

Reference Frame of the Ventriloquism Aftereffect

Norbert Kopčo,^{1,2,3} I-Fan Lin,² Barbara G. Shinn-Cunningham,² and Jennifer M. Groh¹

¹Center for Cognitive Neuroscience, Duke University, Durham, North Carolina 27708, ²Hearing Research Center, Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts 02215, and ³Department of Cybernetics and Artificial Intelligence, Technical University of Košice, 04001 Košice, Slovakia

Seeing the image of a newscaster on a television set causes us to think that the sound coming from the loudspeaker is actually coming from the screen. How images capture sounds is mysterious because the brain uses different methods for determining the locations of visual versus auditory stimuli. The retina senses the locations of visual objects with respect to the eyes, whereas differences in sound characteristics across the ears indicate the locations of sound sources referenced to the head. Here, we tested which reference frame (RF) is used when vision recalibrates perceived sound locations. Visually guided biases in sound localization were induced in seven humans and two monkeys who made eye movements to auditory or audiovisual stimuli. On audiovisual (training) trials, the visual component of the targets was displaced laterally by 5–6°. Interleaved auditory-only (probe) trials served to evaluate the effect of experience with mismatched visual stimuli on auditory localization. We found that the displaced visual stimuli induced ventriloquism aftereffect in both humans (~50% of the displacement size) and monkeys (~25%), but only for locations around the trained spatial region, showing that audiovisual recalibration can be spatially specific. We tested the reference frame in which the recalibration occurs. On probe trials, we varied eye position relative to the head to dissociate head- from eye-centered RFs. Results indicate that both humans and monkeys use a mixture of the two RFs, suggesting that the neural mechanisms involved in ventriloquism occur in brain region(s) using a hybrid RF for encoding spatial information.

Introduction

The “ventriloquism effect” involves the perception that a sound arises from the location of a visual stimulus, even when the two cues are actually in different places (Jack and Thurlow, 1973; Alais and Burr, 2004). In the “ventriloquism aftereffect,” repeated pairings of spatially mismatched visual and auditory stimuli produce a shift in perceived sound location that persists when the sound is presented alone (Canon, 1970; Recanzone, 1998; Woods and Recanzone, 2004). These effects pose a computational puzzle because the brain uses different methods for localizing visual and auditory stimuli: the retina provides a map of the visual scene with respect to the eyes, whereas differences in sound loudness and arrival time across the two ears indicate the locations of sounds with respect to the head (Brainard and Knudsen, 1995; Razavi et al., 2007). Here, we tested which of these two reference frames (RFs) is used by the brain when visual stimuli recalibrate perceived sound locations.

Persistent visually driven biases in perceived sound location were induced in seven humans and two monkeys. Analogous experimental procedures were used to assess the similarity of

audiovisual (AV) recalibration across species. Such comparisons are important for determining whether physiological studies in nonhuman primates can provide insight into multisensory spatial processing in humans. Subjects made eye movements to audiovisual or auditory-only (A-only) stimuli (Knudsen and Knudsen, 1985). On audiovisual (training) trials, the visual component of the stimuli was displaced laterally. Interleaved auditory-only (probe) trials served to evaluate how exposure to mismatched audiovisual stimuli affected sound localization.

First, we tested whether training in a subregion of audiovisual space causes local, but not global changes in localization. We used one initial eye fixation position on training trials and presented the discrepant audiovisual stimuli from a restricted spatial range (see Fig. 1A, top). Because the aftereffect was spatially specific, we could test the reference frame of the recalibration by shifting fixation on probe trials. Specifically, on interleaved auditory-only probe trials, we varied initial eye position with respect to the head (which was fixed) and presented sounds from locations spanning both the same head-centered locations and the same eye-centered locations as on the training trials (see Fig. 1A, bottom).

If visually induced spatial plasticity occurs in a brain area using a head-centered RF, then shifts in perceived sound location should occur only for sounds at the same head-centered locations (in Fig. 1B, solid blue line matches the red line). Conversely, if plasticity occurs in an eye-centered RF, then visually induced shifts should occur only for sounds at the same eye-centered locations (dotted blue line is shifted to the left of the red line). A third possibility is that the neural mechanism involves an intermediate mixture of both RFs (a “hybrid” frame). The predicted outcomes for head- and eye-centered RFs are displayed in Figure

Received June 13, 2009; revised Sept. 24, 2009; accepted Sept. 28, 2009.

