

1

2

3

4 Hemisphere-Specific Properties of the Ventriloquism Aftereffect5 Norbert Kopčo^{1,2}, Peter Lokša¹, I-fan Lin³, Jennifer Groh⁴, and Barbara Shinn-Cunningham⁵6 ¹ Institute of Computer Science, Faculty of Science, P. J. Šafárik University, 04001 Košice,

7 Slovakia

8 ² Hearing Research Center and Department of Biomedical Engineering, Boston University,

9 Boston, Massachusetts 02215

10 ³ Taipei Medical School Shuang Ho Hospital, New Taipei City, 235 Taiwan11 ⁴ Center for Cognitive Neuroscience, Duke University, Durham, North Carolina 2770812 ⁵ Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 1521313 norbert.kopco@upjs.sk

14

15

16 Running Title: Hemisphere-Specificity of Ventriloquism

17

18 Portions of this work were presented at the 153rd Meeting of the Acoustical Society of

19 America.

20

21

22

23 **ABSTRACT (100 words max)**

24 Visual calibration of auditory space requires re-alignment of representations differing
25 in 1) format (auditory hemispheric channels vs. visual maps) and 2) reference frames (head-
26 centered vs. eye-centered). Here, a ventriloquism paradigm from Kopčo *et al.* [J Neurosci, 29,
27 13809-13814] was used to examine these processes in humans for ventriloquism induced within
28 one spatial hemifield. Results show that 1) the auditory representation can be adapted even by
29 aligned audio-visual stimuli, and 2) the spatial reference frame is primarily head-centered, with
30 a weak eye-centered modulation. These results support the view that the ventriloquism
31 aftereffect is driven by multiple spatially non-uniform, hemisphere-specific processes.

32

33 © 2019 Acoustical Society of America

34

35 PACS numbers: 43.66.Pn, 43.66.Qp, 43.66.Mk

36

37

38 **1. Introduction**

39 Vision plays an important role in calibration of auditory spatial perception. In the
40 “ventriloquism aftereffect” (VAE), repeated presentations of spatially mismatched visual and
41 auditory stimuli produce a shift in perceived sound location that persists when the sound is
42 presented alone (Canon, 1970; Recanzone, 1998; Woods and Recanzone, 2004; Bertelson *et*
43 *al.*, 2006). The brain mechanisms that support this process are mysterious because spatial
44 representations seem to differ in vision and in hearing in two ways.

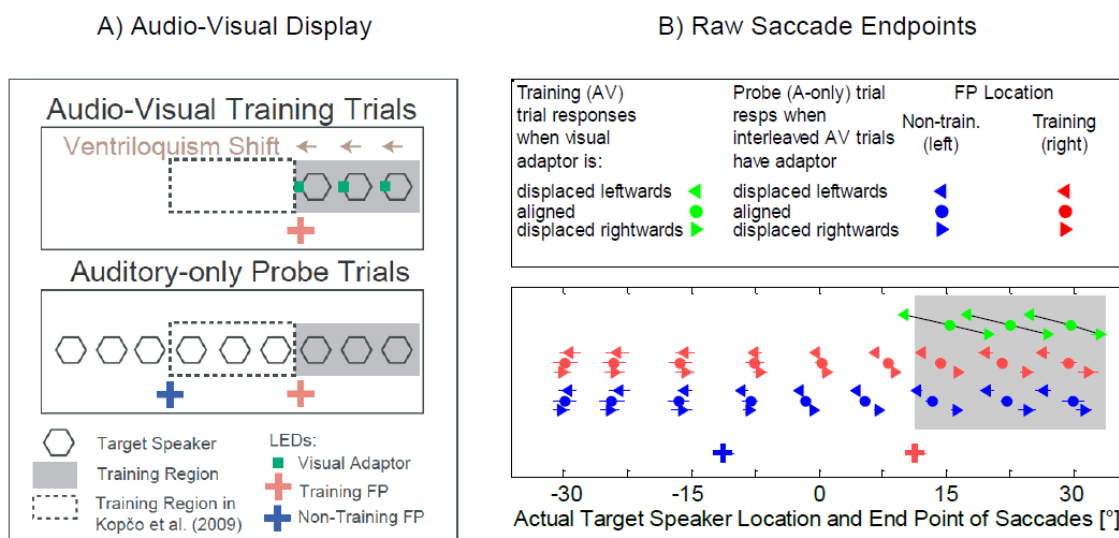
45 First, visual space is initially encoded relative to the direction of the eye gaze, while
46 the cues for auditory space are first computed relative to the orientation of the head (Groh &
47 Sparks, 1992). A means of reconciling this discrepancy in reference frames (RF) is necessary
48 to achieve correct recalibration. Our previous study suggests that a mixture of eye-centered
49 and head-centered RFs are associated with recalibration in the central region of the
50 audiovisual field (Kopco *et al.*, 2009).

