PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Visual recalibration of auditory spatial perception decays at different time scales

Patrick BRUNS[1]; Brigitte RÖDER[1]

[1] University of Hamburg, Germany

ABSTRACT

In human observers, exposure to audiovisual stimuli with a consistent spatial disparity results in an adjustment of auditory spatial representations that corrects for the mismatch between auditory and visual input (known as the ventriloquism aftereffect). Our recent findings suggest that such adjustments occur in parallel at multiple time scales. However, it is not clear if recalibration of auditory space, once emerged, remains stable until new audiovisual evidence is encountered, or would decay over time. Here we reanalyzed data from a previous study to answer this question. In this study, human participants localized sounds in complete darkness, immediately before and after they were exposed to 200 audiovisual stimuli in which the visual stimulus was consistently displaced to the right of the sound source. A trial-by-trial analysis of the localization responses in the auditory localization posttest suggested that the size of the ventriloquism aftereffect was strongest at the beginning and decreased over the course of the 96 unimodal localization trials. Thus, cross-modal recalibration decayed although participants did not experience any new audiovisual evidence. Despite this rapid decay, a stable carryover effect on sound localization was observed on the next day, suggesting distinct underlying mechanisms for immediate and sustained cross-modal recalibration effects.

Keywords: Multisensory, Spatial hearing, Ventriloquism aftereffect

1. INTRODUCTION

Visual spatial information is directly encoded in the retinal image, whereas the auditory system has to infer the location of an auditory object from binaural and monaural cues that are generated by the interaction of the sound waves with the head and external ears (1). Visual spatial processing is usually more precise and dominates auditory spatial processing. For example, when an auditory and a visual stimulus are presented synchronously but at discrepant spatial locations, the perceived location of the sound source is usually shifted toward the visual location, a phenomenon commonly referred to as the ventriloquism effect (2,3). In addition to multisensory spatial integration evident in the ventriloquism effect, the brain uses visual information to constantly recalibrate unisensory auditory spatial representations. This can be demonstrated by briefly exposing human observers to audiovisual stimuli with a consistent spatial discrepancy. In what has become widely known as the ventriloquism aftereffect, the localization of subsequently presented unimodal auditory test stimuli was found to be shifted to correct for the previously encountered cross-modal spatial mismatch (4,5). The behavioral ventriloquism aftereffect has been associated with adjustments in auditory cortical processing (6,7). However, it is currently unclear how stable such rapid adjustments are after cessation of audiovisual stimulation.

Cross-modal recalibration affects unimodal sound localization even after a single audiovisual exposure stimulus (8), although the size of the ventriloquism aftereffect increases if several audiovisual exposure trials with a consistent spatial disparity precede the auditory test trials (9-12). In fact, cumulative recalibration seems to depend on different underlying mechanisms than immediate trial-by-trial recalibration. When two tones of different frequencies were paired with opposite directions of audiovisual spatial mismatch (leftward vs. rightward), a corresponding frequency-specific cumulative aftereffect was obtained, although, on a trial-by-trial level, sound localization was additionally modulated by the direction of audiovisual disparity in the directly preceding audiovisual trial, indicating a sound frequency-invariant immediate aftereffect (13). Thus,

[1] patrick.bruns@uni-hamburg.de

cumulative evidence for a new cross-modal spatial association seems to result in adjustments of spatial representations in lower-level, tonotopically organized auditory regions which are presumably more stable than immediate adjustments mediated by higher-level, sound frequency-invariant brain regions (6,7,14).

In a recent study, we directly tested the longevity of the cumulative ventriloquism aftereffect after cessation of audiovisual discrepancy training (15). In this study, participants were exposed to spatially discrepant audiovisual stimuli on three consecutive days. On each day, their unimodal sound localization was tested before and after the audiovisual exposure phase. The audiovisual spatial discrepancy was either constant (always 13.5°) or varied from day to day (i.e., incrementally increased from 4.5° to 13.5°). We found that repeated exposure to audiovisual stimuli with a constant spatial disparity enhanced the ventriloquism aftereffect over the course of the three days, whereas no enhancement was seen after training with varying audiovisual discrepancies. Importantly, aftereffects in the constant group were still present at the beginning of the second and third session (i.e., 24 hours after cessation of the previous audiovisual exposure phase) and accumulated with additional audiovisual discrepancy training. Thus, despite exposure to congruent audiovisual input in the natural environment, the ventriloquism aftereffect was retained and consolidated between sessions. This finding is seemingly at odds with earlier studies which reported a rapid decline of the ventriloquism aftereffect when a short delay was introduced between audiovisual exposure and unimodal auditory test trials (11,12).