This work was supported by the following grants: Monkey experiments were supported by National Eye Institute Grant R01 EY016478, National Science Foundation Grant 0415634, and National Institute of Neurological Disorders and Stroke Grant R01 NS050942 (J.M.G.). Human experiments were supported by National Institutes of Health (NIH) Grant R01 DC05778 (B.G.S.-C.). N.K. received additional support for travel on this collaboration from NIH Grant R03 TW007640 and Slovak Cultural and Educational Grant Agency (KEGA) Grant 3/7300/09. We thank Jessi Cruger for her help with the monkey data collection and Nate Greene and Abigail Underhill for their help with preliminary studies.

Correspondence should be addressed to Norbert Kopčo, Department of Cybernetics and Artificial Intelligence, Technical University of Košice, Letná 9, 04001 Košice, Slovakia. E-mail: kopco@bu.edu.

DOI:10.1523/JNEUROSCI.2783-09.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/2913809-06\$15.00/0

1B, bottom, which summarizes the potential effect as the difference between the induced bias on trials involving the training fixation and the induced bias on trials involving the nontraining fixation point (FP).

Materials and Methods

General methods. Subjects made eye movements from a visual fixation point to a broadband noise delivered from loudspeakers in darkness. On training trials (Fig. 1A, top), visual stimuli were presented simultaneously with the sounds, using light-emitting diodes (LEDs) displaced from the locations of the speakers. On randomly interleaved probe trials (Fig. 1A, bottom), only the auditory stimuli were presented (50% of all trials).

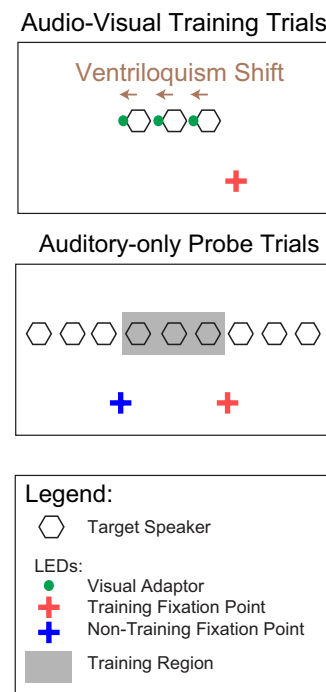
Subjects. Seven human subjects (four females; three males) and two adult male rhesus monkeys participated. The human and animal experimental protocols were approved by the institutional review committees at Boston University and Duke University, respectively.

Setup. Subjects were seated in a quiet darkened room in front of an array of speakers and LEDs (Fig. 1). To keep the head-centered RF fixed, the subjects' heads were restrained (humans, chin rest; monkeys, implanted head post). Subjects' behavior was monitored and responses were collected by an infrared eye tracker (humans) or implanted scleral eye coil (monkeys). The eye-tracking system was calibrated using visually guided saccades to selected target locations at the beginning of each session.

Stimuli. Sounds were broadband noises with 10 ms on/off ramps [humans, 100 ms, 0.2–6 kHz, 70 dB sound pressure level (SPL) (A); monkeys, ~500–1000 ms, 0.5–18 kHz, 50 dB SPL(A)] presented from speakers mounted on the horizontal plane ~1.2 m (humans) or 1.45 m (monkeys) from the center of the listener's head. Spacing between speakers was 7.5° (humans) or 6° (monkeys). The LEDs for the AV stimuli were mounted so that they were either horizontally aligned with the speakers or displaced (either to the left or to the right) by 5° (humans) or 6° (monkeys). They were turned on and off in synchrony with the corresponding speakers. Two additional LEDs 10° (humans) or 8° (monkeys) below the speaker array served as fixation locations (azimuths of $\pm 11.8^\circ$ in humans, $\pm 8^\circ$ in monkeys).

Procedures. Trials began with the onset of one of the two fixation LEDs. After subjects fixated the LED for 150 ms (humans) or 500 ms (monkeys), the fixation LED was turned off and the AV or A-only stimulus was presented. The subjects performed a saccade to the perceived location of the stimulus (humans were instructed to look to the location of the auditory component of the stimulus; monkeys were rewarded for a saccade that ended within a 16°-wide rectangular window centered on the auditory component and covering the visual component on the AV trials). Training (AV) and probe (A-only) trials were randomly interleaved at a ratio of 1:1 (in the monkeys, 12.5% of the total trials were AV-aligned and presented from the $\pm 30^\circ$ locations, just outside the range of the A-only test trials, to keep the monkeys aware of the possible stimulus range and to reinforce spatial specificity of the induced aftereffect). Trials were run in blocks with a consistent AV pairing (leftward, rightward, or no shift). For the monkeys, multiple blocks were conducted per session, with shifts in a particular direction for that session interleaved with no-shift blocks. For the humans, each session contained only one block and the order of blocks was random across the subjects. Each monkey performed a total of 128–160 blocks of ~600 trials each. Each human performed 12 sessions of ~720 trials each.

