Visually Induced Plasticity of Auditory Spatial Perception in Macaques

Timothy M. Woods[1] and Gregg H. Recanzone[1,2,*]
[1]Center for Neuroscience and
[2]Section of Neurobiology
Physiology and Behavior
University of California at Davis
Davis, California 95616

Summary

When experiencing spatially disparate visual and auditory stimuli, a common percept is that the sound originates from the location of the visual stimulus, an illusion known as the ventriloquism effect [1]. This illusion can persist for tens of minutes, a phenomenon termed the ventriloquism aftereffect [2–5]. The underlying neuronal mechanisms of this rapidly induced plasticity remain unclear; indeed, it remains untested whether similar multimodal interactions occur in other species. We therefore tested whether macaque monkeys experience the ventriloquism aftereffect similar to the way humans do. The ability of two monkeys to determine which side of the midline a sound was presented from was tested before and after a period of 20–60 min in which the monkeys experienced either spatially identical or spatially disparate auditory and visual stimuli. In agreement with human studies, the monkeys did experience a shift in their auditory spatial perception in the direction of the spatially disparate visual stimulus, and the aftereffect did not transfer across sounds that differed in frequency by two octaves. These results show that macaque monkeys experience the ventriloquism aftereffect similar to the way humans do in all tested respects, indicating that these multimodal interactions are a basic phenomenon of the central nervous system.

Results and Discussion

Most real world events consist of multiple sensory attributes. Although each sensory system initially processes information independently, this information is ultimately combined to form a unified percept. However, under the appropriate conditions both spatial and temporal mismatches between two or more modalities can give rise to the illusion that there is no disparity. In the spatial realm, visual stimuli usually dominate the localization of both auditory stimuli [1], the perception of limb position [6], and tactile localization [7]. In the temporal realm, auditory stimuli usually dominate the perception of visual temporal rate [8–10].

These capture effects can also persist when unimodal stimuli are presented. For example, in the ventriloquism aftereffect [2–5], the perception of acoustic space can be reliably shifted in one direction following a period of exposure in which visual stimuli are presented at a

*Correspondence: ghrecanzone@ucdavis.edu

consistent spatial disparity to auditory stimuli. Similar aftereffects have also been described for the influence of visual stimuli on the perception of limb position [11] and for the influence of auditory stimuli on visual temporal rate perception [10].

The ventriloquism aftereffect can be induced using complex stimuli but also with simple stimuli, such as tones or noise and light flashes [2–5]. Because simple stimuli are effective in generating the aftereffect, and aftereffects are ubiquitous across sensory modalities, it is likely that these enduring illusions are manifest early in the sensory pathways. This suggests that nonhuman animals also experience similar aftereffects, but to our knowledge, there have been no attempts to determine if enduring illusions such as the ventriloquism aftereffect exists in species other than humans, and if so, how different stimulus parameters influence the aftereffect.

To test this possibility, we trained two rhesus monkeys to perform two behavioral tasks. The first was a sound lateralization task that was run immediately before and after a training task (see the Supplemental Data, available with his article online, for details). Each trial was initiated by the monkey fixating a small red or green light-emitting diode (LED) to within 2°. After a variable delay (500–1000 ms), a single 200 ms acoustic stimulus (tone or Gaussian noise, 65 dB SPL [sound pressure level], 5 ms on/off ramp) was presented from one location in azimuth ($+/-$ 18 or 26° in 4° increments). The fixation light was extinguished 300 ms after the auditory stimulus was completed, and two target lights appeared at $+/-$ 6° in azimuth. The monkey was rewarded if it made a saccade to the target light nearest the sound stimulus and fixated it to within 2°. Each location was tested on 12–20 randomly interleaved trials. Immediately after this lateralization task, the monkey performed the training task. Here, the monkey was required to fixate the LED and depress a lever to initiate a trial. Three to seven simultaneous auditory and visual (200 ms duration) stimuli were presented from one of the pretraining locations. The monkey was required to maintain fixation and depress the lever until the acoustic stimulus was presented 15 dB louder (from the same location), and the monkey then had to release the lever to obtain a reward. This paradigm is similar to that previously used for human subjects [4], and it ensured that the monkey was actively attending to the auditory stimulus but not to any spatial aspect of either stimulus. Within a session, the visual stimulus was presented at either −4° ("sound-right"), +4° ("sound-left"), or 0° ("zero-disparity") relative to the auditory stimulus on every trial. The location of these stimuli varied between trials to encompass the range of locations tested in the pretraining task such that each location was tested 20–50 times (20–60 min).

