BRILL

Seeing and Perceiving 25 (2012) 1–14

brill.nl/sp

The Aftereffects of Ventriloquism: The Time Course of the Visual Recalibration of Auditory Localization *

Ilja Frissen [1,2,], Jean Vroomen [1] and Beatrice de Gelder [3,4]**

[1] Department of Medical Psychology and Neuropsychology, Tilburg University, Tilburg, The Netherlands

[2] LUNAM Université, Ecole Centrale de Nantes, Institut de Recherche en Communications et Cybernétique de Nantes, Nantes, France

[3] Cognitive and Affective Neuroscience Laboratory, Tilburg University, Tilburg, The Netherlands

[4] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA

Received 2 February 2011; accepted 25 September 2011

Abstract

Exposure to synchronous but spatially discordant auditory and visual inputs produces adaptive recalibration of the respective localization processes, which manifest themselves in measurable aftereffects. Here we report two experiments that examined the time course of visual recalibration of apparent sound location in order to establish the build-up and dissipation of recalibration. In Experiment 1 participants performed a sound localization task before and during exposure to an auditory–visual discrepancy. In Experiment 2, participants performed a sound localization task before and after 60, 180 or 300 exposures to the discrepancy and aftereffects were measured across a series of post-adaptation sound localization trials. The results show that recalibration is very fast. Substantial aftereffects are obtained after only 18–24 exposures and asymptote appears to be reached between 60 and 180 exposures. The rate of adaptation was independent of the size of the discrepancy. The retention of the aftereffect was strong, as we found no dissipation, not even after as few as 60 exposure trials.

Keywords

Sensory perception, multisensory processes, perceptual adaptation, sound localization, ventriloquism after effect, time course

---

* This article is part of the Multisensorial Perception collection, guest edited by S. Wuerger, D. Alais and M. Gondan.

** To whom correspondence should be addressed. E-mail: ilja.frissen@irccyn.ec-nantes.fr

 DOI:10.1163/187847611X620883

1. Introduction

The visual and the auditory system maintain coordinated representations of external space, which is presumably achieved by systematically cross-checking between the two modalities. Such interactions come to light when presenting an observer with synchronous but spatially discrepant auditory and visual information. This typically creates a percept in which sound is located nearer to the location of the visual input (Bermant and Welch, 1976; Bertelson and Radeau, 1981; Klemm, 1909). This visual effect on auditory location is generally known as the ventriloquist effect (Bertelson, 1999; Recanzone, 2009). Exposure to the ventriloquism situation also leads to compensatory aftereffects, consisting in post-exposure shifts in auditory localization (Canon, 1970, 1971; Frissen *et al.*, 2003, 2005; Lewald, 2002; Radeau, 1973; Radeau and Bertelson, 1974; Recanzone, 1998), and sometimes also in visual localization (e.g., Radeau, 1973; Radeau and Bertelson, 1974; 1976: Experiment 1). The effect has also been demonstrated in non-human primates (e.g., Kopco *et al.*, 2009; Recanzone, 1998). It is generally agreed that aftereffects reflect a recalibration process that results in a reduction of the perceived discrepancy, and could play an important role in achieving and maintaining a coherent intersensory representation of space (Held, 1965; Welch, 1978).