A Audio-Visual Display



B Predicted Results

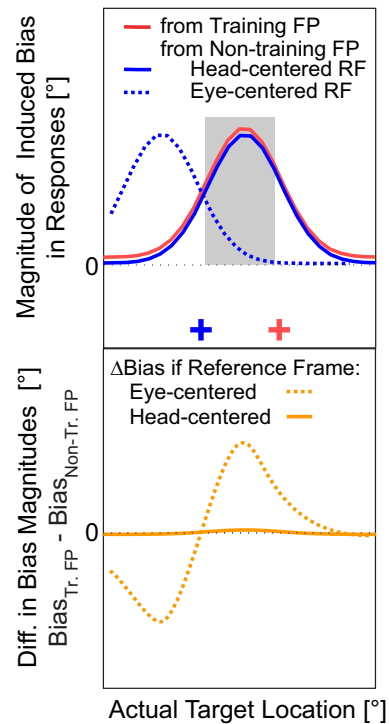


Figure 1. Experimental setup and predictions of behavior based on the two candidate RFs. **A**, Audiovisual display used to present the AV training stimuli in one experimental block. At the beginning of each AV training trial (top), the subject had to fixate on the same initial FP; then, the training stimulus was presented from one of the three center locations, keeping the direction of the induced shift the same (by consistently presenting the visual adaptor displaced to the left, to the right, or aligned with the target speaker). On the auditory-only probe trials (bottom), the same nine speaker locations and two FPs were used in all blocks in both experiments. The probe trials were randomly interleaved among the training trials and the FP and target locations varied randomly from trial to trial. **B**, Predicted results. Red line, Expected pattern of biases induced in the A-only probe responses when the eyes fixate the training FP (i.e., the same FP location as on the AV training trials). Blue lines, The expected pattern of biases in the responses from the nontraining FP if the RF of adaptation is head-centered (solid blue line) or eye-centered (dotted blue line). The orange lines show the differences between the expected bias magnitudes from the training versus the nontraining FPs in the two RFs.

Data analysis. Data from the first quarter of each block were excluded from the analysis to remove any rapid auditory localization adjustments observed at the start of each block. Basic analysis of the temporal profile of the aftereffect is provided in supplemental Figure S1 (available at www.jneurosci.org as supplemental material).

One noticeable difference between the humans and the monkeys was that the monkey responses to the peripheral targets were centrally biased (by 2–6°) (Fig. 2B, C), whereas no such bias was observed in the humans (Fig. 2A). Both the relatively large response bias and the larger response variability (supplemental Fig. S2, available at www.jneurosci.org as supplemental material) of the monkey results compared with the human results may help explain the weaker aftereffect in the monkey data. Previous reports involving auditory saccades in monkeys have suggested that monkeys sometimes make two saccades to reach an auditory target (Jay and Sparks, 1990), and this would appear in our results as a central bias for targets peripheral to the fixation positions. In our study, monkeys sometimes but not always made more than one saccade toward the target. Therefore, for the monkeys, the first saccade (or the second saccade if the delay between the first and second saccade was <300 ms) was considered a response. Since this 300 ms criterion was a conservative criterion, many second saccades were rejected, resulting in an overall central bias in the monkey responses.

Results

Saccades to auditory targets

Each experimental session started with a control block on which auditory-only (humans) or auditory-only and visual-only (mon-