51 Second, there is growing evidence that, in mammals, auditory space is encoded non-
52 homogeneously, based on two (or more) spatial channels roughly aligned with the left and
53 right hemifields of the horizontal plane (Grothe *et al.*, 2010; Groh, 2014). This is markedly
54 different from visual spatial codes, in which the retinal surface provides a map of the position
55 of stimuli in the environment.

56 Thus, the process of using visual information to recalibrate auditory space is multifaceted,
57 and may operate differently in different portions of the environmental scene. Indeed, differential
58 patterns of adaptation across auditory space have been observed (Phillips and Hall, 2005; Maier
59 *et al.*, 2010), suggesting that the auditory code in humans likely employs the same two-channel
60 scheme that has been observed in animal species (Salminen *et al.*, 2009).

61 Here, we tested whether the spatial characteristics of the ventriloquism aftereffect
62 induced in the audiovisual periphery (i.e., in a single hemifield) differ from those occurring

63 when the aftereffect is induced in the central region (i.e., covering both hemifields; Kopco *et*
 64 *al.*, 2009). Persistent visually driven biases in perceived sound location were induced. As in
 65 Kopco *et al.* (2009), we presented mismatched (5°-shifted) audio-visual (AV) stimuli in only
 66 a subregion of space (Fig. 1A, top panel), but this time the training region was peripheral,
 67 rather than central, to the fixation point used for these trials. We evaluated the effects of this
 68 pairing on saccade accuracy for interleaved auditory-only trials both from that fixation point
 69 and a non-training fixation point in the opposite hemifield (Fig. 1A, bottom panel).
 70 As was the case for our previous study involving central training, the pairing of a displaced
 71 visual stimulus induced a local aftereffect in the peripheral trained region. Contrary to the
 72 previous study, this aftereffect appeared to be mostly in the head-centered reference frame, as
 73 the contribution of an eye-centered component was not readily apparent. However, we also
 74 observed biases related to the location of the fixation point, even when the AV stimuli were
 75 aligned. Together, these findings confirm the contribution of multiple signals related to
 76 different reference frames and representational formats across the horizontal space.



77
 78 **Figure 1: Experimental set-up and raw experimental data. A) Audiovisual display used to present**
 79 **the AV training stimuli in one experimental block. At the beginning of each AV training trial (top), the**
 80 **subject had to fixate on the same initial fixation point (FP); then, the training stimulus was presented from**

81 one of three locations lateral to the FP, keeping the direction of the induced shift the same within a block
82 (by consistently presenting the visual adaptor displaced to the left, to the right, or aligned with the target
83 speaker). On the auditory-only probe trials (bottom), the same nine speaker locations and two FPs were
84 used in all blocks. The probe trials were randomly interleaved among the training trials and the FP and
85 target locations varied randomly from trial to trial. Dashed frame indicates the central training region
86 used in Kopco et al. (2009). B) Raw saccade endpoints of the responses to the AV training stimuli and
87 auditory-only probe stimuli as a function of the actual target speaker location, collapsed across time. The
88 symbols represent across-subject mean responses (± 1 SEM indicated by horizontal lines) in different
89 audiovisual conditions (see legend), separately for the training trials (green), probe trials starting at the
90 training fixation (red), and probe trials starting at the non-training fixation (blue). Graphs for each
91 measurement type are plotted in one row, vertically offset from data for other types, for visual clarity. The
92 A-only data corresponding to each target location are approximately aligned with that target location. For
93 the AV data, the dashed lines connect symbol triplets for the same auditory target when presented with
94 one of the three different visual adaptors (the AV-aligned data are located approximately at the
95 corresponding target location).

96

97 2. Methods

98 All procedures and equipment closely matched those used in Kopco *et al.* (2009).

99 *General methods.* Experiments were performed in an experimental lab in the Boston
100 University Hearing Research Center. Subjects made eye movements from a visual fixation
101 point to a broadband noise delivered from loudspeakers in darkness. On training trials (Fig.
102 1A, top), visual stimuli were presented simultaneously with the sounds, using light-emitting
103 diodes (LEDs) displaced from the locations of the speakers or aligned with them. On
104 randomly interleaved probe trials (Fig. 1A, bottom), only the auditory stimuli were presented.

105 *Subjects.* Seven young adults with normal hearing by self-report participated. The
106 experimental protocols were approved by the Boston University institutional review
107 committee.

108 *Setup.* Subjects were seated in a quiet darkened experimental room in front of an array
109 of speakers and LEDs (Fig. 1). To keep the head-centered RF fixed, the subjects' heads were
110 restrained by a chin rest. Subjects' behavior was monitored and responses were collected by
111 an infrared eye tracker, calibrated using visually guided saccades to selected target locations
112 at the beginning of each session.