What is not known in any kind of detail is the time course of this recalibration. It has been suggested that visual recalibration of auditory localization occurs very rapidly (Lewald, 2002; Recanzone, 1998). In fact, in earlier studies conducted in our laboratory we observed that one minute of exposure is sufficient to establish a reliable aftereffect (Bertelson *et al.*, 2006; Frissen *et al.*, 2003, 2005). Although it was not the main focus of the study, Radeau and Bertelson (1976) report several acquisition functions obtained in two experiments under different experimental conditions. The two experiments were essentially the same except for the particular task the participants performed during exposure to the auditory–visual spatial conflict. In the first task, participants pointed at the apparent location of the visual input, and in the second at that of the auditory input. The corresponding acquisition functions showed evidence for very fast adaptation. In the first experiment, visual aftereffects reached asymptote of approximately 1° after as little as five exposure blocks (each consisting in five single exposure trials). In the second experiment, auditory aftereffects reached asymptote of approximately 2° apparently somewhat later, after 20–25 exposure blocks. One study to look systematically at the acquisition function of ventriloquism aftereffects is by Bertelson (1993), which also confirmed that recalibration is indeed very fast. After as little as 5–8 exposure episodes (each consisting in six single presentations of spatial conflict) to an auditory–visual spatial discrepancy recalibration appeared to have reached asymptote, which seemed to depend only on the size of the spatial discrepancy: the larger the discrepancy the larger the asymptote. Unfortunately, since it was a conference presentation, only very little information is available about the experimental details. Most recently, Wozny and Shams (2011) conducted an experiment that allowed them to investigate recalibration at the shortest time scale so far. By interleaving short (35 ms) unisensory auditory and visual stimuli with auditory–visual stimuli they were able to measure small (i.e., 5% of the discrepancy) but significant aftereffects after a single exposure to an auditory–visual discrepancy. Interestingly, this was true despite the fact that in the experiment size and direction of the discrepancy was continuously changing. The authors conclude that the perceptual system is in a continuous state of recalibration, although the process is apparently supervised by the perceived unity of the multisensory stimuli. In other words, recalibration occurs only then when the stimuli are judged as belonging together.

Whereas little is known about the acquisition of the aftereffect of ventriloquism, virtually nothing is known about its dissipation. There have been some informal observations that the effect lasts tens of minutes (e.g., Recanzone, 2009) but that is after typically 20–30 min of adaptation, and no systematic study is available on the relationship between adaptation duration and dissipation. Wozny and Shams (2011) on the other hand found very rapid dissipation in the order of seconds, although that observation was based on aftereffects established after a single exposure. Knowledge of dissipation times is not only of great practical use when studying recalibration; it is also of theoretical value as it can point to the locus of adaptation. Very fast dissipation betrays a peripheral or sensory locus whereas extremely long retention times indicate the involvement of central processes. For instance, aftereffects of a peripheral locus of adaptation, such as the color afterimage, tend to decay in a matter of seconds whereas more complex aftereffects, such as the contingent color aftereffect can still be effective days after exposure (e.g., McCollough, 1965). Moreover, acquisition and dissipation functions, either by themselves or in concert, can also be very effective tools in distinguishing between perceptual processes. The work of Bertelson and colleagues on auditory–visual speech perception provides a relevant example. They showed that exposure to incongruent auditory–visual speech (i.e., a McGurk type situation; McGurk and MacDonald, 1976) can lead to the recalibration of auditory speech identification and that this effect went in the opposite direction of another already known effect, that of selective speech adaptation (Bertelson *et al.*, 2003). This contrast already provided an indication that different perceptual processes were at play. Two subsequent time course studies, one on acquisition and another on dissipation gave further evidence of this. The acquisition study (van Linden *et al.*, 2004) showed that, whereas recalibration quickly reached asymptote and after a while even decreased back to baseline, the selective speech adaptation effect continued to increase slowly as exposure continued. Similarly, the dissipation study showed differential patterns of decay (Vroomen *et al.*, 2004).

Thus, the aim of the present study was to study both the acquisition and dissipation functions of the visual recalibration of auditory spatial perception. Both experiments were conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and participants gave their informed consent.

2. Experiment 1: Acquisition Functions

The experiment was modeled on that of Bertelson (1993). Accordingly the experimental procedure consisted of three consecutive phases: a pretest, an exposure phase to an auditory–visual spatial discrepancy, and an 'erasure' phase. The erasure phase was installed as a kind of perceptual 'reset' to counter the aftereffects established during the second experimental phase in which the participant was exposed to a spatial discrepancy. Thus the erasure phase enabled us to test different conflicts within a single session.

2.1. Methods

2.1.1. Participants

Twenty students from Tilburg University (age 19–29, eleven female), all naïve as to the purpose of the experiment, and with normal hearing and normal or corrected to normal vision, participated in two sessions each.