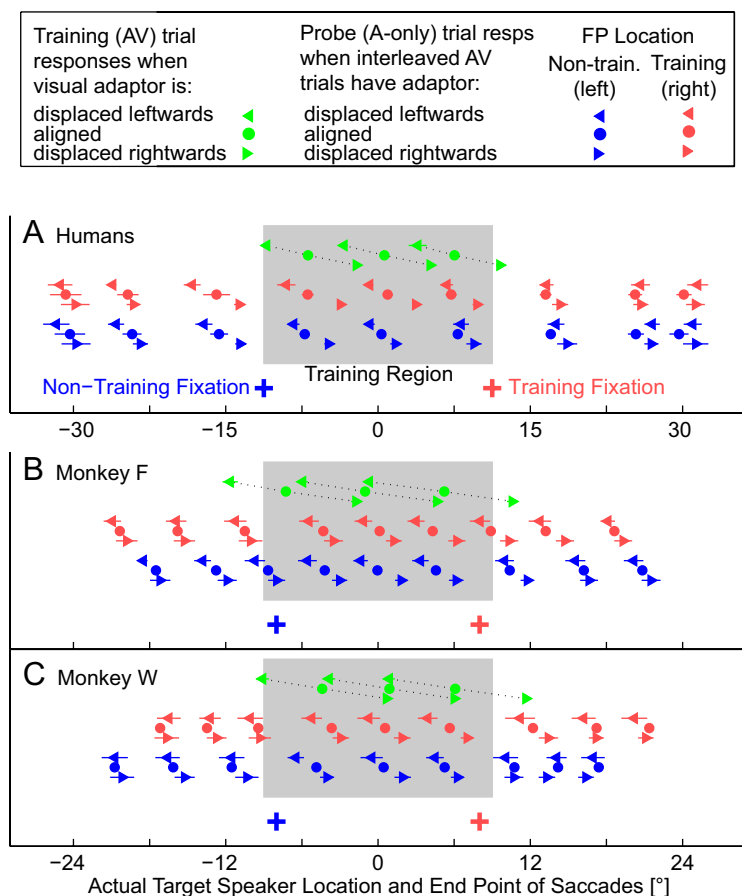


Figure 2. Raw saccade endpoints of the responses to the AV training stimuli and auditory-only probe stimuli as a function of the actual target speaker location, collapsed across time. The symbols represent responses in different audiovisual conditions (see legend), separately for the training trials (green), probe trials starting at the training fixation (red), and probe trials starting at the nontraining fixation (blue). **A**, Across-human-subject mean (± 1 SEM) of responses. **B**, Monkey F’s across-block means (± 1 SEM) of responses. **C**, Same for monkey W. The dashed lines connect symbol triplets for the same auditory target when presented with one of the three different visual adaptors (symbol triplets for A-only responses corresponding to the same target location are not explicitly connected as they are not confusable). Graphs for each measurement type are plotted in one row, vertically offset from data for other types, for visual clarity. For display purposes, a *post hoc* calibration of the eye position data based on saccades to the full range of visual targets from supplemental Figure S2 (available at www.jneurosci.org as supplemental material) was applied to the monkey results in this figure.

keys) stimuli were presented in random order from different target locations. Performance on these control trials provided baseline data on the performance of both the monkeys and the humans on the auditory localization task (supplemental Fig. S2, available at www.jneurosci.org as supplemental material). The average SD of the A-only responses was 3.0° for the humans and 4.3 and 5.0°, respectively, for monkeys F and W.

Ventriloquism effect

An almost-complete ventriloquism effect was observed in the AV training trials in both the humans and the monkeys. A connected triplet of green symbols at the top of each panel of Figure 2 represents the responses to the AV training stimuli with a single target speaker and the three different visual adaptor locations (the actual target speaker location is not explicitly shown in the figure but can be easily determined by finding, for each circle, the nearest tick mark along the *x*-axis). For clarity, the symbols are offset vertically, so that the visually induced shift appears as a tilt in the triplet of symbols for each target location. In both species and all conditions, the triangles are displaced toward the visual adaptor, with the magnitude of the displacement at least 80% of

the imposed offset of the visual adaptor relative to the auditory stimulus.

Ventriloquism aftereffect

Experience with spatially mismatched AV stimuli caused both humans and monkeys to mislocalize sounds in the direction of the previously presented visual stimuli. The red and blue symbols in Figure 2 show responses to A-only targets starting at the training and nontraining FPs, respectively. As for the AV responses, the responses to the same A-only targets form triplets in which the triangles are vertically displaced from the corresponding circles for clarity. In the training region and with eyes at the training fixation, the effect of interleaved, mismatched AV stimuli was to shift the saccade endpoints to auditory-only stimuli by up to 2.7° (or 54% of the AV displacement) in humans and 1.4° (or 23%) in monkeys. Graphically, this can be seen by comparing the horizontal positions of the red triangles with the corresponding red circles in the gray training regions in Figure 2 (also see Fig. 3, discussed below).

The first question we asked was whether the ventriloquism aftereffect generalized to locations that were not presented in training trials. We found that, in both species, the interleaved ventriloquism trials affected localization judgments adjacent to the trained region, but only modestly so. For example, in the humans, one or more of the target locations directly adjacent to the set of targets in the training region also showed a shift in saccade endpoints, but the effect diminished with increasing distance from the trained region. Graphically, the red triplets are more vertically aligned the farther they are from the gray area in Figure 2. In other words, training with mismatched auditory–visual spatial cues affected localization judgments locally, rather than globally. This spatial specificity was similar in both humans and monkeys, despite the fact that, unlike the monkeys, the humans received no “reinforcing” trials with coincident AV stimuli from locations at the edges of the test range. This consistent spatial specificity enabled us to explore the spatial frame in which auditory–visual stimuli are aligned.