113 *Stimuli.* Sounds were 100-ms broadband noises (0.2–6 kHz) with 10 ms on/off ramps
114 presented at 70 dBA from speakers mounted in the horizontal plane ~1.2 m from the center of
115 the listener's head. Spacing between speakers was 7.5°. For the training AV stimuli, only the
116 speakers at the locations 15°, 22.5°, and 30° were used (Fig. 1A). The LEDs for the AV
117 stimuli were mounted so that they were either horizontally aligned with the speakers or
118 displaced (either to the left or to the right) by 5°. They were turned on and off in synchrony
119 with the corresponding speakers. Two additional LEDs 10° below the speaker array served as
120 fixation locations (azimuths of $\pm 11.8^\circ$).

121 *Procedures.* Trials began with the onset of one of the two fixation LEDs. After
122 subjects fixated the LED for 150 ms, the fixation LED was turned off and the AV or A-only
123 stimulus was presented. The subjects performed a saccade to the perceived location of the
124 stimulus. The saccade end point was recorded at the saccade end, i.e., when the eye fixation
125 was sustained at the same location for 150 ms, at which point the experiment continued with
126 the next trial. In both AV and A-only trials, the subjects were instructed to look to the location
127 of the auditory component of the stimulus.

128 Training (AV) and probe (A-only) trials were randomly interleaved at a ratio of 1:1.
129 Training stimuli were presented from one of the 3 training locations while the subject fixated
130 the training fixation point (FP; top panel of Fig. 1A). Probe stimuli were presented from one
131 of the 9 speakers, while the subject fixated either the training or the non-training FP (bottom
132 panel of Fig. 1A).

133 Trials were run in sessions with a consistent AV pairing (leftward, rightward, or no
134 shift). Each session started with a pre-adaptation reference measurement (18 A-only trials
135 from the training fixation point), followed by 720 trials in which the training fixation point
136 and the AV shift direction was fixed. Each subject performed 12 sessions (2 fixation points x
137 3 shift directions x 2 repeats) in order that was randomized across the subjects.

138 *Data analysis.* Data from the first quarter of each session were excluded to remove
139 transitory values observed during the initial buildup of VAE. Within-session averages were
140 computed from the remaining data separately for each combination of target location, training
141 FP location, fixation position, and condition. Since no large left-right differences were
142 observed, data with training FP on the left were mirror-flipped and combined with the data
143 with training FP on the right (see Table 1). All data are presented as across-subject means and
144 standard errors of the mean, with the training FP always shown on the right and the non-
145 training FP on the left. Repeated measures ANOVAs were used to assess statistical
146 significance of the observed effects.

147 **3. Overall Design and Results**

148 As in Kopco et al. (2009), we presented paired visual-auditory stimuli in a subregion
149 of audiovisual space, fixed in both eye- and head-centered coordinates. We used one initial
150 eye fixation position on training trials and presented the discrepant audiovisual stimuli from a
151 restricted spatial range that was lateral with respect to the fixation point (see Fig. 1A, top).
152 Because the visual training was local, we could test the spatial attributes of the resulting
153 recalibration by shifting fixation on probe trials. Specifically, on interleaved auditory-only
154 probe trials, we varied initial eye position (FP) with respect to the head (which was fixed) and
155 presented sounds from all target locations spanning both the same head-centered locations and
156 the same eye-centered locations as on the training trials (see Fig. 1A, bottom). We first

157 consider the effects observed on the AV training trials themselves before turning to aspects of
158 how the effects generalize to the auditory-only conditions across both the trained and
159 untrained regions of space as a function of eye-referenced vs. head-referenced fixation
160 position.