Here we reanalyzed the data from our previous study (15) to test the possibility that cross-modal recalibration involved two separate processes operating at distinct time scales, a strong but rapidly decaying within-session aftereffect and a weaker but more stable between-session aftereffect (12). To this end, we quantified the change in the size of the ventriloquism aftereffect over the course of the 96 localization trials in each unimodal auditory localization test. This allowed us to assess whether the aftereffect remained stable immediately after the end of each audiovisual exposure phase, or whether it started to decay despite the carryover effect on sound localization that was observed on the next day.

2. METHOD

2.1 Participants

The datasets of 42 healthy adult volunteers (29 women and 13 men; mean age: 24.4 years; age range: 19-46 years) from our previous study (15), publicly available in the APA repository hosted by the Center for Open Science (16), were reanalyzed for the present study. The original study featured two experimental conditions, one in which participants were exposed to a constant audiovisual spatial discrepancy of 13.5° on all three days of the experiment (constant group), and one in which the presented degree of audiovisual spatial disparity increased over the course of the three days (incremental group). Only the data from participants who had received constant training (n = 42) were considered for the present analysis because a between-session aftereffect was only obtained with constant training but not with incremental training (15). Experimental testing had been approved by the ethics commission of the German Psychological Society (DGPs) and was performed in accordance with the ethical standards laid down in the Declaration of Helsinki.

2.2 Experimental Procedure

The experimental procedure has been described in detail elsewhere (15). In brief, participants completed three experimental sessions on consecutive days. Experimental sessions were conducted in a dark sound-attenuated room. Each session consisted of three unimodal auditory localization test blocks and two audiovisual exposure blocks (see Figure 1). Unimodal sound localization performance was measured at the beginning of each session (Pre), immediately after the first audiovisual exposure block (Post1), and again immediately after the second audiovisual exposure block (Post2).

Auditory stimuli were 750-Hz tones with a duration of 200 ms (including 5-ms linear rise/fall envelopes), which were presented from eight different loudspeaker locations (spanning ±31.5° in steps of 9°) at 65 dB(A). The loudspeakers were mounted on a semicircular frame at a distance of 90 cm and were hidden from view by an acoustically transparent curtain. In the auditory localization test blocks, each location was presented 12 times in a randomized order, resulting in 96 trials per test block. In each trial, participants indicated the perceived location of the sound source with a rotatable hand pointer which was mounted in front of them. They had to align the hand pointer within ±10° before each trial to ensure a constant starting position for the pointing movements.

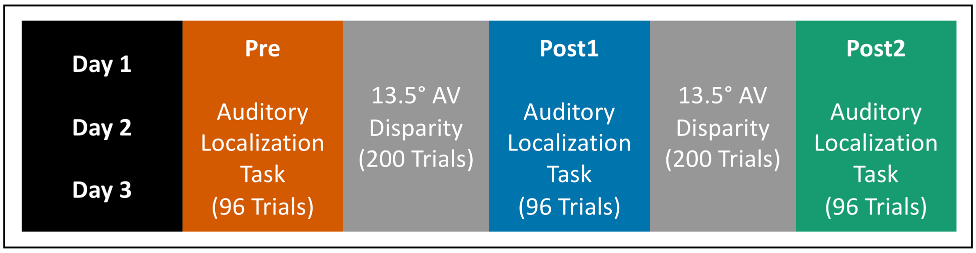

Figure 1 – Schematic illustration of the study design

Audiovisual exposure blocks consisted of 200 trials each (25 trials per loudspeaker location), in which the sound was presented together with a synchronous visual stimulus (red laser dot) that was displaced by 13.5° to the right of the sound source. Audiovisual stimuli were presented at a rate of one stimulus per second. In 4% of the trials, either the visual or the auditory stimulus was interrupted for 150 ms. Participants had to detect these rare deviant stimuli, but did not engage in an active localization task during audiovisual exposure.