2.1.2. Apparatus and Stimuli

Testing was carried out in a dark and sound attenuated booth. Participants sat in front of a table with their head restrained by a fixed chinrest at 40 cm above the tabletop. The setup involved nine display units, an array of push buttons and a response box. Display units, which were occluded by means of an acoustically transparent black cloth, each consisted of a loudspeaker (Visaton, FRWS 5, $Ø = 5$ cm) with an LED ($Ø = 1$ cm) over its center. All units were arranged in a horizontal array, at 90 cm distance and 20 cm below eye level, spanning from -20 to $+20°$ at $5°$ intervals. The three most central loudspeakers (-5, 0, $+5°$) were used for auditory localization trials, while of the remaining units only the LEDs were used. To collect localization responses, 108 pushbuttons were arranged on the tabletop along another semi-circular array, at $1°$ intervals, and placed just comfortably at arm length. Thus, the pushbuttons are located at the end of the pointing movement. Performance on catch trials (see procedure) was recorded with a separate response box, placed 20 cm directly in front of the participant.

The auditory stimulus was a 200 ms long 750 Hz pure tone, with 5 ms linear on and offsets, presented at 64 dB (A). The (synchronous) LED flashes also lasted 200 ms, and were clearly visible through the occluding cloth when lit.

2.2. Design and Procedure

Two within-subjects factors were manipulated. One of these was the direction of visual distracter. The visual distracter was either to the left or to the right of the sound stimulus. The other factor was discrepancy. The spatial discrepancy between the auditory and the visual stimuli was 5, 10, 15 or 20°. The resulting eight conditions were run in a blocked fashion and divided over two sessions with the restriction that the direction of the visual distracter was always the same within a session. Sessions were run on separate days. Half the participants started with the distracter to the left, and the other half with the distracter to the right. In both sessions the four different discrepancies were administered in four consecutive and balanced (Latin square) runs. Between runs there was a break for saving the intermediate data and initiating the next run.

Each run was made up of three consecutive phases: a pretest, an exposure phase and an erasure phase, and lasted about 5 min. The auditory stimulus in the pretest was a train of six tones extending over a period of several seconds. The pretest phase consisted of 18 randomized auditory localization trials, 6 for each of the three central loudspeakers. On each trial a 2200 ms train of 6 tones (inter-stimulus interval: 200 ms) was presented. Participants were allowed to point as soon as the train started and were allowed another 2500 ms after the train had ended. The instruction was to always press the push button that was in the apparent direction of the sound using the dominant hand for all pointing and catch trial (see below) responses. Occasionally a participant pressed two buttons, in which case the apparatus was set up to record the button that was pressed first. The exposure phase was divided into 12 blocks, each consisting of a number of exposure trials followed by a single localization trial at one of the three test locations. Exposure trials were 6 presentations, at 1 s intervals, of the condition's particular auditory–visual discrepancy with the adapter sound from the median loudspeaker (i.e., 0°) and the visual distracter to its left or to its right, depending on the session. A single localization trial (identical to the pretest) followed the exposure trials after 1300 ms. In this manner, each of the three loudspeakers were tested four times, in a quasi-random order, across the 12 blocks in the spatial discrepancy phase.

The erasure phase was similar to the discrepancy phase. It was divided into 6 erasure blocks, each consisting in a number of exposure trials and a single localization trial. Exposure trials were now 6 presentations of the auditory and visual stimuli from the same location (i.e., both in the median plane), at 1 s intervals. The localization trial followed the exposure trials after 1300 ms and the participant was once again allowed 2500 ms to respond. In this manner, each of the three loudspeakers was tested twice, in a quasi-random order, across the 6 blocks in the erasure phase.

To ensure that the participant attended to the stimuli, there were occasional catch trials across both the auditory–visual discrepancy phase (a total of four) and the erasure phase (two), which consisted of the single omission of a visual distracter. This could occur in any of the adaptation blocks except for the very first one. It was the participant's task to detect these occurrences and to indicate this by pressing a button on the response box.

Before starting with the actual experiment the experimenter demonstrated the pointing task and the catch trial detection task to the participant by running a truncated version of an experimental run. This version consisted of six pretests and five erasure trials, with catch trials on four of these erasure trials in the four possible positions, which the experimenter indicated to the participant.