Figure 1 shows the results from three experiments using the sound-right (Figure 1A), sound-left (Figure 1B), and zero-disparity training (Figure 1C). Following the sound-right training, the monkey had a clear shift to the right of the psychometric function, resulting from a bias to respond "left" near the midline. Similarly, there is a

A Sound Right Training

B Sound Left Training
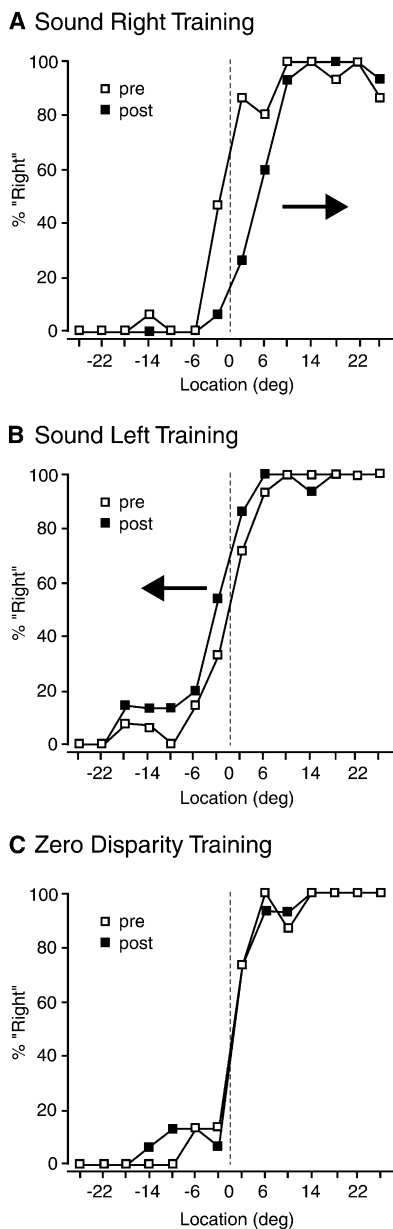

C Zero Disparity Training

Figure 1. Single-Experiment Examples of the Aftereffect

(A) Data from a single pretraining (pre; open squares) and posttraining (post; closed squares) session using a sound-right training paradigm (30 trials/location; 1 kHz tone). The midline is shown as the vertical dashed line. (A) There is a clear shift to the right of the psychometric function following sound-right training, indicating that the monkey perceived stimuli to be to the left of their actual location. (B) Data from an experiment using the sound-left training paradigm (30 trials/location; 2 kHz tone). The posttraining psychometric function is shifted to the left, indicating a rightward bias by the monkey. (C) Data from an experiment using the zero-disparity training paradigm (30 trials/location; 1 kHz tone). In this case, the pretraining and posttraining psychometric functions are nearly identical. For the posttraining condition only, the monkey was rewarded for selecting either target for locations at +/− 2°. This prevented the monkeys from biasing their choice based on reward availability. All other locations required a correct response to receive a reward.
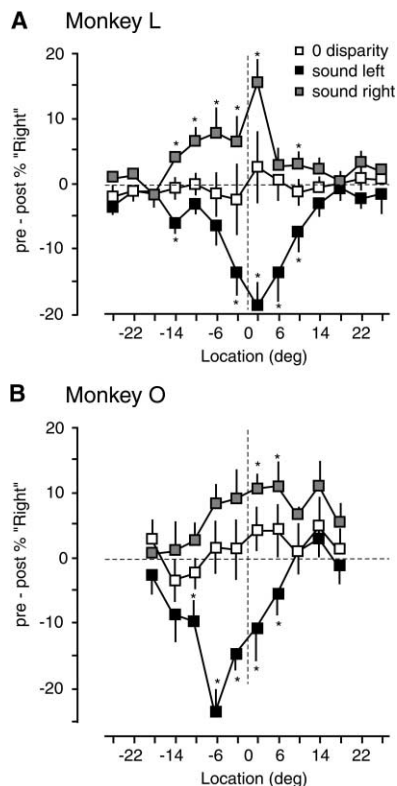
A Monkey L