These results were confirmed by performing two separate three-way repeated-measures ANOVAs (one on the human and one on the monkey A-only data), with the factors of target speaker location (nine levels), fixation point of the trials (training vs nontraining FP), and the direction of induced shift (left vs right). The results of this analysis, summarized in Table 1, show that the main effect of location was always significant, confirming that the ventriloquism aftereffect is spatially specific and does not automatically generalize to the whole audiovisual field. The location by FP interaction was also significant in both species, confirming that the reference frame of visual–auditory recalibration is not purely head-centered. However, visual inspection of the data shown in Figure 2 shows that the reference frame is not purely eye-

Table 1. Three-way repeated-measures ANOVAs of the human and monkey data

Factor	Humans (<i>n</i> = 7)			Monkeys (<i>n</i> = 2)		
	df	<i>F</i>	Significance	df	<i>F</i>	Significance
Location (speakers 1–9)	8, 48	13.4	***	8, 8	12.1	***
Fixation point on A-only trials (training vs nontraining)	1, 6	28.8	***	1, 1	14.0	
Direction of induced shift (left vs right)	1, 6	1.2		1, 1	0.01	
Location by FP	8, 48	5.47	***	8, 8	6.86	**
Location by direction	8, 48	1.9		8, 8	0.31	
FP by direction	1, 6	0.35		1, 1	0.63	
Location by FP by direction	8, 48	1.36		8, 8	2.42	

Significance levels are as follows: **p* < 0.05, ***p* < 0.01, and ****p* < 0.005.

centered either. Specifically, if ventriloquism arose in this reference frame, it would produce a much stronger aftereffect at the three left-most locations in the nontraining fixation data (blue symbols), as predicted by the blue dotted line in Figure 1*B*; however, this was not observed.

Reference frame of visual–auditory recalibration

To analyze the effect that moving the eyes from the training to the nontraining initial fixation position has on reference frame of recalibration, Figure 3, top panels, shows the magnitude of the aftereffect after collapsing the data across the two directions of the induced shift (note that no main effect or interaction involving the direction factor were significant in the ANOVA analysis) (Table 1). In both species, the peaks of the induced shift became smaller and moved leftward when the initial fixation moved from the right, training FP, to the left, nontraining FP (Fig. 3, top panels, compare red and blue traces). Thus, the observed results are inconsistent with visual–auditory recalibration occurring in solely auditory, and head-centered, brain regions. However, the leftward shift of the blue versus the red traces was never as large as the angular distance between the two fixation points, as would be expected if the representation was purely eye-centered.

To compare the current results more directly to the predictions of the two models, a difference between the shift magnitudes from the two FPs was computed (Fig. 3, black traces) and compared with predictions based on the two models (Fig. 3, orange traces). Again, the results fall between the predictions of the two models, suggesting that both the head- and eye-centered signals contribute to visual calibration of auditory space, resulting in a mixed-reference frame representation.

Since the monkey ANOVA in Table 1 had only two subjects, two additional one-way ANOVA analyses were performed separately, one for each monkey, on the difference data shown by triangles in Figure 3*B*. In this analysis, the only factor was the target location, and the data from each block were treated as a repeat. Again, these ANOVAs found a significant effect in both monkeys.

