS values (as measured via TOJs) a marked dissociation is frequently observed both within (Tappe, Niepel, & Neumann, 1994) and across (Jaskowski, 1996; Pieron, 1952; Rutschmann &
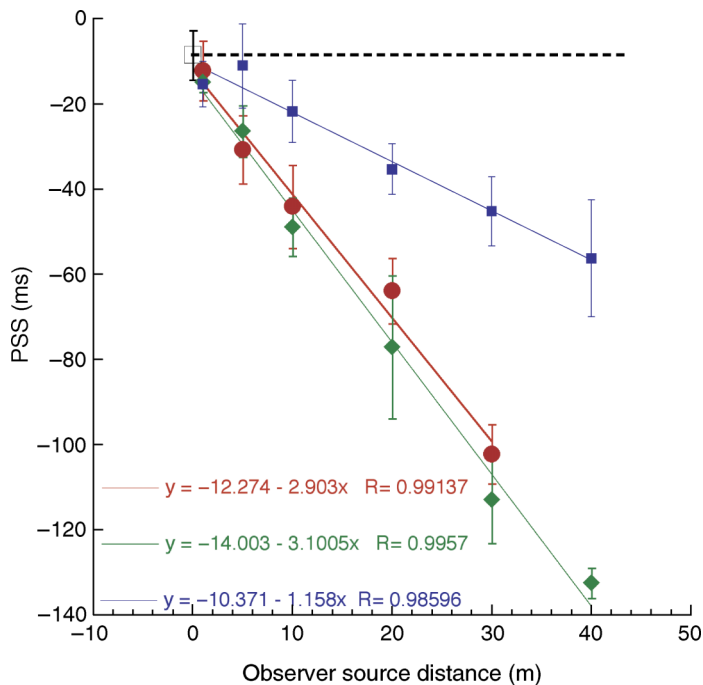
Figure 1. Point of subjective simultaneity (PSS) values (averaged across observers) gathered via temporal order judgments (TOJs) as a function of observer–source distance (OSD). Negative values indicate that an auditory temporal lead is necessary for audiovisual perceptual synchrony. The unfilled square data point represents the averaged PSS gathered with headphones (effectively an OSD of zero). The horizontal short-dashed line represents PSS values predicted if observers are able to fully maintain the perception of synchrony across distance. The unadapted data from Experiment 1 form the solid green sloping line (diamonds). The solid red line (circles) forms the results of a control experiment performed in a large reverberant environment (see text for details). The blue line (squares) represents adapted PSS values from Experiment 2. Error bars represent one standard deviation (variance between observers) either side of the parameter value.

Link, 1964; Stone et al., 2001) the modalities. The reasons behind this dissociation remain unclear.

Returning to Figure 1, as OSD increases, observer's PSS values become increasingly negative (Figure 1—green line, diamonds). This reflects the fact that at larger OSD values, observers require the auditory "click" to be delivered prior to the visual "flash" in order to maintain perceived audiovisual synchrony. Having measured the mean corridor temperature to be 18°C, the calculated gradient predicted by the differential speed of light and sound in air is −2.92 ms/m. Our data are fitted by a gradient of −3.10 ms/m. Although observers had robust visual distance cues (depth, size, ambient lighting) at their disposal, it could be argued that cues to absolute auditory distance were more ambiguous. We therefore measured temporal impulse response functions for our auditory stimulus at each of the OSDs investigated. Figure 2 shows

the two extreme conditions (1 and 40 m). Observers are known to use the ratio of direct-to-reverberant energy when making absolute auditory distance judgments (Zahorik, Brungart, & Bronkhorst, 2005). At 40 m (in red), the direct and reverberant portions of the stimulus are clearly discriminable from one another because the higher order reflections are partly absorbed and diffused along the corridor by the acoustically treated ceiling, carpeted floor, and door coves. However, at 1 m (in blue) the direct portion is, to some extent, merged with early reverberations from the nearby walls, floor, and ceiling, which is typical for near-source propagation in a narrow tunnel. Might this difference in the two response functions underlie the lack of perceptual compensation in our data (Figure 1—solid green line, diamonds)?