161 *3.1 Ventriloquism effect*

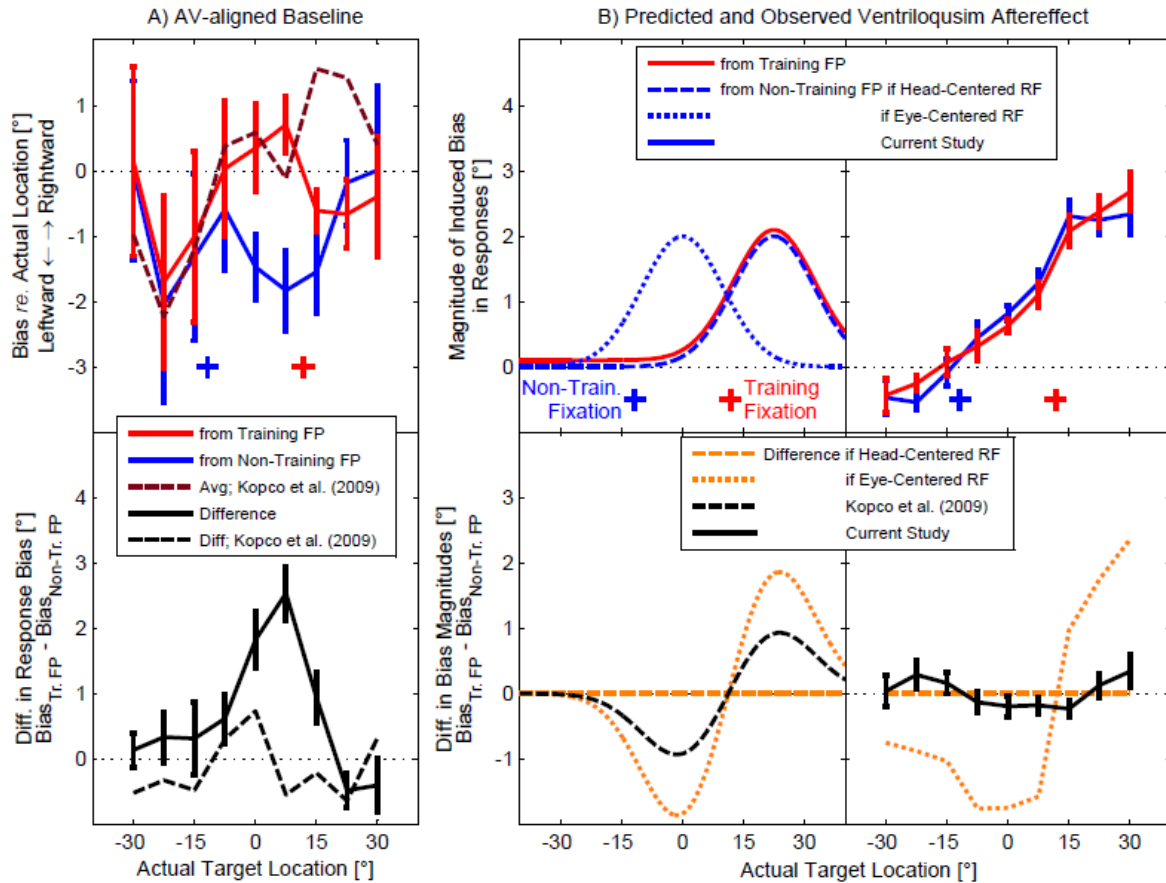
162 A strong ventriloquism *effect* – or capture of the auditory stimulus location by the
163 visual stimulus on combined AV trials - was observed. Green symbols in Fig. 1B show the
164 raw responses. When the AV stimuli were aligned, the average responses were not biased at
165 all. The relative strength of the ventriloquism effect was evaluated as percent of shift in
166 responses towards the visual (V) component re. the A-component on misaligned AV trials,
167 which was for each A target location and V-component shift computed as $(resp_{V-misalign} - resp_{V-align}) / (stim_{V-misalign} - stim_{V-align})$, where *stim* is the actual location of the V-component. The
168 strength ranged from 96% for the target at 15° to 82% for the target at 30° (averaged across 2
169 directions of induced shift). Even though there was a slight decrease in the strength of the
170 ventriloquism effect for the most lateral targets, it was expected that, as in Kopco *et al.*
171 (2009), this strong ventriloquism *effect* would be associated with a clear local ventriloquism
172 *aftereffect*.

174 *3.2 Gaze-dependent Effects During AV-aligned Baseline*

175 We next assessed the auditory-only responses interleaved with the spatially aligned
176 AV stimuli. The red and blue circles in Fig. 1B show these responses. Overall, the pattern of
177 results shows that the subjects accurately localized the auditory targets, showing a systematic
178 displacement of the responses with the actual target locations. To analyze the impact of the
179 visual training in more detail, the top panel of Fig. 2A shows the biases in these responses
180 relative to the actual target location, separately for the two fixation points. A gaze-direction-
181 dependent adaptation is seen when comparing the responses from the training FP (red) to

182 those from the non-training FP (blue). Specifically, the responses to the targets at azimuth of
183 0° - 15° were biased to the left by 1° - 2.5° when performed from the non-training FP (blue “+”
184 symbol) compared to the responses from the training FP (red “+” symbol). A dashed line in
185 this panel represents the same data from the central-adaptation experiment of Kopco *et al.*
186 (2009), averaged across the two FP locations as no large FP-dependent differences were
187 observed in that study. A solid black line in the bottom panel of Fig. 2A shows the difference
188 between the red and blue lines from the top panel, while the dashed line represents the
189 difference from the central-adaptation experiment of Kopco *et al.* (2009). These panels show
190 that responses to auditory-only stimuli from AV-trained locations that are lateral and near the
191 training FP differ depending on whether eyes fixate within the same hemifield or the opposite
192 hemifield. On the other hand, when the AV training locations are in the center, covering both
193 hemifields, no such differential effect of fixation location is observed (dash-dotted line). A
194 one-way repeated measures ANOVA performed on the difference data showed a significant
195 effect of target location ($F_{8,48} = 9.45$, $p < 0.001$). This effect of eye fixation direction is strong,
196 of size comparable to the VAE (see next section); thus, there is some eye-gaze-dependent
197 contribution to responses to auditory-only stimuli even when vision is not used to induce any
198 recalibration of the auditory spatial representation. However, this contribution is only visible
199 if the AV stimuli are presented within one spatial hemifield. Overall, the pattern of results in
200 the top panel of Fig. 2A for both experiments is that, independent of FP location, the
201 responses are mostly accurate in the trained region (all errors are much smaller than 1° , except
202 for the blue data point at 15°), while they tend to be biased away from the training region
203 outside of it (except for the left-most data point). This bias away is observed in all the non-
204 training subregions for both FPs and both experiments, with the exception of the trained-FP
205 data in the central region in the current experiment (3 red central targets in Fig. 2A are

206 approximately at 0°). Thus, the gaze-specific adaptation, which is observed in the same
 207 region, is likely caused by this lack of repulsion in the trained-FP central data.