2.3 Data Analysis

As in our previous study (15), two separate analyses were performed to assess within-session recalibration effects (measured at Post1 and Post2) and between-session recalibration effects (measured at Pre on Day 2 and 3). Within-session recalibration effects, reflecting the effect of the directly preceding audiovisual exposure block, were calculated by subtracting mean localization responses at Pre from single-trial Post1 responses and mean Post1 responses from single-trial Post2 responses in each session. Single-trial shifts in sound localization relative to the preceding test block were then averaged separately for three bins of trials (Trials 1-32, 33-64, and 65-96) to assess a potential decay of recalibration during auditory test blocks (i.e., after cessation of audiovisual stimulation). The resulting values were submitted to a three-way repeated measures analysis of variance (ANOVA) with factors of Bin (1-32, 33-64, 65-96), Block (Post1, Post2), and Day (1, 2, 3). Similarly, between-session recalibration effects were calculated by subtracting mean localization responses at baseline (Pre) on Day 1 from single-trial responses at Pre on Day 2 and at Pre on Day 3 and averaging the resulting values separately for each bin of trials. Between-session shifts in sound localization were submitted to a two-way repeated measures ANOVA with factors of Bin (1-32, 33-64, 65-96) and Day (2, 3).

Binning of trials was done to increase the signal-to-noise ratio of individual data for inferential statistics. In addition, we inspected the trial-by-trial trend of within-session and between-session recalibration effects at a group level. For this purpose, shifts in sound localization at each of the 96 trial positions were averaged across days and participants, separately for Pre (Days 2-3), Post1 (Days 1-3), and Post2 (Days 1-3). Simple linear correlation analyses were performed to assess whether there was a linear relationship of localization shifts (i.e., recalibration effects) with time (i.e., trial number) at the group level.

3. RESULTS

Sound localization responses relative to Pre on Day 1 are shown in Figure 2, separately for each localization test and bin of trials. Within each session, there was a clear shift in sound localization after the first audiovisual exposure block at Post1 compared to Pre which was particularly pronounced on Day 1. Sound localization shifts further increased, but to a lesser extent, after the second audiovisual exposure block at Post2. Accordingly, a repeated measures ANOVA on the within-session recalibration effects revealed a significant main effect of Block, $F(1, 41) = 24.62$, $p < .001$, as well as a significant interaction between Block and Day, $F(2, 82) = 3.70$, $p = .029$. In addition, a significant main effect of Bin was obtained, $F(2, 82) = 6.03$, $p = .004$. Crucially, there was no interaction between Bin and any of the other factors (all ps $\geq .230$). The main effect of Bin indicates that across the three days and the two posttest blocks, within-session aftereffects significantly decreased from the first bin of trials (M = 2.3°, 95% CI [1.7, 2.9]) over the second bin (M = 1.8°, 95% CI [1.3, 2.4]) to the third bin (M = 1.4°, 95% CI [0.9, 1.9]).