2.3. Results and Discussion

The data of four participants were excluded from further analysis either because of sub-normal performance on catch trials ($<75\%$; two participants) or for not being able to reliably discriminate between the three test locations (two other participants). The remaining participants' catch trials scores were high, ranging from 92 to 100%. All p-values for individual t-tests were Bonferroni corrected. Violations of sphericity assumption were dealt with by applying the Greenhouse–Geisser (for epsilons < 0.75) or Huyn–Feldt (for epsilons $\geqslant 0.75$) correction.

Aftereffects of the post-test and erasure phase were calculated by subtracting individual localization responses from the pretest phase. Aftereffects were counted as positive when they went in the direction of the visual distracter (during the discrepancy phase). Aftereffects were normalized such that the magnitude was expressed as a percentage of the auditory–visual discrepancy. Whereas aftereffects are normally calculated on a relatively large series of post-tests, here we necessarily had only a single localization test to assess the aftereffect at any particular position in the block (see procedure). We tried to reduce the consequent noise by pooling the aftereffects across the three test locations and the direction of the visual distracter, and finally binning aftereffects for two consecutive trials.

The results are shown in Fig. 1 in which the acquisition curves are shown (panel a) and the overall aftereffects across time series (panel b). A number of observations are made. First, after as little as 3–4 blocks of six exposures to an auditory visual discrepancy we find substantial aftereffects in the expected direction for all discrepancy sizes except for 5°. This means that recalibration occurs within less than half a minute. Second, the curves in panel a follow somewhat different time courses. The curves for the 10 and 20° discrepancy first increase rapidly, but show a drop at or after 5–6 adaptation blocks only to level off at around 20%. The 15° curve shows a steady, nearly monotonously increasing trend. The 5° curve appears negative over the entire test period. However, even uncorrected one-sample t-tests showed that none of the six post-tests were significantly different from zero (all t-values < 1.91, all p-values > 0.075), and neither was the aftereffect after pooling across post-test (see Fig. 1(b)), $t(15) = 1.85$, $p = 0.084$. A 6 (time) $\times$ 3 (discrepancy) repeated measures ANOVA, which excluded the data from the 5° discrepancy condition showed no significant effect for discrepancy ($F(2, 20) = 1.79$, $p = 0.19$), time ($F(5, 50) = 2.27$, $p = 0.071$), or the interaction term ($F(10, 100) = 1.09$, $p = 0.38$). A trend analysis revealed a significant cubic trend ($F(1, 10) = 1.73$, $p = 0.031$) which is consistent with the above description of the curve.

Third, erasure trials were successful in that the overall aftereffects were not significantly different from zero (Fig. 1(b)). One-sample t-tests, showed that the overall aftereffect in the erasure phase was not significantly different from zero for the three largest discrepancies (all p-values > 0.4). In keeping with the generally anomalous results in the 5° condition, there was a very large negative aftereffect after the first two erasure trials, although there was a tendency for the aftereffect