B Monkey O

Figure 2. Summary of the Aftereffect across Experiments and Monkeys

Each panel shows the difference between the mean pretraining responses minus the mean posttraining responses, with the data from monkey L in panel (A) and monkey O in panel (B.) The vertical dashed line shows the midline, the horizontal dashed line shows zero difference. Shaded squares show the difference following the sound-right training paradigm (n = 14 and 11 for monkeys L and O, respectively). At points near the midline, there is an increase in this difference, indicating that the monkey had a leftward bias. Solid squares show the difference following the sound-left training paradigm (n = 14 and 10 in monkeys L and O, respectively). In this case there is a decrease of this difference near the midline, indicating a rightward bias by the monkey. Open squares show the results from the zero-disparity training paradigm (n = 15 and 10 for monkeys L and O, respectively). There is little consistent change in these differences following this training paradigm. Error bars represent the SEM. and are shown in only one direction at some locations for clarity. Points where error bars are not shown were smaller than the symbol size. Stars indicate $p < 0.05$ from posthoc analysis (two-tailed t test). Experiments in which the pretraining session showed a 50% "right" response greater than 2° different than zero indicated an initial response bias by the monkey and were not included in this analysis.

clear shift to the left of the psychometric function following sound-left training (Figure 1B), resulting from increased "right" responses. For both cases, the bias was in the direction of the visual stimulus during the training paradigm. For the zero-disparity training, there was no systematic shift in the psychometric function (Figure 1C).

The results from all sessions from both monkeys are summarized in Figure 2. In each plot, the difference between the percent "right" responses during the pre-training and posttraining sessions are plotted at each

location tested. There is a consistent upward shift following sound-right training (shaded squares), indicating fewer "right" responses (corresponding to the rightward shift in the psychometric function as in Figure 1A). Similarly, following the sound-left training period (solid squares), there is a consistent downward shift, resulting in a greater percentage of "right" responses following training. Sessions following the zero-disparity training (open squares) showed very little shift and all mean values were within a 5% difference between the pre- and posttraining conditions.

The differences between the sound-right, sound-left, and zero-disparity training conditions were statistically significant (ANOVA: $F_{(2, 61.99)}$; $p < 0.001$ and $F_{(2, 38.95)}$; $p < 0.001$ for monkeys L and O, respectively), and posthoc analysis (t test) showed that the significant differences were located near the midline (Figure 2, stars indicate $p < 0.05$). This midline effect was expected due to the nature of this lateralization task, as shifts in the perception of the stimulus location at far eccentricities would still be correctly identified as "left" or "right." However, significant effects were still noted in some cases well beyond the 4° disparity presented during the training period. These shifts in the psychometric functions are consistent with results in human observers using head- [4] or manual-pointing [5] tasks or a button-release task [4] and indicate that similar shifts in auditory spatial perception are consistent between these two species. Interestingly, the two monkeys showed the effect in different degrees, with monkey L being fairly symmetric and monkey O showing a much greater effect for sound-left training compared to sound-right training. This is likely a result of monkey O having a slight rightward bias of 0.5–1.5° in the pretraining sessions; this bias would tend to magnify the effect for sound-left training. Nonetheless, the main finding is that both monkeys showed a statistically significant effect for both sound-right and sound-left training and no effect for zero-disparity training.