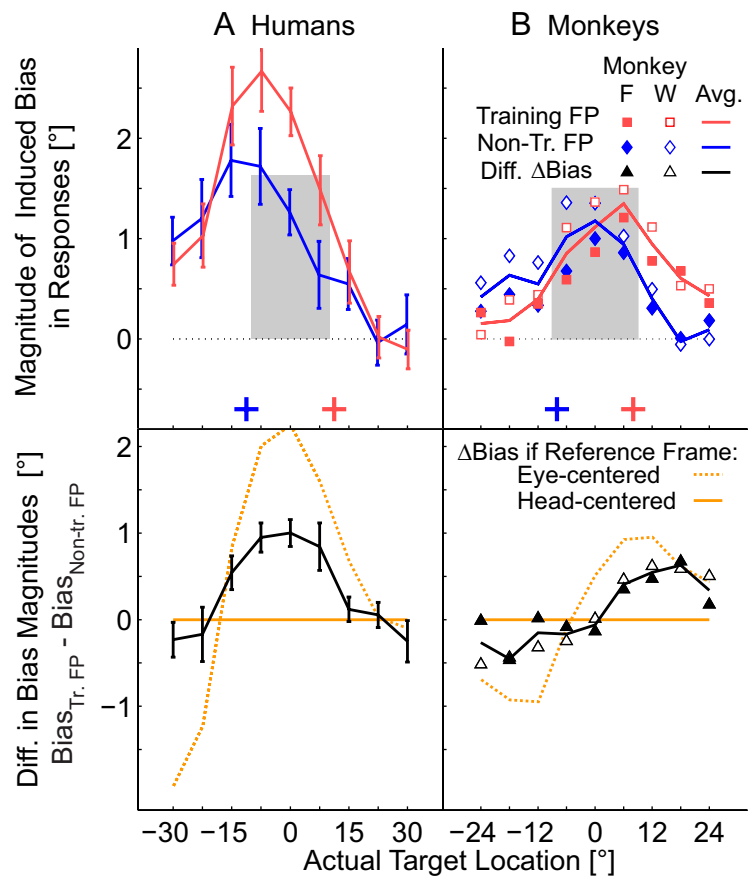


Figure 3. Magnitude of visually induced shifts in auditory saccades (top panels) and comparison of the spatial characteristics of the shifts to predictions based on eye- and head-centered RFs (bottom panels). The graphs in the top panels show the difference between the saccade endpoint locations on auditory-only probe trials interleaved with spatially displaced AV stimuli (Fig. 2, triangles) and the endpoints on probe trials interleaved with aligned AV stimuli (Fig. 2, circles). Data are collapsed across the direction of the AV displacement and across time, excluding the first quarter of each block (for more detailed temporal analysis, see supplemental Fig. S1, available at www.jneurosci.org as supplemental material). Bottom panels, The effect of the initial fixation position on the magnitude of the induced shift plotted as the difference between the shifts from the training and nontraining FPs (i.e., each black line in a bottom panel plots the difference between the red and blue lines from the corresponding top panel). The reference frame predictions (orange lines) are based on the training FP responses (red lines). **A**, Across-human-subject means (± 1 SEM). **B**, Across-monkey-subject mean and individual monkey data.

Discussion

Here, we show that when humans or monkeys repeatedly perform saccades to an auditory target presented simultaneously with a spatially displaced visual adaptor, a short-term adaptation takes place. This adaptation causes auditory location judgments to be biased toward the visual adaptor location, even on interleaved trials on which no visual adaptor is present. Specifically, saccades to auditory-only targets presented in an $\sim 20^\circ$ -wide horizontal subregion of space centered on the locations trained with

interleaved audiovisual targets were shifted by up to 1.5° (monkeys) or 2.5° (humans). The similarity in these across-species results, despite minor methodological differences (such as differences in the duration of the auditory stimuli, which were up to 1000 ms for the monkeys and only 100 ms for the humans; the presence of “reinforcing” trials at the test range edges for monkeys, but not humans; etc.), suggests that the mechanisms underlying audiovisual spatial calibration are similar in monkeys and humans. This in turn suggests that physiological studies of audiovisual spatial integration in monkeys can provide insight into human perception and behavior.

Overall, the size of the aftereffect, corresponding to 25–50% of the audiovisual adaptor displacement, is consistent with previous human studies involving either head pointing responses (Recanzone, 1998) or identifying target locations via a categorical button press response (Bertelson et al., 2006). These past studies report adaptation from 30% (Bertelson et al., 2006) to 85% of the induced visual–auditory discrepancy (Recanzone, 1998). The similarity in these results is striking, given methodological differences. For instance, the current study and these past studies each differed in the training region width and in the spatial sampling of the training region [current study, three locations in a 20° region; Bertelson et al. (2006), five locations in a 100° region; Recanzone (1998), 15 locations in a 60° region]. Comparison of these results suggests that, although it is possible to induce a local ventriloquism aftereffect (as in the current study), the magnitude of the effect is weaker than when a large region of audiovisual space is trained using a dense spatial sampling of training locations [as in Recanzone (1998)].

Saccade shifts were observed for sounds originating near to, not just within, the training region. The shift magnitudes gradually decreased with increasing separation from the training region, showing that the ventriloquism aftereffect can cause a spatially specific recalibration. For our monkeys, the presence of the “reinforcing” AV trials from the edges of the test region (on 12.5% of the trials) may have contributed to the spatial specificity we found. However, we observed similar specificity in the humans, who were not presented with reinforcing trials. This, to our knowledge, is the first report showing that the ventriloquism aftereffect is spatially specific and that it can be induced in a subregion of the audiovisual space. In contrast, a previous study in which bias was induced by compressing vision (Zwiers et al., 2003) found that shifts generalized to locations outside the adapted region without any decrease in the shift magnitude outside the trained region. It may be that spatial specificity is more likely to occur when audiovisual targets are only presented in a narrow range of space, such as we used, compared with the larger range used by Zwiers et al. (2003). Another possible explanation of this difference is that the plasticity induced in the current study was short term (on the order of minutes to hours), not days (Zwiers et al., 2003); it may even be that different neural structures are affected by short- rather than long-term training (Shinn-Cunningham, 2001).