208

209

210 **Figure 2. Adaptation induced by AV stimuli. A) Average bias in A-only responses in the AV-**

211 **aligned baseline condition as a function of the actual target location. Top panel shows mean response**

212 **biases (\pm SEM) when eyes are fixated at the training FP (red) and the non-training FP (blue). In addition,**

213 **the across-FP average data for central adaptation from Kopco et al. (2009) are shown for comparison**

214 **purposes (dashed line). The solid line in the bottom panel shows the difference between responses from**

215 **training FP and the non-training FP. The dashed line shows the difference for taken from Kopco et al.**

216 **(2009). B) Predicted and observed ventriloquism aftereffect. The top left panel plots the expected pattern**

217 **of biases induced in the A-only probe responses when preceding AV trials are presented in the training**

218 **region (15° - 30°). Red line shows predictions when the eyes fixate the training FP (i.e., the FP location**

219 **used during AV training trials). Dash-dotted blue line shows expected results from the non-training FP if**

220 **the RF of adaptation is head-centered, while dashed blue line shows expected results for an eye-centered**

221 **RF. The bottom panel shows the differences between the expected bias magnitudes from the training**
222 **versus the non-training FPs in the two RFs in orange. For comparison, the black dashed line sketches the**
223 **results corresponding to the mixed RF observed after VAE was induced in the central region in (Kopco**
224 **et al., 2009) . Top right panel shows the across-subject mean (\pm SEM) difference between the auditory**
225 **saccade end point locations when interleaved with spatially displaced AV stimuli vs. when interleaved**
226 **with AV-aligned stimuli, collapsed across the direction of the AV displacement. The solid black line in the**
227 **bottom right panel shows the effect of initial fixation position on the magnitude of the induced shift as the**
228 **across-subject mean (\pm SEM) difference between the shifts from the training and non-training FPs (i.e the**
229 **difference between the red and blue lines). Orange lines show the predictions of the difference for the two**
230 **reference frames based on the training FP data (red) from the top right panel.**

231

232 *3.3 Ventriloquism Aftereffect and its Reference Frame*

233 The expected pattern of ventriloquism aftereffect, and the predictions about the
234 reference frame based on it, are illustrated in the left-hand panels of Fig. 2B. The red line in
235 the top left panel shows the predicted magnitude of the aftereffect induced by the AV stimuli,
236 peaking in the trained region ($15^\circ - 30^\circ$) when assessed with eyes fixating the training FP. If
237 visually induced spatial plasticity occurs in a brain area using a head-centered RF, then shifts
238 in perceived sound location should occur mainly for sounds at the same head-centered
239 locations (in Fig. 2B, dash-dotted blue line matches the red line). Conversely, if plasticity
240 occurs in an eye-centered RF, then visually induced shifts should occur mainly for sounds at
241 the same eye-referenced locations (dotted blue line is shifted to the left of the red line by the
242 same displacement as the non-training FP is shifted relative to the training FP). The bottom
243 left panel summarizes the predicted results if evaluated as a difference between the responses
244 from the training and non-training FPs. The dash-dotted orange line shows the difference
245 between the red and dash-dotted blue lines, corresponding to the expected results if the
246 reference frame is head-centered. The dashed orange line shows the difference between the

247 red and dashed blue lines, corresponding to the expected results if the reference frame is eye-
248 centered. The dashed black line shows the predicted difference in the biases expected if the
249 RF is mixed, as observed in Kopco *et al.* (2009), in which case it should fall approximately in
250 the middle of the predictions of the two RFs shown in orange.

251 We assessed the auditory-only responses interleaved with the spatially mis-aligned AV
252 stimuli against these predictions. The red and blue triangles in Fig. 1B show the raw
253 responses in the conditions in which the ventriloquism aftereffect was induced in a leftward
254 direction (leftward-pointing triangles) or rightward direction (rightward-pointing triangles).
255 Overall, exposure to spatially mismatched AV stimuli resulted in a shift of responses to
256 sounds in the direction of the previously presented visual stimuli (compare the corresponding
257 triangles to the respective circles). To allow a detailed analysis of the results comparable with
258 the predictions of Fig. 2B, the red line in the top right panel of Fig. 2B plots the magnitude of
259 the bias in responses measured with eyes fixating the trained FP (red plus sign) *re.* no-shift
260 baseline from Fig. 1B, as a function of target location and averaged across the two directions
261 of induced shift (note that no main effect or interaction involving the direction factor were
262 significant in the ANOVA analysis, supporting this way of collapsing the data for
263 visualization; Table 1). The effect was strongest for the three right-most targets, i.e., in the
264 trained region, reaching approximately 2.3° (51% of the ventriloquism effect strength). It was
265 also location-specific, decreasing quickly toward zero outside of the trained region. These
266 results are consistent with the results of Kopco *et al.* (2009), confirming that the VAE can be
267 induced locally, so that it can be used to assess the VAE RF.