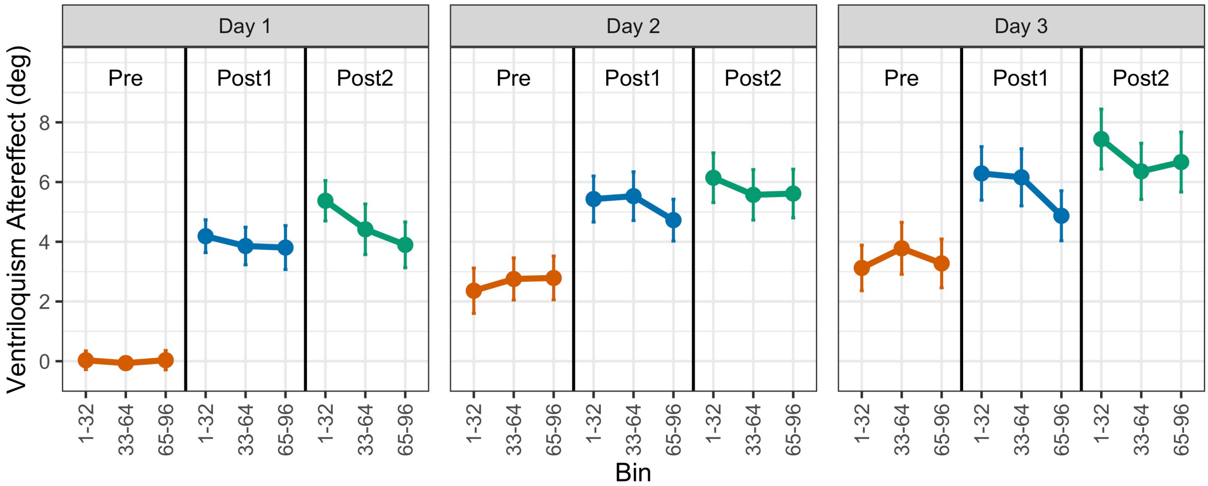

Figure 2 – Mean shifts in sound localization (with *SEM*s) relative to Pre on Day 1

In addition to the within-session sound localization shifts observed immediately after audiovisual exposure (Post1 and Post2), there was a clear between-session recalibration effect. Compared to Pre on Day 1, sound localization was shifted both at Pre on Day 2 and at Pre on Day 3, despite the 24-hour interval since the last audiovisual exposure block (see Figure 2). Accordingly, a repeated measures ANOVA on the between-session recalibration effect yielded a highly significant intercept, $F(1, 41) = 21.80$, $p < .001$, indicating that across days and bins, localization shifts were significantly larger than zero. The size of the shift did not differ significantly between Day 2 and 3, $F(1, 41) = 1.60$, $p = .213$. Moreover, and in contrast to the within-session aftereffects measured at Post1 and Post2, between-session aftereffects at Pre did not differ between the three bins of trials. Neither the main effect of Bin, $F < 1$, nor the interaction between Bin and Day, $F < 1$, was significant.

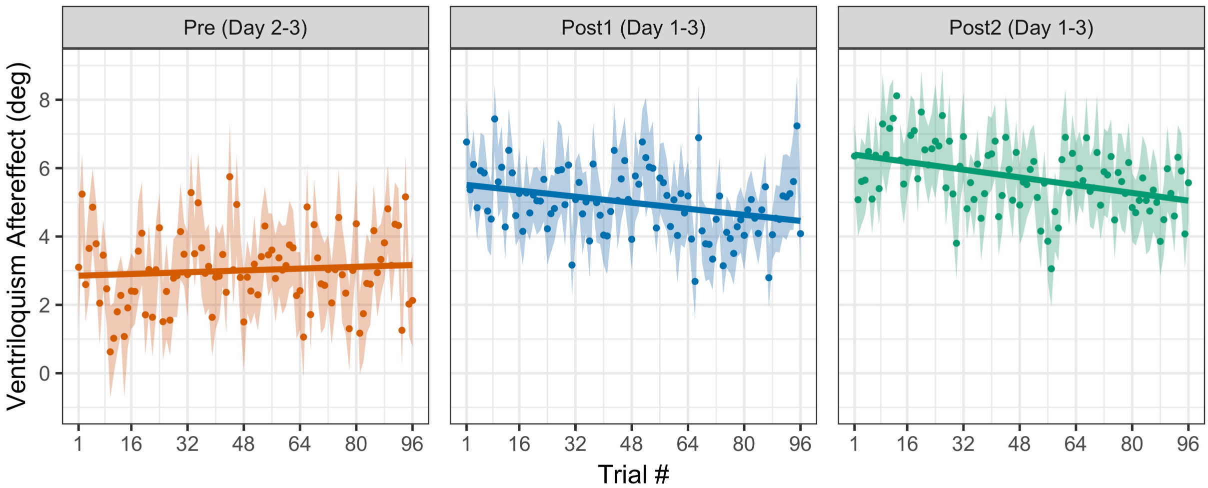

Figure 3 – Mean shifts in sound localization per trial (with *SEM*s) and best-fitting regression lines

To further investigate the time course of between-session aftereffects during Pre blocks and of within-session aftereffects during Post1 and Post2 blocks, we averaged localization shifts for each trial position across days and participants. As can be seen in Figure 3, the size of between-session aftereffects measured at Pre was uncorrelated with time (i.e., trial number), $r = .08$, $p = .422$. By contrast, there was a significant negative correlation between localization shifts and time both for Post1, $r = -.32$, $p = .001$, and Post2, $r = -.42$, $p < .001$.