Reference frame of recalibration and brain loci underlying ventriloquism aftereffect

In both species, the direction of eye gaze (i.e., the FP location) influenced the pattern of induced biases on the probe trials. It was not possible to align the eye-dependent bias patterns in either head- or eye-centered RF. Thus, results support the interpretation that visually guided spatial adaptation occurred in a mixed RF. This kind of mixed representation can arise in various ways (e.g., multiple structures might undergo adaptation, each using a

different frame; the adapted neural structure might receive both head-centered and eye-centered signals; etc.). Moreover, if multiple structures undergo adaptation, the character of the representation may change over time. Consistent with this idea, supplemental Figure S1 (available at www.jneurosci.org as supplemental material) shows that the aftereffect arises in a predominantly head-centered representation early during training and a more mixed representation as training progresses.

Plasticity underlying this adaptation could in principle occur in the auditory pathway, association areas, the oculomotor pathway, or some combination of the above. The number of potential sites encompassed in this list is large, but information on the multimodal properties and reference frame is only available for a limited subset of the list. Some form of hybrid representation or mixed auditory and visual signals has been reported in several areas of auditory pathway, the posterior parietal cortex, and two areas responsible for planning saccades. Specifically, signals relevant to the integration of visual and auditory space such as overt visual responses (Porter et al., 2007) as well as eye position-dependent modulations of auditory responses (Groh et al., 2001; Zwiers et al., 2004; Porter and Groh, 2006) have been found in the inferior colliculus (IC) in the primate. Visual and eye position signals have also been reported in auditory cortex (Werner-Reiss et al., 2003; Fu et al., 2004; Brosch et al., 2005; Ghazanfar et al., 2005). In both the IC and A1, eye position modulation of neural response was sufficient to cause the representation to be classified as a hybrid of head- and eye-centered information, in conflict with classical views that auditory information is generally encoded in a head-centered reference frame. However, there are no studies of the reference frame of earlier areas on the auditory pathway, so it is not known how early auditory signals might be transformed from the native head-centered frame of auditory spatial cues (interaural time and level differences) into a reference frame more appropriate for integration with visual information.

In the parietal cortex, both visual (Andersen and Mountcastle, 1983) and auditory (Stricanne et al., 1996; Schlack et al., 2004; Cohen et al., 2005) space are represented. However, that representation is also a hybrid representation that reflects a mixture of head- and eye-centered information (Stricanne et al., 1996; Duhamel et al., 1997; Mulette-Gillman et al., 2005, 2009). The superior colliculus (SC) and frontal eye fields also contain both visual and auditory signals (Meredith and Stein, 1983; Wallace et al., 1993; Wallace and Stein, 1994) and are thought to be essential for the generation of all saccades (Schiller et al., 1979). Jay and Sparks (1987a,b) examined visual and auditory sensory activity in the SC and reported that visual signals were eye-centered and auditory signals showed partially shifting receptive fields (a type of hybrid reference frame). It is not known whether the auditory saccade-related activity (as opposed to sensory activity) is also hybrid or whether it is eye-centered, as is the case for visual saccades. However, the ventriloquism aftereffect is not likely to consist purely of saccade adaptation [e.g., as has been studied by Desmurget et al. (1998) or Hopp and Fuchs (2004)] because (1) previous studies have observed the aftereffect in paradigms that did not involve saccades (Recanzone, 1998; Zwiers et al., 2003; Bertelson et al., 2006) and (2) the main effect of a purely oculomotor adaptation would likely depend on whether the induced shift results in longer or shorter saccades (Hopp and Fuchs, 2004), which we did not observe.

Overall, the current results suggest that, in both human and monkey, auditory–visual spatial recalibration occurs in a hybrid reference frame, after auditory spatial information has been partially transformed from a head-centered representation. Addi-

tional behavioral and neurophysiological studies (e.g., looking at the temporal profile of the ventriloquism aftereffect) are necessary to fully understand the mechanism and brain areas underlying the recalibration.