This effect was not due to a change in the monkey's acuity or diligence in performing the task, which would be expected to result in either an increased or decreased slope of the psychometric function, respectively. The ratio of the pre- and posttraining slope of the psychometric function, defined as the slope of the line drawn between the two data points that bracketed the 50% "right" response, are shown in Figure 3A. This distribution is largely symmetrical, indicating that there was no consistent effect on the slope of the psychometric functions following sound-right or sound-left training. The ratios for the zero-disparity training showed a similar distribution (Figure 3B). There was no statistically significant difference between the zero-disparity and the sound-left and sound-right disparity training distributions (Figures 3A versus 3B; Wilcoxon signed rank; $p > 0.05$). This analysis indicates that the training period did not systematically alter the monkey's motivational state or its diligence in performing the task; rather, the psychometric functions were essentially the same except for a shift in the direction corresponding to the visual stimulus during the training period. In the case of the zero-disparity training, there was no statistically significant improvement (paired t test; $p > 0.05$), although the

A 4 degree disparity training
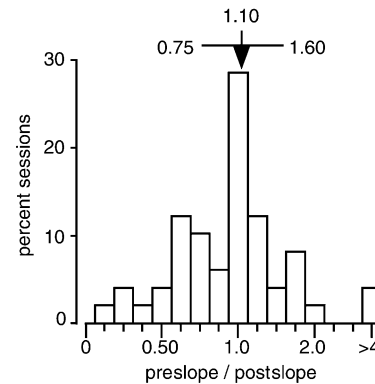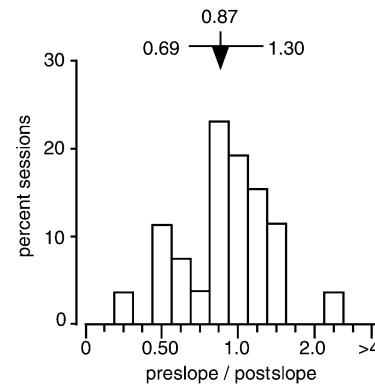
B 0 degree disparity training

Figure 3. Frequency Distribution of the Slope Ratios

(A) The ratio of the pretraining slope divided by the posttraining slope following either sound-left or sound-right training. There were no significant differences between sound-left or sound-right training or between monkeys (t tests, $p > 0.05$ in all cases), so all data were pooled (49 total experiments). The median is shown as the arrowhead, and the first and third quartiles are spanned by the horizontal line. This distribution is largely symmetrical and is centered near 1.0, indicating no consistent change in the slope of the psychometric function as a consequence of the training paradigm. (B) Frequency distribution of the same ratio for the zero-disparity training paradigm. Conventions as in panel (A) and data from both monkeys are pooled (25 total experiments). The median is slightly below 1.0 but the first and third quartiles evenly span 1.0 (no difference). The two distributions are not statistically significantly different from each other (see text).

median ratio was below 1.0, suggesting a trend for better sound-lateralization ability.

Effects of Acoustic Stimulus Spectrum
We noted that the magnitude of the aftereffect varied between sessions in both monkeys. We reasoned that this variability may be due to variability in the monkeys' ability to discriminate the stimuli, because there are clear differences in sound localization ability depending on the stimulus spectrum [12]. Each monkey was tested at a number of different tone frequencies as well as broadband noise at different intensities, which resulted in a range of the slope of the psychometric functions in the pretraining condition. We compared these slopes to
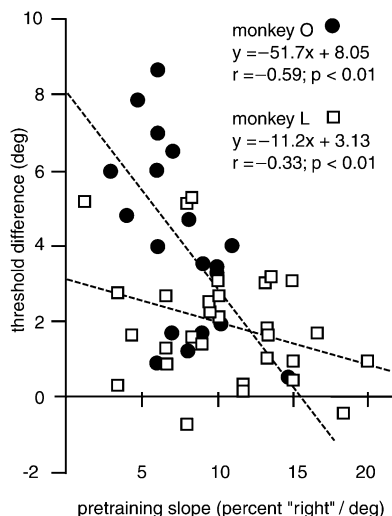
Figure 4. Aftereffect Strength as a Function of Lateralization Performance