To test this possibility, we repeated Experiment 1 in a large reverberant environment (a high auditorium with a volume of approximately 13000 m^3). All stimuli and procedures were identical to those described above with the exception that the maximum possible OSD was now 30 m. Figure 3 shows the auditory temporal impulse response functions for the 1 and the 30 m conditions measured in the auditorium. The contrast with Figure 2 is clear—at both OSDs there is a rapid cessation of the direct portion giving way to a common reverberant tail. Thus, compared to the corridor environment, observers should have more robust auditory distance information available to them in the form of grossly different direct-to-reverberant energy ratios. It also appears that the
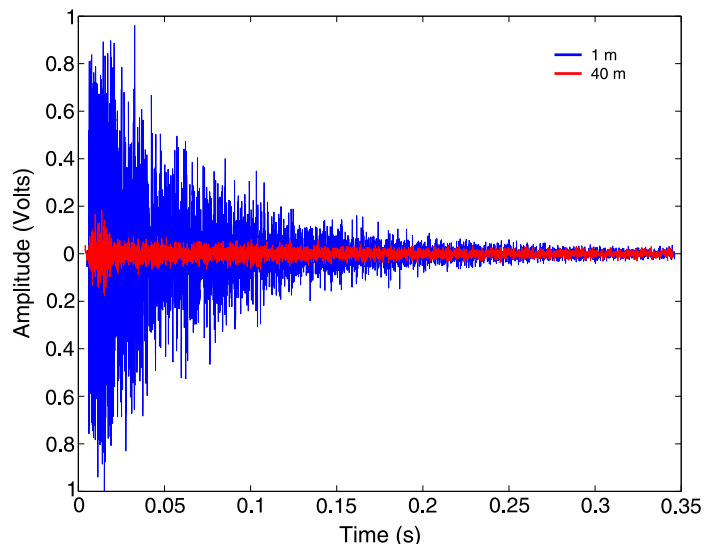

Figure 2. A-weighted temporal impulse response functions recorded (Bruel & Kjaer type 2250 sound level meter) in a long, narrow corridor at 1 and 40 m from the source (13-ms white noise "click" delivered by a Sony SRS730 loudspeaker). Beyond ~0.25 s, the reverberant tails are similar for both distances whereas the incident portion is substantially attenuated for the 40-m function.
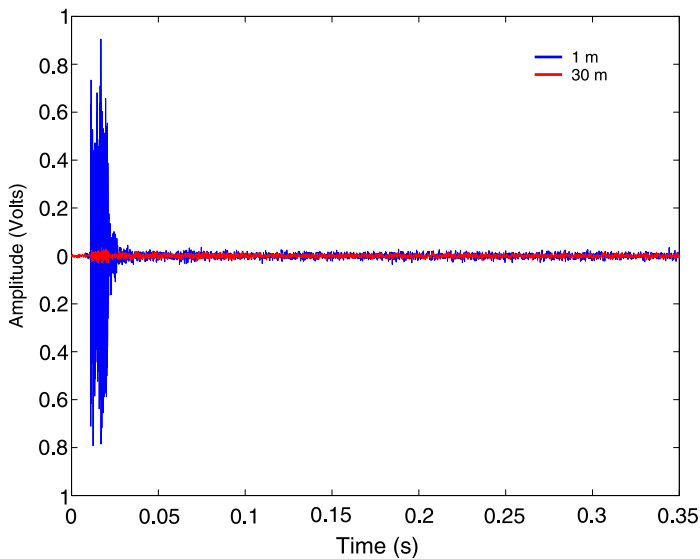
Figure 3. A-weighted temporal impulse response functions recorded (Bruel & Kjaer type 2250 sound level meter) in a 13000-m^3 auditorium at 1 and 30 m from the source (13-ms white noise "click" delivered by a Sony SRS730 loudspeaker). Comparison with Figure 2 reveals more distinct direct and reverberant portions of the stimulus.

direct portion of the 40 m corridor condition (Figure 2—in red) is greater than its 30 m counterpart in the auditorium (Figure 3—in red). This is in keeping with the slower falloff in intensity in narrow reflective environments (e.g., our corridor; Zahorik, 2002).