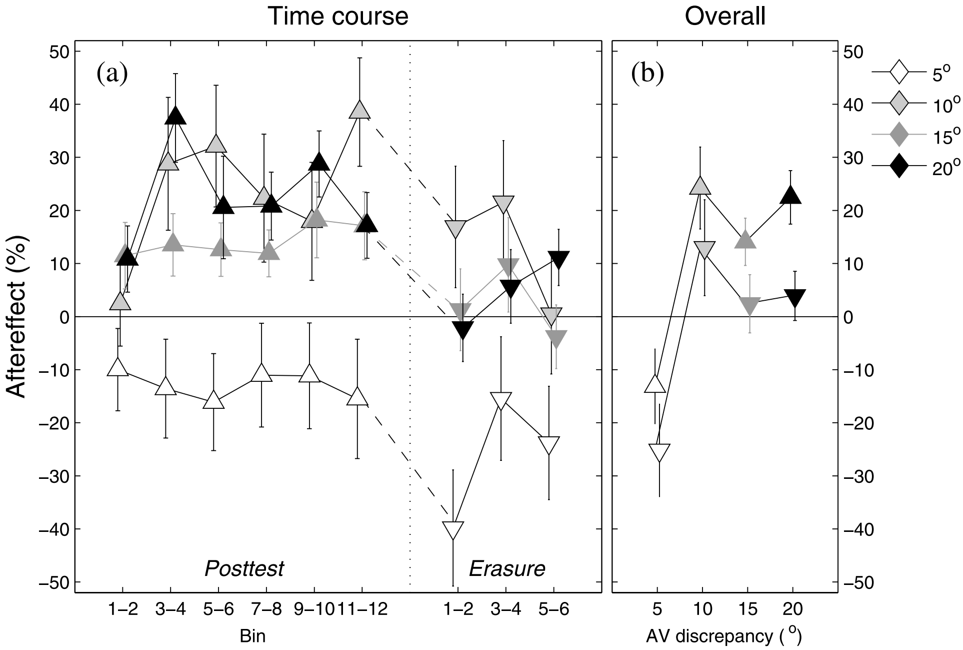

Figure 1. Experiment 1: acquisition functions. (a) Mean aftereffects (as a percentage of the auditory–visual discrepancy) are plotted as a function of the amount of exposure to an auditory–visual stimulus. Error bars represent the SEM. On the abscissa are bins of two blocks of adaptation trials each block corresponding to six unique presentations of the auditory–visual stimulus presented at a rate of 1 s^{-1} (see also method). The panel is divided into two parts (vertical dotted line). The left hand part (labeled post-test) corresponds to those trials where there was a spatial discrepancy between the auditory and the visual stimulus during adaptation. The parameter is the size of the auditory–visual spatial discrepancy (see legend). The right hand part (erasure) corresponds to those trials in which there was no discrepancy during adaptation. (b) Mean aftereffects average across the entire post-tests (▲) and erasure (▼) phase of the experiment, as a function of the spatial discrepancy during post-test.

to go to zero after that. The overall aftereffect for the erasure trials for 5° was still significantly different from zero ($t(15) = 2.89$, $p = 0.044$).

In sum, for the three largest discrepancies tested here recalibration was fast and the acquisition did not seem to be dependent on the size of the discrepancy. Quite different results were obtained for adaptation to a 5° discrepancy where we failed to observe an aftereffect in the expected direction of the visual stimulus. The aftereffects were not significantly different from zero but with a tendency to go in the opposite direction; a particular pattern that has been observed before (Bertelson *et al.*, 2006; Eramudugolla *et al.*, 2011; Lewald, 2002; Passamonti *et al.*, 2009).

3. Experiment 2: Dissipation Functions

In Experiment 2 we aimed at examining the rate of dissipation as a function of adaptation time. We expected that adaptation would consolidate with more exposure: in other words, the more exposure, the slower the dissipation rate. To test this, participants were adapted to 60 (i.e., 1 min), 180 (3 min) and 300 (5 min) presentations of an auditory–visual spatial discrepancy of 15° and aftereffects were then measured across a series of 27 post-tests (i.e., 1.5 min).

3.1. Method

3.1.1. Participants

Nine new students from Tilburg University (age 17–26, 3 male), all naïve to the purpose of the experiment, and with normal hearing and normal or corrected to normal vision, participated in three sessions each.

3.1.2. Apparatus and Stimuli

The setup was the same as in Experiment 1.

3.2. Procedure

We ran a complete within-subject design with two factors, exposure duration (60, 180 or 300 exposures) and direction of the visual distracter (15° to the left or to the right of the sound). All six conditions were run twice, for a total of 12 runs. Each run was equally divided over three sessions with each session dedicated to one level of the exposure duration factor. All was counterbalanced except for the direction of the visual distracter which alternated over runs within each session.

A run was made up of three consecutive parts: a pretest, exposure to the auditory–visual spatial discrepancy, and a post-test. The pretest consisted of 27 completely randomized auditory localization trials, 9 from each of the three central loudspeakers. On each trial a single tone was presented, and participants were allowed a fixed period to respond (3.330 s, including the 200 ms of the tone). The participant was instructed to always press the push button that was in the apparent direction of the sound. The post-test was the same as the pretest except for the randomization of the trials. Post-test runs were organized in nine blocks of three trials with one trial for each test location. Within and across blocks care was taken that each position was tested in all sequential positions. Six different permutations of post-test trial orders were created which were rotated across runs.