The y axis plots the difference in threshold, defined as the location in azimuth at 50% right responses linearly interpolated from the two data points that bracketed that value, between the pretraining and posttraining session. Stimuli tested were either 0.5, 1, 2, 4, and 8 kHz at 65 dB SPL or Gaussian noise at 65 or 25 dB SPL. The *x* axis shows the slope of the psychometric function in the pretraining condition and reflects the ability of the monkey to lateralize those stimuli. There was a difference between monkeys (monkey O: solid circles; monkey L: open squares) but the correlation was statistically significant for each. Dashed lines show the regression lines (equations and r values are shown in the inset). The stimuli more difficult to lateralize (low slope values) produced the largest effects.

the difference between the pretraining and posttraining threshold, taken as the 50% "right" response rate, with thresholds from the sound-left experiments changed in sign so that positive values correspond to shifts in the expected direction. The data from the two monkeys are shown independently (Figure 4), and both sets of data were statistically significantly correlated (see inset). These results indicate that the magnitude of the aftereffect illusion was greatest for stimuli that the monkey had most difficulty in discriminating, suggesting that a greater aftereffect occurs when the auditory stimulus is more easily "captured" by the visual stimulus during the training session. This is consistent with our observations in human subjects, where the subjects must experience the ventriloquism illusion in order for the aftereffect to occur (unpublished data).

Transference of the Aftereffect across Frequency
Previous studies in human subjects indicate that there is little if any aftereffect when the stimulus used in the training session is two octaves different in frequency than the test frequency used in the pre- and posttraining tasks [4, 5]. To explore the transference of the aftereffect across large frequency differences in monkeys, we conducted pre- and posttraining sessions for both 1 kHz and 4 kHz stimuli in seven separate experiments in monkey L. An example of the data from one of these experiments is shown in Figure 5. The monkey was first tested at 4 kHz (open squares, Figure 5A) and then 1 kHz (open
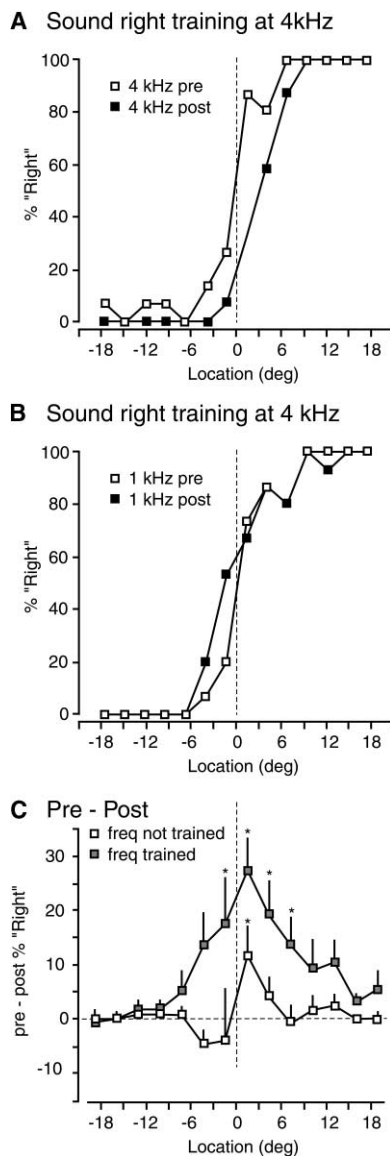
Figure 5. Minimal Transference across Frequency

In this example, the monkey was tested at both 4 kHz (A) and 1 kHz (B) before (pre; open squares) and after (post; closed squares) sound-right training using the 4 kHz stimulus. The 1 kHz tone was tested immediately before and after the training session. There was little if any effect of the training at 4 kHz in this example (B). The psychometric function did show a clear shift to the right for the 4 kHz stimulus (A), consistent with previous experiments. (C) Differences in the percent "right" responses between the pretraining and posttraining sessions following sound-right training (seven experiments total). The trained frequency was always 4 kHz (solid squares) and showed a consistent increase as was seen in other experiments (Figure 2). The untrained frequency was always 1 kHz (open squares) and showed little effect. Conventions as in Figure 2.