PSS values from TOJs in the auditorium are shown in Figure 1 (red line, circles). To maintain perceptual synchrony, observers require a progressively larger auditory lead that is approximately equivalent to that required in the corridor environment (Figure 1—green line, diamonds). This is highlighted by the similarity in the gradients of the two straight lines: −3.1 versus −2.9 ms/m. Thus, the absence of perceptual synchrony in Experiment 1 cannot be attributable to a paucity of auditory distance information.

Experiment 2 proceeds to examine an alternative method by which perceived synchrony might be attained at significant OSDs. The results of Experiment 1 show that, in both environments, the asynchrony arising at each OSD is clearly available to observers. Repeated exposure to such asynchrony is known to bring about adaptive shifts in audiovisual timing (Fujisaki et al., 2004; Vroomen et al., 2004). If a series of flash-click stimulus pairs are presented each with a physical offset of zero, they will arrive asynchronously (light leads sound) at significant OSDs. Thus, provided observers retain a fixed OSD, such a stimulus train will provide a fixed adaptation lag (Figure 4). Experiment 2 investigates the effects of adapting to this type of lag.

Experiment 2—Methods

Stimuli

Auditory and visual stimuli were identical to those used in Experiment 1.

Procedures

Experiment 2 involved identical procedures to Experiment 1 except for the introduction of initial adaptation and top-up phases. The adaptation phase involved 100 physically simultaneous click-flash pairs (approximately 1 min in total) separated by an interstimulus interval (ISI) that varied in a pseudorandom fashion between 500 and
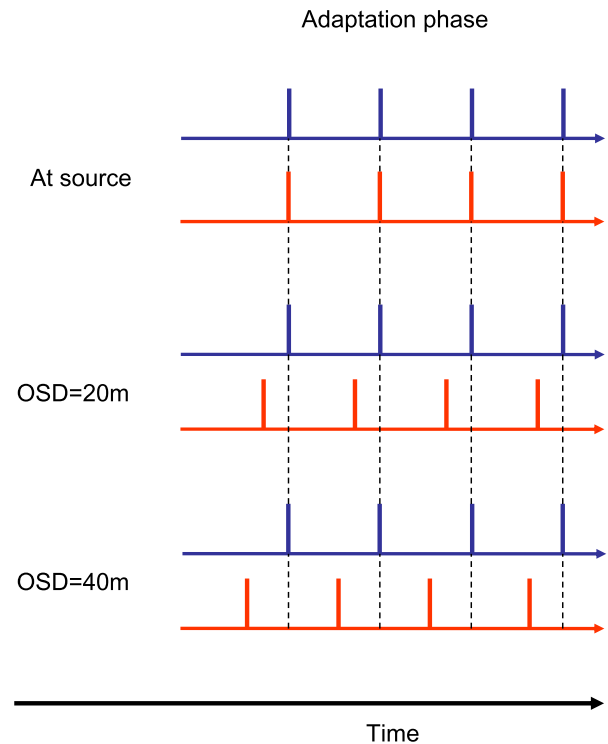

Figure 4. A schematic of the adaptation paradigm employed in Experiment 2. Both auditory (red vertical lines) and visual (blue vertical lines) stimuli leave the source (loudspeaker and computer monitor, respectively) simultaneously. Due to the differential velocities of light and sound in air, auditory stimuli become temporally delayed relative to visual stimuli as OSD increases. These "distance induced" asynchronies provide a fixed adaptation lag to which observers are repeatedly exposed during the adaptation phase. For simplicity, the interstimulus interval (ISI) between click-flash pairs is shown as constant whereas in reality it was pseudorandom. Following the adaptation phase, observers were presented with a test stimulus consisting of a click-flash pair of variable asynchrony (not shown).