268 The reference frame of the VAE was examined by shifting the initial FP to a new
269 location and examining how the observed VAE changed. The blue line in the top right panel
270 plots the bias in responses measured with eyes fixating the new, non-trained FP (blue plus
271 sign), shifted by approximately 23° to the left from the trained FP. There was very little

272 difference in the measured VAE for the two FPs (blue line lies approximately on top of the
273 red line). Thus, the observed results are consistent with visual–auditory recalibration
274 occurring in a predominantly head-centered coordinate frame.

275 To compare the current results more directly to the predictions of the two models and
276 to the data of Kopco *et al.* (2009), a difference between the shift magnitudes from the two FPs
277 was computed (bottom right of Fig. 2B, black traces) and compared with predictions based on
278 the two models (orange traces). Again, the results are very close to the predictions of the
279 head-centered RF.

280 These results were confirmed by performing a 4-way repeated-measures ANOVA with
281 the factors of target speaker location (nine levels), fixation point of the trials (training vs. non-
282 training FP), AV-trial fixation point location (left vs. right), and the direction of induced shift
283 (left vs. right). The results of this analysis, summarized in Table 1, show that the main effect
284 of location was always significant, confirming that the ventriloquism aftereffect is spatially
285 specific and does not automatically generalize to the whole audiovisual field. The location by
286 FP interaction was also significant, showing that the reference frame of visual–auditory
287 recalibration is not purely head-centered, even though the eye-centered modulation is
288 relatively small.

289 **4. Discussion and Conclusions**

290 The current study examined the spatial properties of the ventriloquism aftereffect
291 induced by AV stimuli presented in only one spatial hemifield in the peripheral audio-visual
292 field. The goal was to ascertain how the ventriloquism aftereffect unfolds as a function of
293 multiple different spatial attributes: fixation position, generalization in head- vs. eye-centered
294 coordinates, and training within one spatial hemifield in contrast to training in both hemifields
295 (as in Kopco *et al.*, 2009). The results indicate that the ventriloquism aftereffect is a

296 multifaceted process, dependent on both the format of the neural representation of space in
297 hearing and vision, and on the reference frame used by the two senses.

298 In terms of the representational format, the location of the fixation position impacted
299 the pattern of adaptation induced by the AV stimuli, even when the AV-stimuli were
300 presented from matching locations and no VAE was induced. This unexpected adaptation was
301 not observed in the previous central-adaptation study (Kopco *et al.*, 2009). And, it is difficult
302 to identify its cause, since a baseline measurement with no AV stimulation was not
303 performed. However, a comparison of the central-adaptation and peripheral-adaptation data
304 suggests that adaptation away from the training region was observed in the AV-aligned data
305 in both experiments. Such expansion of space is consistent with previously observed inherent
306 biases towards the periphery (Razavi *et al.*, 2007). The current data shows that the inherent
307 biases might be more correctly described as biases away from the AV-training region, rather
308 than towards the periphery, and that the biases might be modulated by eye-gaze direction.
309 Specifically, in the current experiment in which the AV-aligned stimuli were presented in the
310 periphery, there was no repulsive bias in the central region when the gaze was fixated to a
311 point in the training hemifield, but it was observed if the gaze was fixated in the opposite
312 hemifield. At least three other factors of the current experimental design might also contribute
313 to the effect. First, the effect might be a result of adaptation to the auditory stimulus-
314 distribution, which becomes skewed when the training stimuli are included since all of them
315 come from one side (e.g., similar to adaptation reported by Dahmen *et al.* (2010). Second, the
316 visual signal might be causing some global ventriloquism-like adaptation outside the training
317 region, such that the auditory-only responses are shifted towards the region from which the
318 visual stimuli are frequently presented, but only when the FP is in the hemifield ipsilateral to
319 the AV stimulation (and such shift towards the training region cancels out the repulsion
320 observed otherwise). Whatever the specific mechanism, this adaptation effect shows that there

321 is a hemifield-specific integration of visual and auditory spatial signals that differs from the
322 integration occurring when the stimuli are presented centrally, covering both spatial
323 hemifields.