squares, Figure 5B) in the pretraining conditions. The sound-right training session consisted of 4 kHz tones with 35 trials at each location. Immediately following the training session, 1 kHz was tested (Figure 5B) and there was no apparent aftereffect. The 4 kHz task was then immediately performed (Figure 5A), and the psychometric function showed a clear shift to the right, as in previ-

ous experiments. These data are shown pooled with 6 additional sessions in Figure 5C, where there is a clear aftereffect for the 4 kHz stimulus (solid squares), consistent with the results from the previous experiment (Figure 2A). There was also a small but statistically significant effect for one location for the 1 kHz tone stimulus (paired t test; $p < 0.05$). Thus, these results indicate that there is minimal influence of the training stimulus on the localization ability of acoustic stimuli at least 2 octaves different in frequency. This finding is consistent with previous studies in human subjects [4, 5], suggesting that similar mechanisms are in place for both humans and macaques.

Persistence of the Aftereffect
The duration of the effect was not explicitly tested, although our results indicate that the effect endured throughout the duration of the posttraining experiments (from 20–40 min). Our best indication of this is from the transference experiments, in which the nontrained frequency was always tested before the trained frequency in the posttraining condition. Thus, the effect was still clearly evident even after the monkey had performed a similar task at a different frequency for approximately 30 min. However, we saw no evidence that the aftereffect persisted through the next training day. Comparisons between the pretraining thresholds taken the day after no-disparity training were not different from those taken the day after sound-left or sound-right training (t test; all $p > 0.05$).

To address how much training is necessary to generate the aftereffect, we compared training sessions that ranged between 20 to 50 trials for each stimulus location. Multivariant regression analysis with both the slope of the pretraining psychometric function and the number of trials for each training stimulus showed no significant correlation of the number of training trials with the size of the aftereffect in either monkey ($p > 0.05$). Secondly, one monkey was tested in seven different sessions using a 2 kHz test stimulus with a range of 30–50 trials during the training paradigm. Again, there was no significant correlation between the number of trials and the magnitude of the aftereffect ($r = -0.025$; $p > 0.05$). We conclude that there is no influence of the number of trials presented during the training paradigm within the range that we tested.

These findings in macaques indicate that this species may represent an ideal animal model to study the neuronal basis of these illusions. It has been previously suggested that capture effects are the result of the sensory modality with the greater acuity dominating that of the modality with the lower acuity [1–6, 11]. Recent experiments have strengthened this view, because under the appropriate conditions, auditory stimuli can capture the spatial location of visual stimuli [13]. One would therefore expect that either the responses of single neurons and/or neuronal ensembles representing the lower-acuity modality would be altered to a greater extent than those representing the higher acuity modality. This could occur in multimodal areas of the cerebral cortex and/or superior colliculus, which potentially encode the multiple sensory attributes of complex objects and events

[14–20]. One would therefore expect that the neuronal responses in these areas would be consistent with the sensory percept (illusion) and not the actual stimulus locations. Similarly, these effects would be expected to persist in the absence of new information that the disparity no longer exists. This would explain the duration of the effect during the posttraining period in our monkeys but a re-calibration back to normal when the monkey experiences a normal auditory/visual spatial environment in its home cage between sessions.

However, because the transference across frequency was very limited, it is also likely that sharply tuned neurons in unimodal auditory areas could be involved. One promising cortical candidate is the caudomedial area (CM) of auditory association cortex. Neurons in this area have been shown to respond to both somatosensory [21] and visual [22] inputs. This area also has sharp auditory spatial tuning [12], and although many neurons in CM are relatively broadly tuned for frequency, a substantial minority of neurons have sharp frequency tuning consistent with the limited transference of the aftereffect [23]. These factors make CM a potential, previously considered "unimodal," cortical area to reflect the ventriloquism aftereffect.

Finally, although this phenomenon is intriguing, one is forced to wonder what practical benefit such rapidly induced plasticity would provide the animal. Abrupt changes in sound localization ability, for example by an acute partial occlusion of one ear, would generate spatial mismatches between auditory and visual stimuli. If the nervous system can rapidly adapt to such changes, the perception of a unified sensory world would be maintained. It is reasonable to expect that this change would be enduring, as the realignment of acoustic and visual spatial representation would allow accurate localization of real objects when the visual input is not present, such as during occlusion.