1000 ms. This was followed by a 2 s pause signaling to observers that the adaptation phase had finished and that the top-up phase was imminent. The top-up phase contained four further physically simultaneous click-flash pairings (again with a pseudorandom ISI) plus a fifth pair with one of seven asynchronies separated by 40 ms steps as in Experiment 1. This final fifth stimulus pair constituted the "test" phase to which observers (authors J.H., D.W., and naive observer F.K.) made their postadaptation TOJ. This process was repeated 30 times for each of the seven test asynchronies.

The resultant data were analyzed (in the same fashion as Experiment 1) to give postadaptation PSS and threshold values. This procedure was carried out at a number of different OSDs (in the corridor environment), which were blocked and randomized. At larger OSDs, the physically simultaneous adapting pairs became increasingly perceptually asynchronous (Figure 4).

Experiment 2—Results

The resultant PSS values (extracted from the psychometric functions and averaged across observers—one for each OSD value) are shown in Figure 1 (blue line, squares). Comparison with the unadapted data (green line,
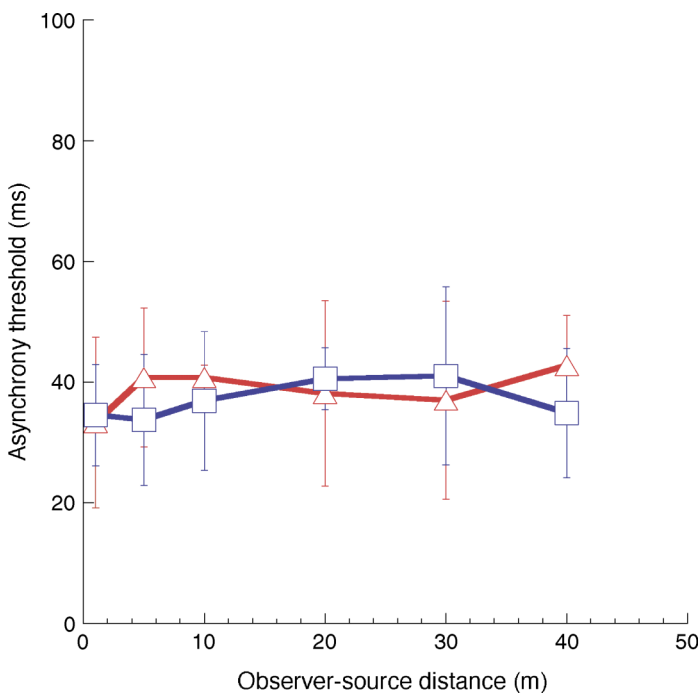

Figure 5. Threshold values (averaged across observers) for adapted (squares) and unadapted (triangles) temporal order judgments (TOJs) across distance. Error bars represent one standard deviation (variance between observers) either side of the parameter value.

diamonds) reveals a very different picture. Postadaptation TOJs now require a much smaller auditory temporal lead for perceptual synchrony. This is reflected in the much flatter gradient of the linear regression line fitted to the data (Figure 1—blue line). The net effect of such shifts in PSS is to bring perceived timing much closer to the physical timing between stimuli as they leave the source. A second important feature of the data is that sensitivity to audiovisual asynchrony is invariant across distance for all conditions (Figure 5). A repeated measures analysis of variance revealed no significant effect of distance upon threshold, $F(5, 35) = 0.29$, $p > .05$. In addition, there was no dependence of thresholds on adapted versus unadapted state, $F(1, 35) = 0.09$, $p > .05$. In other words, despite large changes in the midpoint of the psychometric functions, sensitivity to audiovisual asynchrony remains unchanged.

Discussion

The results of Experiments 1 and 2 provide an important insight into how the nervous system processes audiovisual time. The fact that observers are unable to utilize perceived visual and/or auditory distance cues to recalibrate perceived timing is, to some extent, unsurprising. As highlighted by others, this impressive feat would require accurate knowledge of absolute distance as well as implicit knowledge of the speed of sound in air (Arnold et al., 2005). Also, such recalibration conflicts with everyday experience: For example, when observing the "flash" and the "bang" of a distal firework, the fact that the auditory "bang" palpably follows the visual "flash" is confirmation that (a) the asynchrony is suprathreshold and (b) the perceived distance plays little role in perceived timing.