324 Regarding reference frames, the current results together with those of Kopco et al.
325 (2009) show that in humans the RF of VAE is a mixture of eye-centered and head-centered
326 coding. In the central region, the effect is a fairly even mixture of these two reference frames,
327 whereas in the periphery, the pattern more closely fits the head-centered predictions, but also
328 shows an interaction with eye position. This shows that the transformation of the visual and
329 auditory signals into an aligned reference frame, thought to be necessary for the ventriloquism
330 aftereffect to work, is non-uniform. While it is not immediately clear what form of non-
331 uniformity might be causing this pattern of results, it may be related to the hemispheric-
332 difference channel models of auditory space representation (Salminen *et al.*, 2009; Grothe *et*
333 *al.*, 2010; Groh, 2014).

334 Kopco et al. (2009) performed the central-adaptation ventriloquism experiments in
335 two rhesus monkeys in addition to the humans (footnote 1). In the monkeys, the RF was
336 mixed between head- and eye-centered frames, consistent with most neurophysiological
337 observations in the same species (Moriya *et al.*, 2013). Overall, these differences across
338 training regions (and, possibly, across species) suggest that the locations in the brain that are
339 recruited to accomplish this recalibration of auditory space may be widely varied. Some are
340 likely head-centered, some are eye-centered, some may involve the position of the eyes in the
341 orbits per se. These sites of plasticity may be recruited differently depending on the training
342 region and whether it spans both head-centered hemifields or is contained within one.

343 Additional experimental and/or modeling studies are needed to test alternative
344 explanations about the different reference frames of the ventriloquism aftereffect as well as
345 about the unexpected AV-aligned adaptation effect. However, the current results demonstrate

346 that there are hemisphere-specific adaptation processes in visual recalibration of auditory
347 space, resulting in different FP-dependent patterns of adaptation depending on the region in
348 which adaptation is induced.

349

350 **Acknowledgments**

351 This work was supported by the SRDA, project DS-2016-0026, EU H2020-MSCA-
352 RISE-2015 grant no. 69122, and by the EU RDP projects TECHNICOM I, ITMS:
353 26220220182, and TECHNICOM II, ITMS2014+:313011D23. Barbara Shinn-Cunningham
354 was supported by NIH R01 DC013825 Jennifer Groh was supported by NIH NS50942 and
355 NSF 0415634.

356

357 **Footnotes**

358 ¹ The current experiments were also performed in two rhesus monkeys. A detailed treatment
359 of these effects can be found in the bioRxiv preprint BIORXIV/2019/564682.

360

361 **References and links**

362 Bertelson, P., Frissen, I., Vroomen, J., and de Gelder, B. (2006). "The aftereffects of
363 ventriloquism: Patterns of spatial generalization," *Percept. Psychophys.* **68**, 428-436.
364 Canon, L. K. (1970). "Intermodality inconsistency of input and directed attention as
365 determinants of the nature of adaptation," *J. Exp. Psych.* **84**, 141-147.
366 Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L., and King, A. J. (2010). "Adaptation to
367 Stimulus Statistics in the Perception and Neural Representation of Auditory Space,"
368 *Neuron* **66**, 937-948.
369 Groh, J. M. (2014). "Making space: how the brain knows where things are.," Cambridge, MA:
370 Harvard University Press.
371 Grothe, B., Pecka, M., and McAlpine, D. (2010). "Mechanisms of sound localization in
372 mammals.," *Physiol Rev* **90**, 983-1012.
373 Kopco, N., Lin, I. F., Shinn-Cunningham, B. G., and Groh, J. M. (2009). "Reference Frame of
374 the Ventriloquism Aftereffect," *J. Neurosci.* **29**, 13809-13814.
375 Maier, J. K., McAlpine, D., Klump, G. M., and Pressnitzer, D. (2010). "Context Effects in the
376 Discriminability of Spatial Cues," *Jaro-J Assoc Res Oto* **11**, 319-328.

- 377 Moriya, T. J., Groh, J. H., and Meynet, G. (2013). "Episodic modulations in supernova radio
378 light curves from luminous blue variable supernova progenitor models," *Astron*
379 *Astrophys* **557**.
- 380 Phillips, D. P., and Hall, S. E. (2005). "Psychophysical evidence for adaptation of central
381 auditory processors for interaural differences in time and level," *Hear. Res.* **202**, 188-
382 199.
- 383 Razavi, B., O'Neill, W. E., and Paige, G. D. (2007). "Auditory Spatial Perception
384 Dynamically Realigns with Changing Eye Position," *J. Neurosci.* **27**, 10249-10258.
- 385 Recanzone, G. H. (1998). "Rapidly induced auditory plasticity: The ventriloquism
386 aftereffect," *Proceedings of the National Academy of Sciences* **95**, 869-875.
- 387 Salminen, N. H., May, P. J., Alku, P., and Tiitinen, H. (2009). "A population rate code of
388 auditory space in the human cortex. ," *Plos One* **4:e7600**.
- 389 Woods, T. M., and Recanzone, G. H. (2004). "Visually Induced Plasticity of Auditory Spatial
390 Perception in Macaques," *Current Biology* **14**, 1559-1564.
- 391
- 392

393 **Table 1.** Four-way repeated-measures ANOVA of the VAE magnitude data.