Conclusions

This report demonstrates that the same stimuli that generate the ventriloquism aftereffect in humans [4] also generate the aftereffect in macaque monkeys. The aftereffect did not systematically change the slope of the psychometric function but rather shifted the perception of acoustic space in the expected direction. Also similar to human studies [4, 5], there was no transference of this effect for stimuli two octaves different in frequency. Thus, the macaque monkey is likely sharing the same alteration in the perception of acoustic space that occurs in humans following brief periods of exposure to auditory and visual spatial disparity. These results pave the way for future studies directed at exploring the underlying neuronal mechanisms of multimodal integration in the mammalian brain.

References

1. Howard, I.P., and Templeton, W.B. (1966). Human Spatial Orientation, (New York, NY: Wiley).
2. Radeau, M., and Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. Percept. Psychophys. *22*, 137–146.
3. Radeau, M., and Bertelson, P. (1978). Cognitive factors and adaptation to auditory-visual discordance. Percept. Psychophys. *23*, 341–343.
4. Recanzone, G.H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. Proc. Natl. Acad. Sci. USA *95*, 869–875.
5. LeWald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. Learn. Mem. *9*, 268–278.
6. Hay, J.C., Pick, H.L., Jr., and Ikeda, K. (1965). Visual capture produced by prism spectacles. Psychon. Sci. *2*, 215–216.
7. Pavini, L., Spence, C., and Driver, J. (2000). Visual capture of touch: out-of-the-body experiences with rubber gloves. Psychol. Sci. *11*, 353–359.
8. Shipley, T. (1964). Auditory flutter-driving of visual flicker. Science *145*, 1328–1330.
9. Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. Nature *408*, 788.
10. Recanzone, G.H. (2003). Auditory influences on visual temporal rate perception. J. Neurophysiol. *89*, 1078–1093.
11. Welch, R.B. (1999). Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. In: Cognitive Contributions to the Perception of Spatial and Temporal Events, G. Aschersleben, T. Bachmann, and J. Musseler, eds, (Amsterdam: Elsevier), 371–387.
12. Recanzone, G.H., Guard, D.C., Phan, M.L., and Su, T.K. (2000). Correlation between the activity of single auditory cortical neurons and sound localization behavior in the macaque monkey. J. Neurophysiol. *83*, 2723–2739.
13. Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. Curr. Biol. *14*, 257–262.
14. Bruce, C., Desimone, R., and Gross, C.G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J. Neurophysiol. *46*, 369–384.
15. Bruce, C.J., Desimone, R., and Gross, C.G. (1986). Both striate cortex and superior colliculus contribute to visual properties of neurons in superior temporal polysensory area of macaque monkey. J. Neurophysiol. *55*, 1057–1075.
16. Mazzoni, P., Bracewell, R.M., Barash, S., and Andersen, R.A. (1996). Spatially tuned auditory responses in area LIP of macaques performing delayed memory saccades to acoustic targets. J. Neurophysiol. *75*, 1233–1241.
17. Cohen, Y.E., and Andersen, R.A. (2002). A common reference frame for movement plans in the posterior parietal cortex. Nat. Rev. Neurosci. *3*, 553–562.
18. Russo, G.S., and Bruce, C.J. (1994). Frontal eye field activity preceding aurally guided saccades. J. Neurophysiol. *71*, 1250–1253.
19. Kadunce, D.C., Vaughan, J.W., Wallace, M.T., and Stein, B.E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. Exp. Brain Res. *139*, 303–310.
20. Perrault, T.J., Jr., Vaughan, J.W., Stein, B.E., and Wallace, M.T. (2003). Neuron-specific response characteristics predict the magnitude of multisensory integration. J. Neurophysiol. *90*, 4022–4026.
21. Schroeder, C.E., Lindsley, R.W., Specht, C., Marcovici, A., Smiley, J.F., and Javitt, D.C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. J. Neurophysiol. *85*, 1322–1327.
22. Schroeder, C.E., and Foxe, J.J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Brain Res. Cogn. Brain Res. *14*, 187–198.
23. Recanzone, G.H., Guard, D.C., and Phan, M.L. (2000). Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. J. Neurophysiol. *83*, 2315–2331.