Although the outcome of Experiment 1 agrees with those of Arnold et al. (2005), Lewald and Guski (2004), and Stone et al. (2001), it contrasts sharply with other studies demonstrating perceptual maintenance of audiovisual synchrony across distance (Alais & Carlile, 2005; Engel & Dougherty, 1971; Kopinska & Harris, 2004; Sugita & Suzuki, 2003). Why then does such a dichotomy exist between the results of previous studies? Possible answers lie in the different stimuli and the methodology employed. Alais and Carlile (2005) used an adaptive staircase to track their subjects PSS. Given the dramatic, rapid adaptive PSS shifts observed in the current study and elsewhere (Fujisaki et al., 2004; Vroomen et al., 2004), it is conceivable that repeated exposure to very similar asynchrony levels may have adaptively shifted PSS itself. Considering the relatively shallow gradient of our adapted data, this could offer a partial (approximately 50% on average) explanation for their results.

In addition to procedural issues, it is reasonable to speculate that the choice of stimuli and its amplitude in the

presence of background noise may also be a factor. With the exception of their 5 m stimulus, Alais and Carlile's (2005) study used simulated auditory stimuli whose time–intensity profiles mimicked those expected to influence auditory distance perception in enclosed reverberant environments (Bronkhorst & Houtgast, 1999), but neither visual (computer monitor at a fixed viewing distance) nor auditory (static speakers) stimuli contained any actual physical distance cues. Similarly, Sugita and Suzuki's (2003) auditory stimuli were presented over headphones at all distances. Finally, both Engel and Dougherty's (1971) and Sugita and Suzuki's observers were instructed to "imagine" that visual and auditory stimuli were spatially colocalized at different distances. Strategic factors similar to this have recently been shown to produce PSS shifts comparable to those expected if observers use distance cues to maintain the perception of synchrony (Arnold et al., 2005). Although it is not possible to partition the literature along any one of these issues in isolation, it is noteworthy that one commonality between the current study and those finding no evidence of the aforementioned effects is the use of "real-world" auditory and visual distance cues (but see Kopinska & Harris, 2004). The results of Experiment 1 are in agreement with those of Arnold et al. (2005), who had an almost identical experimental set-up to the current study; with Lewald and Guski. (2004), who conducted their experiment outdoors; and with Stone et al. (2001), who used a "dimly lit room." These three studies have produced very similar results to those found here (Experiment 1) despite their different audiovisual (yet "real-world") environments.

Returning to the example of the firework, the results of Experiments 1 and 2 suggest that the asynchronous flash and bang may only form the initial percept (Experiment 1). With repeated exposure to such events (e.g., by the end of a fireworks display), it seems that observers adaptively shift their PSS toward the naturally occurring distance-related asynchrony (Experiment 2). That such adaptation is incomplete (approximately 50% on average) suggests a similar a mechanism to that described by Fujisaki et al. (2004). In our study and that of Fujisaki et al., the effect of adaptation appears to be a PSS shift that forms a fixed proportion of the adaptation asynchrony. This is confirmed by the linear nature of the postadaptation data set. This is perhaps surprising considering that the magnitude of many audiovisual interactions is maximal when the temporal disparity is relatively small (Heron, Whitaker, & McGraw, 2004; Roach, Heron, & McGraw, 2006; Sekuler, Sekuler, & Lau, 1997). Our largest OSD (40 m) corresponds to an auditory temporal lag of approximately 118 ms—well within the range investigated by Fujisaki et al. who show that adaptation breaks down at asynchronies beyond approximately 250 ms (Fujisaki et al., 2004). The implication for the adapted data (Figure 1—blue line, squares) is that the linearity of this effect would dissipate at very large distances where the nervous system is likely to treat such large asynchronies as arising from unrelated

events (Heron et al., 2004; McDonald, Teder-Salejarvi, & Ward, 2001; Meredith, Nemitz, & Stein, 1987; Roach et al., 2006).