The exposure phase to spatial discrepancy consisted of 60, 180 or 300 exposures to the condition's particular auditory–visual discrepancy, at a presentation rate of 1/s (i.e., 1, 3 or 5 min, respectively). The spatial discrepancy was presented across the three central speakers positions (-5, 0 and $+5°$). In particular, the auditory–visual stimulus was presented five times at one location after which it moved to a random new location. To ensure that the participant attended to the exposure stimuli, there were occasional catch trials (2, 6 or 9, depending on the number of exposures), which consisted in the omission of one visual distracter. It was the participant's task to detect these occurrences by pressing a button on the response box.

3.3. Results and Discussion

Overall, performance on the catch trials was high (>80%) and none of the participants was excluded. Aftereffects were calculated as before by subtracting the individual post-test localization responses from the mean localization response on the corresponding speaker location in the pretest and counted positive when they went in the direction of the visual distracter. The results are shown in the two panels of Fig. 2.

The most striking observation from Fig. 2(a) is that there is no evidence of any dissipation over the time sampled. All three functions are as good as level and show no sign of decline. The curve for 60 exposures appears to start at zero. This was attributable to a single outlying point ($-11.7°$, participant 6, $z = -2.12$), which was subsequently treated as a missing value in the statistical analysis. The square marker shows the mean aftereffect with this one point excluded. The lack of dissipation means that retention after as little as 1 min exposure is already very strong. The aftereffects was entered in a 3 (exposures: 60, 180, 300 min) $\times$ 27 (serial position of the post-test) repeated measures ANOVA, which indeed showed no significant effect of serial position ($F(26, 130) = 1.56$, $p = 0.165$).

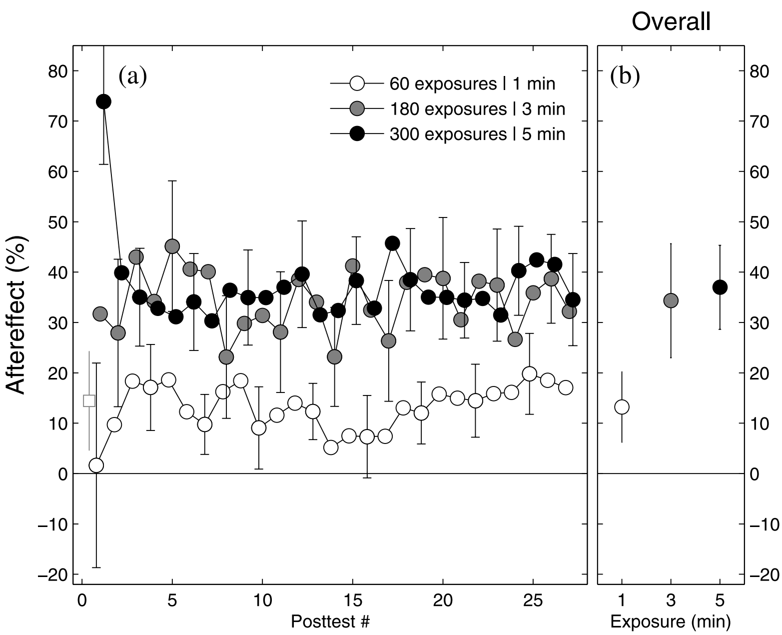

Figure 2. Experiment 2: dissipation functions. (a) Mean aftereffects (as a percentage of the auditory–visual discrepancy) as a function of time (post-test number) after adaptation. Error bars represent the SEM. The parameter is the duration of adaptation (see legend). With the exception of the very first trial standard error was relatively uniform across the series and therefore, for the sake of clarity, only a subset of error bars are shown. The single square represents the mean aftereffect with the exclusion of one single outlying data point at the very first post-test (participant #6, see also results). (b) Overall mean aftereffects as a function of adaptation duration.