References

- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14:257–262.
- Andersen RA, Mountcastle VB (1983) The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *J Neurosci* 3:532–548.
- Bertelson P, Frissen I, Vroomen J, de Gelder B (2006) The aftereffects of ventriloquism: patterns of spatial generalization. *Percept Psychophys* 68:428–436.
- Brainard MS, Knudsen EI (1995) Dynamics of visually guided auditory plasticity in the optic tectum of the barn owl. *J Neurophysiol* 73:595–614.
- Brosch M, Selezneva E, Scheich H (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25:6797–6806.
- Canon LK (1970) Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J Exp Psychol* 84:141–147.
- Cohen YE, Russ BE, Gifford GW 3rd (2005) Auditory processing in the posterior parietal cortex. *Behav Cogn Neurosci Rev* 4:218–231.
- Desmurget M, Pélisson D, Urquizar C, Prablanc C, Alexander GE, Grafton ST (1998) Functional anatomy of saccadic adaptation in humans. *Nat Neurosci* 1:524–528.
- Duhamel JR, Bremmer F, BenHamed S, Graf W (1997) Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389:845–848.
- Fu KM, Shah AS, O’Connell MN, McGinnis T, Eckholdt H, Lakatos P, Smiley J, Schroeder CE (2004) Timing and laminar profile of eye-position effects on auditory responses in primate auditory cortex. *J Neurophysiol* 92:3522–3531.
- Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25:5004–5012.
- Groh JM, Trause AS, Underhill AM, Clark KR, Inati S (2001) Eye position influences auditory responses in primate inferior colliculus. *Neuron* 29:509–518.
- Hopp JJ, Fuchs AF (2004) The characteristics and neuronal substrate of saccadic eye movement plasticity. *Prog Neurobiol* 72:27–53.
- Jack CE, Thurlow WR (1973) Effects of degree of visual association and angle of displacement on the “ventriloquism” effect. *Percept Mot Skills* 37:967–979.
- Jay MF, Sparks DL (1987a) Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *J Neurophysiol* 57:22–34.
- Jay MF, Sparks DL (1987b) Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J Neurophysiol* 57:35–55.
- Jay MF, Sparks DL (1990) Localization of auditory and visual targets for the initiation of saccadic eye movements. In: *Comparative perception, Vol I, Basic mechanisms* (Berkley M, Stebbins W, eds), pp 351–374. New York: Wiley.
- Knudsen EI, Knudsen PF (1985) Vision guides the adjustment of auditory localization in young barn owls. *Science* 230:545–548.
- Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. *Science* 221:389–391.
- Mullette-Gillman OA, Cohen YE, Groh JM (2005) Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J Neurophysiol* 94:2331–2352.
- Mullette-Gillman OA, Cohen YE, Groh JM (2009) Motor-related signals in the intraparietal cortex encode locations in a hybrid, rather than eye-centered, reference frame. *Cereb Cortex* 19:1761–1775.
- Porter KK, Groh JM (2006) The “other” transformation required for visual-auditory integration: representational format. *Prog Brain Res* 155:313–323.
- Porter KK, Metzger RR, Groh JM (2007) Visual- and saccade-related signals in the primate inferior colliculus. *Proc Natl Acad Sci U S A* 104:17855–17860.
- Razavi B, O’Neill WE, Paige GD (2007) Auditory spatial perception dynamically realigns with changing eye position. *J Neurosci* 27:10249–10258.
- Recanzone GH (1998) Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc Natl Acad Sci U S A* 95:869–875.
- Schiller PH, True SD, Conway JL (1979) The effects of frontal eye field and superior colliculus ablations on eye movement. *Science* 206:590–592.
- Schlack A, Sterbing-D’Angelo SJ, Hartung K, Hoffmann KP, Bremmer F (2004) Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci* 25:4616–4625.
- Shinn-Cunningham B (2001) Models of plasticity in spatial auditory processing. *Audiol Neurootol* 6:187–191.
- Stricanne B, Andersen RA, Mazzoni P (1996) Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J Neurophysiol* 76:2071–2076.
- Wallace MT, Stein BE (1994) Cross-modal synthesis in midbrain depends on input from cortex. *J Neurophysiol* 71:429–432.
- Wallace MT, Meredith MA, Stein BE (1993) Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J Neurophysiol* 69:1797–1809.
- Werner-Reiss U, Kelly KA, Trause AS, Underhill AM, Groh JM (2003) Eye position affects activity in primary auditory cortex of primates. *Curr Biol* 13:554–562.
- Woods TM, Recanzone GH (2004) Visually induced plasticity of auditory spatial perception in macaques. *Curr Biol* 14:1559–1564.
- Zwiers MP, Van Opstal AJ, Paige GD (2003) Plasticity in human sound localization induced by compressed spatial vision. *Nat Neurosci* 6:175–181.
- Zwiers MP, Versnel H, Van Opstal AJ (2004) Involvement of monkey inferior colliculus in spatial hearing. *J Neurosci* 24:4145–4156.