Asynchrony adaptation of the type observed here and in other studies (Fujisaki et al., 2004; Vroomen et al., 2004) no doubt has its roots in Helson's (1964) adaptation-level theory. Helson proposed that human sensations are judged relative to an ensemble of previous experience. This experience forms a sensory baseline or "adaptation level" that is constantly updated by new experiences. The product of the adaptation effects in the current study is to make asynchronous audiovisual pairs perceptually more synchronous following repeated exposure (a process of assimilation). Whereas Restle (1971) has argued that these types of effects are largely cognitive in nature, it is difficult to explain our results in terms of a simple response bias. Clearly, a cognitive preference for "vision first" or "sound first" could shift the baseline PSS for TOJs. Following adaptation to "vision first," observers would need to consciously adopt a strategy of responding "audition first" to subsequent test stimuli. The strength of this preference would need to increase linearly with increasing OSD, an explanation made more implausible by the observation that (apart from very large OSDs) observers were not consciously aware of the sign of the adapting audiovisual asynchrony.

It is of interest to note that the lag adaptation effects described here and by others (Fujisaki et al., 2004; Vroomen et al., 2004) are in the opposite direction to those predicted by "Bayesian recalibration" (Knill & Pouget, 2004). A Bayesian process would predict that exposure to a consistent time lag in one direction would lead to the prior expectation of a subsequent event (for example, a synchronous pairing) possessing the *same* temporal order, not the opposite. Miyazaki, Yamamoto, Uchida, and Kitazawa (2006) find exactly such a Bayesian prediction for somatosensory TOJs. They argue that both Bayesian and lag adaptation effects might contribute to TOJs, and that the perceptual outcome depends upon which effect dominates. The dominance of "lag adaptation" may be peculiar to audiovisual perception due to our long-term exposure to audiovisual asynchrony (Miyazaki et al., 2006).

Although this hypothesis is confirmed by our postadaptation data, it fails to explain why—given repeated exposure—observers are equally adept at shifting their PSS toward auditory leads and lags (Fujisaki et al., 2004; Vroomen et al., 2004), with the former requiring sound to travel faster than light in air. Other examples of asynchronous modality pairings which are unlikely to be encountered "in the real world" include auditory–somatosensory and visual–somatosensory pairs. Further experiments examining potential temporal recalibration mechanisms pertaining to these pairs will provide insight into how and why the nervous system modulates perceived timing.

The effects observed here may well provide "real-world" ecological benefits. It is well established that audiovisual

perceptual binding is critically dependent on the temporal discrepancy between the modalities (Fendrich & Corballis, 2001; Heron et al., 2004; Roach et al., 2006; Sekuler et al., 1997). By adaptively shifting their PSS values, observers alter their perceived timing *away* from that distorted by the differential velocities of light and sound and toward veridicality—which, in turn, correctly identifies distal audiovisual signals as emanating from a common source.

From a neural perspective, the lack of any systematic change in sensitivity to audiovisual asynchrony has important implications. Bushara, Grafman, and Hallett (2001) identified the superior colliculus, insula, and prefrontal cortex as potential brain regions responsible for the detection of audiovisual asynchrony. If adaptation simply fatigued asynchrony detectors within such structures, a subsequent loss of sensitivity might be expected to accompany any shift in PSS. The absence of such effects suggests an alternative explanation where asynchrony adaptation represents a higher level mechanism that involves a more central, attentive process (Fujisaki, Koene, Arnold, Johnston, & Nishida, 2006; Fujisaki & Nishida, 2005), a concept given credence by an apparent lack of stimulus or task specificity (Fujisaki et al., 2004). Although the underlying mechanisms mediating audiovisual adaptive timing shifts remain elusive, from a functional standpoint, we demonstrate that such a mechanism presents an elegant solution to minimizing the consequences of distance-related delays in an environment where unavoidable physical inconsistencies abound.