The ANOVA also showed a clear effect of the number of exposures ($F(2, 16) = 6.05$, $p = 0.012$). After 180 and 300 presentations aftereffects were larger than after 60 presentations. This was confirmed by contrast analysis that showed a significant difference between 60 and 180 exposures ($F(1, 8) = 5.35$, $p = 0.049$), but not between 180 and 300 exposures ($F < 1$). This is further illustrated in Fig. 2(b) where the means across the whole post-test are plotted.

4. General Discussion

The two experiments reported here explored the acquisition and dissipation of the aftereffects of ventriloquism. The acquisition of aftereffects and therefore visual recalibration of auditory localization was found to be very fast. Experiment 1 showed that reliable aftereffects were obtained after as few as 18–24 individual exposures to the auditory–visual discrepancy, while asymptote was reached after 180 exposures (i.e., 3 min). Experiment 2 showed that aftereffects became larger with increasing exposure time (Fig. 2(b)). This supports the claim made after the first experiment that, for a 15° discrepancy, asymptote has not yet been reached after 60 exposure trials (i.e., 1 min). Since there was no significant difference for 180 and 300 exposure trials, this may represent an asymptote; however, further research will be required to confirm this conjecture. These results make explicit and qualify what has been hinted at or suggested by earlier studies (Bertelson, 1993; Lewald, 2002; Radeau and Bertelson, 1976; Recanzone, 1998). Recanzone (1998) and later Lewald (2002) refer to the recalibration as rapid but use an adaptation period of 20–30 min (see also Frissen *et al.*, 2003). The present findings clearly surpass these notions of rapidity by showing that recalibration actually occurs on a time scale that is an order of magnitude faster (see also Wozny and Shams, 2011).

The retention of recalibration is long relative to the duration of the adaptation. Aftereffects were retained for longer than the 27 post-adaptation localization trials, even after only 60 exposure trials. At first, this seems to be contradictory with the results reported by Wozny and Shams (2011) who report (although do not quantify) a rapid decrease of the aftereffect. However, in their case, the aftereffect was the product of a single exposure to an auditory–visual discrepancy and therefore arguably the weakest (i.e., least robust) recalibration possible. One possibility is that aftereffects produced by slightly longer exposure, as in the present case, are consolidated more efficiently and thereby essentially permanent — that is, permanent until new information about the mapping between auditory and visual perceptual space becomes available, such as, for instance, the erasure trials in Experiment 1. Incidentally, this is not inconsistent with the idea that the system is in a constant state of recalibration (e.g., Wozny and Shams, 2011) since without a new auditory–visual discrepancy there is no new error signal to calibrate to and the current calibration is maintained. It is also possible that the time range of post-exposure testing was too short to detect a decrease in aftereffects. In that case, dissipation apparently occurs after at least 27 trials (i.e., 90 s). The relatively long retention has been interpreted as an indication of a shift in auditory space (Recanzone, 2009). That is, the shift is implemented at a level where the various auditory localization cues (e.g., interaural time and level differences, and spectral cues) have been integrated into a coherent representation. This is in accord with other studies that provided evidence for a locus of recalibration that is at least beyond the level of the peripheral inter-aural localization mechanisms (Frissen *et al.*, 2003, 2005; Lewald, 2002; Passamonti *et al.*, 2009; Recanzone, 1998).