4. DISCUSSION

The present study examined the time course of the visual recalibration of auditory localization in the ventriloquism aftereffect, both within and across three experimental sessions that were conducted on consecutive days. In each session, unimodal sound localization was measured before and after each of two brief audiovisual exposure blocks in which sounds were paired with visual stimuli that had a constant spatial discrepancy of 13.5°. Sound localization was shifted in the direction of the visual stimuli immediately after each audiovisual exposure block, indicative of the typical within-session ventriloquism aftereffect (4,5,9). We found that the size of the within-session aftereffect linearly decayed by about 1° over the course of the 96 auditory test trials in each posttest, but remained significant until the end of each posttest phase. Despite this rapid decay of cross-modal recalibration, a residual shift in sound localization was still present at the beginning of the next session 24 hours later (15). This between-session aftereffect was smaller than the localization shift at the end of the last posttest phase, but did not further decay over the course of the 96 pretest trials. Thus, our results suggest distinct underlying mechanisms for immediate and sustained cross-modal recalibration effects.

Our finding that the size of within-session aftereffects decreased over the course of the 96 auditory posttest trials could indicate a simple decay of the aftereffect over time. Alternatively, exposure to the unimodal auditory test stimuli (i.e., without the spatially discrepant visual stimuli) might have triggered an additional unisensory recalibration process that counteracted the previous audiovisual recalibration effect (17). However, previous studies also found a decline of the audiovisual ventriloquism aftereffect when a short delay (without any stimulus presentations) was introduced between the audiovisual exposure and subsequent auditory test trials (11,12). Thus, a simple decay of the aftereffect over time seems a more likely explanation for our findings. In contrast to our results, Frissen et al. (9) reported no decline of the ventriloquism aftereffect over the course of the auditory posttest. However, in their study, only 27 posttest trials were administered (compared to 96 trials in the present study), which might have been too short to detect a decline in the size of the aftereffect.

Despite the rapid and approximately linear decay of the ventriloquism aftereffect which we observed during the first few minutes after cessation of audiovisual exposure, recalibration was not fully abolished 24 hours later at the beginning of the next session, although the carryover effect to the next session was smaller than the ventriloquism aftereffect at the end of the last posttest block on the preceding day. If this between-session aftereffect would still represent the same process as the aftereffect observed immediately after audiovisual exposure, one would need to assume that the decay of the aftereffect seen during the posttest trials rapidly tapered off and reached an asymptote (12). We did not measure sound localization in the time period between the second posttest and the pretest on the next day. However, it seems unlikely that recalibration would have affected sound localization throughout this 24-hour period. Participants were exposed to naturally occurring audiovisual stimuli with their usual spatial correspondence between sessions, which should have counteracted any residual recalibration effects from the brief audiovisual exposure during the test sessions.

A more likely explanation for our results seems that recalibration effects were stored and later retrieved in a context-specific manner. This assumption is supported by our observation that between-session aftereffects emerged only when the audiovisual spatial discrepancy remained constant throughout the three sessions (15). Moreover, recalibration effects were found to become more specific for the trained stimuli with cumulative audiovisual exposure (13). Concordantly, in a study by Hofman et al. (18), participants relearned to localize sounds with molds attached to their pinnae which modified spectral sound localization cues. Crucially, after accurate sound localization with the molds was learned, participants could easily switch back and forth between normal and modified spectral cues, suggesting that the new spatial correspondence was stored and did not have to be reacquired each time the participants encountered the modified spectral cues. Taken together, these findings suggest that exposure to a new cross-modal spatial correspondence, as in the present study, might induce two dissociable recalibration processes, an immediate but transient adjustment of sound localization and a sustained, context-specific learning of the altered spatial correspondence which requires consolidation after the audiovisual training (19).