Factor	d.f.	F	Signif.
Speaker Location (1 to 9)	8, 48	33.87	***
A-only Fixation Point (Tr. vs. Non-Tr.)	1, 6	0.99	
Direction of Induced Shift (L vs. R)	1, 6	0.43	
AV Fixation Point (L vs. R)	1, 6	0.27	
Speaker Location X A-only FP	8, 48	0.79	*
Speaker Location X AV FP	8, 48	2.28	
A-only Fixation Point X AV FP	1, 6	0.42	
Speaker Location X Direction	8, 48	0.56	
AV Fixation Point X Direction	1, 6	2.16	
A Fixation Point X Direction	1, 6	0.1	
Speaker Loc. X AV FP X A-only FP	8, 48	0.31	
Speaker Loc. X AV FP X Direction	8, 48	0.52	
Speaker Loc. X A-only FP X Direction	8, 48	1.69	
AV FP X A-only FP X Direction	1, 6	0.12	
Loc. X AV FP X A-only FP X Direct.	8, 48	1.16	

394

395 Significance levels are as follows: * $p < 0.05$, *** $p < 0.005$.

396

397 **Figure Captions**

398 **Figure 2: Experimental set-up and raw experimental data. A) Audiovisual display used to present**
 399 **the AV training stimuli in one experimental block. At the beginning of each AV training trial (top), the**
 400 **subject had to fixate on the same initial fixation point (FP); then, the training stimulus was presented from**
 401 **one of three locations lateral to the FP, keeping the direction of the induced shift the same within a block**
 402 **(by consistently presenting the visual adaptor displaced to the left, to the right, or aligned with the target**
 403 **speaker). On the auditory-only probe trials (bottom), the same nine speaker locations and two FPs were**
 404 **used in all blocks. The probe trials were randomly interleaved among the training trials and the FP and**
 405 **target locations varied randomly from trial to trial. Dashed frame indicates the central training region**
 406 **used in Kopco et al. (2009). B) Raw saccade endpoints of the responses to the AV training stimuli and**
 407 **auditory-only probe stimuli as a function of the actual target speaker location, collapsed across time. The**
 408 **symbols represent across-subject mean responses (± 1 SEM indicated by horizontal lines) in different**
 409 **audiovisual conditions (see legend), separately for the training trials (green), probe trials starting at the**
 410 **training fixation (red), and probe trials starting at the non-training fixation (blue). Graphs for each**
 411 **measurement type are plotted in one row, vertically offset from data for other types, for visual clarity. The**
 412 **A-only data corresponding to each target location are approximately aligned with that target location. For**
 413 **the AV data, the dashed lines connect symbol triplets for the same auditory target when presented with**
 414 **one of the three different visual adaptors (the AV-aligned data are located approximately at the**
 415 **corresponding target location).**

416

417 **Figure 2. Adaptation induced by AV stimuli. A) Average bias in A-only responses in the AV-**
 418 **aligned baseline condition as a function of the actual target location. Top panel shows mean response**
 419 **biases (\pm SEM) when eyes are fixated at the training FP (red) and the non-training FP (blue). In addition,**
 420 **the across-FP average data for central adaptation from Kopco et al. (2009) are shown for comparison**
 421 **purposes (dashed line). The solid line in the bottom panel shows the difference between responses from**
 422 **training FP and the non-training FP. The dashed line shows the difference for taken from Kopco et al.**
 423 **(2009). B) Predicted and observed ventriloquism aftereffect. The top left panel plots the expected pattern**
 424 **of biases induced in the A-only probe responses when preceding AV trials are presented in the training**
 425 **region ($15^\circ - 30^\circ$). Red line shows predictions when the eyes fixate the training FP (i.e., the FP location**

426 used during AV training trials). Dash-dotted blue line shows expected results from the non-training FP if
427 the RF of adaptation is head-centered, while dashed blue line shows expected results for an eye-centered
428 RF. The bottom panel shows the differences between the expected bias magnitudes from the training
429 versus the non-training FPs in the two RFs in orange. For comparison, the black dashed line sketches the
430 results corresponding to the mixed RF observed after VAE was induced in the central region in (Kopco *et*
431 *al.*, 2009) . Top right panel shows the across-subject mean (\pm SEM) difference between the auditory
432 saccade end point locations when interleaved with spatially displaced AV stimuli vs. when interleaved
433 with AV-aligned stimuli, collapsed across the direction of the AV displacement. The solid black line in the
434 bottom right panel shows the effect of initial fixation position on the magnitude of the induced shift as the
435 across-subject mean (\pm SEM) difference between the shifts from the training and non-training FPs (i.e the
436 difference between the red and blue lines). Orange lines show the predictions of the difference for the two
437 reference frames based on the training FP data (red) from the top right panel.
438