Acknowledgments

References

Alais, D., & Carlile, S. (2005). Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 2244–2247. [PubMed] [Article]

Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Research, 45,* 1275–1284. [PubMed]

Bell, A. H., Meredith, M. A., Van Opstal, A. J., & Munoz, D. P. (2006). Stimulus intensity modifies saccadic reaction time and visual response latency in the superior colliculus. *Experimental Brain Research, 174,* 53–59. [PubMed]

Bronkhorst, A. W., & Houtgast, T. (1999). Auditory distance perception in rooms. *Nature, 397,* 517–520. [PubMed]

Bushara, K. O., Grafman, J., & Hallett, M. (2001). Neural correlates of auditory–visual stimulus onset asynchrony detection. *Journal of Neuroscience, 21,* 300–304. [PubMed] [Article]

Engel, G. R., & Dougherty, W. G. (1971). Visual–auditory distance constancy. *Nature, 234,* 308. [PubMed]

Fain, G. L. (2003). *Sensory transduction.* Sunderland, MA: Sinauer Associates.

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

Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proceedings of the Royal Society of London B: Biological Sciences, 273,* 865–874. [PubMed] [Article]

Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony–asynchrony discrimination of audio-visual signals. *Experimental Brain Research, 166,* 455–464. [PubMed]

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7,* 773–778. [PubMed]

Helson, H. (1964). *Adaptation level theory.* New York: Harper & Row.

Heron, J., Whitaker, D., & McGraw, P. V. (2004). Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research, 44,* 2875–2884. [PubMed] [Article]

Jaskowski, P. (1996). Simple reaction time and perception of temporal order: Dissociations and hypotheses. *Perceptual and Motor Skills, 82,* 707–730. [PubMed]

King, A. J. (2005). Multisensory integration: Strategies for synchronization. *Current Biology, 15,* R339–R341. [PubMed] [Article]

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27,* 712–719. [PubMed]

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception, 33,* 1049–1060. [PubMed]

Lewald, J., & Guski, R. (2004). Auditory–visual temporal integration as a function of distance: No compensation for sound–transmission time in human perception. *Neuroscience Letters, 357,* 119–122. [PubMed]

McDonald, J. J., Teder-Salejarvi, W. A., & Ward, L. M. (2001). Multisensory integration and crossmodal attention effects in the human brain. *Science, 292,* 1791. [PubMed]

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons: I. Temporal factors. *Journal of Neuroscience, 7,* 3215–3229. [PubMed] [Article]

Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nature Neuroscience, 9,* 875–877. [PubMed]

Neumann, O., Koch, R., Niepel, M., & Tappe, T. (1992). Reaction time and temporal serial judgment: Corroboration or dissociation? *Zeitschrift fur Experimentelle und Angewandte Psychologie, 39,* 621–645. [PubMed]

Pieron, H. (1952). The Sensations (p. 295). New Haven: Yale University Press.

Restle, F. (1971). Visual illusions. In N. S. Castellan & F. Restle (Eds.), *Adaptation level theory* (vol. 3, pp. 75–91). Hilsdale, NJ: Erlbaum.

Roach, N. W., Heron, J., & McGraw, P. V. (2006). Resolving multisensory conflict: A strategy for balancing the costs and benefits of audio-visual integration. *Proceedings of the Royal Society of London B: Biological Sciences, 273,* 2159–2168. [PubMed] [Article]

Rutschmann, J., & Link, R. (1964). Perception of temporal order of stimuli differing in sense mode and simple reaction time. *Perceptual and Motor Skills, 18,* 345–352. [PubMed]

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

Smith, W. F. (1933). The relative quickness of visual and auditory perception. *Journal of Experimental Psychology, 16,* 239–257.

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology, 13,* R519–R521. [PubMed] [Article]

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society of London Series B: Biological Sciences, 268,* 31–38. [PubMed] [Article]

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature, 421,* 911. [PubMed]

Tappe, T., Niepel, M., & Neumann, O. (1994). A dissociation between reaction time to sinusoidal gratings and temporal-order judgment. *Perception, 23,* 335–347. [PubMed]

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research, 22,* 32–35. [PubMed]

Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *Journal of the Acoustical Society of America, 111,* 1832–1846. [PubMed]

Zahorik, P., Brungart, D. S., & Bronkhorst, A. W. (2005). Auditory distance perception in humans: A summary of past and present research. *Acta Acustica United With Acustica, 91,* 409–420.