5. CONCLUSIONS

In sum, the present study shows that cross-modal recalibration of auditory localization induced by a brief period of exposure to spatially discrepant audiovisual stimuli, known as the ventriloquism aftereffect, decays rapidly after the end of audiovisual stimulation even if no new audiovisual evidence is encountered. In addition, however, a separate mechanism seems to store and consolidate the newly acquired spatial correspondence, allowing for a context-specific retrieval of audiovisual learning that is stable for at least one day after training. Thus, exposure to spatially discrepant audiovisual stimuli seems to trigger adjustments of auditory sensory representations through different mechanisms and at multiple time scales in parallel, a finding which needs to be considered in applied settings such as the design of virtual reality displays involving audiovisual stimulation.

ACKNOWLEDGEMENTS

The study was supported by the German Research Foundation (DFG) through Grant TRR 169/A1. We thank Elisa Bußkamp, Theresa Paumen, and Samantha Schröder for help collecting data.

REFERENCES

1. Recanzone GH, Sutter ML. The biological basis of audition. Annu Rev Psychol. 2008;59:119-142.
2. Chen L, Vroomen J. Intersensory binding across space and time: a tutorial review. Atten Percept Psychophys. 2013;75(5):790-811.
3. Recanzone GH. Interactions of auditory and visual stimuli in space and time. Hear Res. 2009;258:89-99.
4. Lewald J. Rapid adaptation to auditory-visual spatial disparity. Learn Mem. 2002;9:268-278.
5. Recanzone GH. Rapidly induced auditory plasticity: the ventriloquism aftereffect. Proc Natl Acad Sci USA. 1998;95(3):869-875.
6. Bruns P, Liebnau R, Röder B. Cross-modal training induces changes in spatial representations early in the auditory processing pathway. Psychol Sci. 2011;22(9):1120-1126.
7. Zierul B, Röder B, Tempelmann C, Bruns P, Noesselt T. The role of auditory cortex in the spatial ventriloquism aftereffect. Neuroimage. 2017;162:257-268.
8. Wozny DR, Shams L. Recalibration of auditory space following milliseconds of cross-modal discrepancy. J Neurosci. 2011;31(12):4607-4612.
9. Frissen I, Vroomen J, de Gelder B. The aftereffects of ventriloquism: the time course of the visual recalibration of auditory localization. Seeing Perceiving. 2012;25(1):1-14.
10. Mendonca C, Escher A, van de Par S, Colonius H. Predicting auditory space calibration from recent multisensory experience. Exp Brain Res. 2015;233(7):1983-1991.
11. Bosen AK, Fleming JT, Allen PD, O'Neill WE, Paige GD. Accumulation and decay of visual capture and the ventriloquism aftereffect caused by brief audio-visual disparities. Exp Brain Res. 2017;235(2):585-595.
12. Bosen AK, Fleming JT, Allen PD, O'Neill WE, Paige GD. Multiple time scales of the ventriloquism aftereffect. PLoS One. 2018;13(8):e0200930.
13. Bruns P, Röder B. Sensory recalibration integrates information from the immediate and the cumulative past. Sci Rep. 2015;5:12739.
14. Bruns P, Röder B. Cross-modal learning in the auditory system. In: Lee AKC, Wallace MT, Coffin AB, Popper AN, Fay RR, editors. Multisensory processes: the auditory perspective. Springer Handbook of Auditory Research. Vol. 68. Cham, Switzerland: Springer; 2019. p. 221-242.
15. Bruns P, Röder B. Repeated but not incremental training enhances cross-modal recalibration. J Exp Psychol Hum Percept Perform. 2019;45(4):435-440.
16. Bruns P. Data for: Repeated but not incremental training enhances cross-modal recalibration [Internet]. OSF; 2019. Available from: osf.io/xcskn
17. Machulla T-K, Di Luca M, Froehlich E, Ernst MO. Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. Exp Brain Res. 2012;217(1):89-97.
18. Hofman PM, Van Riswick JGA, Van Opstal AJ. Relearning sound localization with new ears. Nat Neurosci. 1998;1(5):417-421.
19. Powers AR 3rd, Hillock AR, Wallace MT. Perceptual training narrows the temporal window of multisensory binding. J Neurosci. 2009;29(39):12265-12274.