One anomalous finding was the aftereffects in the 5°discrepancy condition in Experiment 1, or rather the lack of clear aftereffects, since they never were significantly different from zero. One possible account for this is to consider that whenever an observer makes a perceptual estimate of a stimulus there is a certain amount of error associated with this estimate that has at least two sources. First, every measurement system, whether biological or mechanical, has a certain amount of (random) measurement error (e.g., Ernst, 2006). Second, if left to its own devices, a measurement system's calibration tends to drift. Not only does this lead to systematic errors in the affected system, it also shifts the mapping between the auditory and visual space. Depending on the type of error, different strategies are required and the problem for the perceptual system is to determine which kind of error needs to be accounted for. In the case of a shift in mapping the best strategy is to measure the error and adjust the perceptual estimates to minimize the error. However, if the error is due to measurement noise then any adjustment would in fact introduce a systematic error (see Burge *et al.*, 2008 for a formal discussion of the effect of random measurement error and miscalibration error on the rate of recalibration). Since it is known that the spatial acuity of the auditory system is in the order of several degrees (e.g., Middlebrooks and Green, 1991) a relatively small discrepancy between the auditory and visual estimate, such as the 5° auditory–visual spatial discrepancy in Experiment 1, could be interpreted as random measurement noise in the auditory system. A large error, on the other hand, would be unlikely due to measurement error and therefore interpreted as a miscalibration and (rapidly) adjusted for. Given the relatively short adaptation period, one might speculate that the 5° discrepancy was interpreted as sensory noise and recalibration was considered inappropriate. Only after prolonged exposure would the error become interpretable as a systematic one and therefore be adjusted for. We therefore expect to find (positive) aftereffects for a 5° discrepancy after longer periods of adaptation. Consistent with this are the findings of Recanzone (1998) who found significant aftereffects for a relatively small discrepancy (8°) after 20–30 min of adaptation.

Alternatively, it was suggested to us that the tendency for the negative effect could be because when participants are exposed to offsets to the left (or right) for a given session then participants could be judging the sound source relative to the average offset for that session (hence, in the wrong direction). In other words, participants adopt a new auditory subjective straight ahead (SSA) at the center of the distribution of (perceived) locations, which would tend to counter the effects of the recalibration. However, following this logic, the perceived distribution is more 'skewed' with the larger discrepancies and presumably the SSA should be shifted more as well, and thereby its influence. The results, on the other hand, suggest that the presumed influence of the SSA shift is largest in the condition with the smallest discrepancy condition where it apparently was big enough to outweigh the recalibration. One way to account for this is to make the additional assumption that at some point the effects of the recalibration in turn start outweighing those of the shift in the SSA. The potential influence of the (perceived) spatial distribution of the stimuli on recalibration should be the topic of further research.

In conclusion, the present study is one of the first to look specifically at the time course of visual recalibration of auditory localization. Whereas some researchers already acknowledged recalibration to be fast (e.g., Lewald, 2002; Recanzone, 1998) the rate of the acquisition found in the present study even surpasses these informal observations. The present results may be instructive for the study of the neural underpinnings of auditory–visual spatial interactions. It is becoming increasingly clear that, in addition to the multisensory inputs from 'classic' association areas and the thalamic nuclei, there are direct connections from both the primary and the non-primary visual cortex (Bizley *et al.*, 2007; Budinger *et al.*, 2006). In line with this, sensitivity to visual stimulation has been widely demonstrated in the auditory cortex of humans (Giard and Peronnet, 1999), and neuroimaging studies (Calvert *et al.*, 1999; Kayser *et al.*, 2010) have shown that auditory–visual interactions occur in early auditory areas. Studies looking specifically at ventriloquism have found evidence for the involvement of the planum temporale (Bonath *et al.*, 2007) and primary auditory cortex (Recanzone, 1998) and the geniculo-striate circuit within the visual system (Passamonti *et al.*, 2009). From a methodological point of view, the rapidity of recalibration and its retention can be exploited in future neuroimaging studies as 'lasting' perceptual changes can be acquired in very little time indeed, and might even allow its study in 'real time'.

Acknowledgements

The authors would like to thank Paul Bertelson for being an ever present source of motivation and Jess Hartcher-O'Brien. Constructive criticisms by the anonymous reviewers are gratefully acknowledged.

References

Bermant, R. I. and Welch, R. B. (1976). Effect of degree of separation of visual–auditory stimulus and eye position upon spatial interaction of vision and audition, *Percept. Mot. Skills* **42**, 487–493.

Bertelson, P. (1993). The time-course of adaptation to auditory–visual spatial discrepancy, in: *Proc. 6th Conf. Eur. Soc. Cognit. Psychol.*, Copenhagen, Denmark.

Bertelson, P. (1999). Ventriloquism: a case of crossmodal perceptual grouping, in: *Cognitive Contributions to the Perception of Spatial and Temporal Events*, G. Aschersleben, T. Bachman and J. Müsseler (Eds), pp. 347–362. Elsevier, Amsterdam.

Bertelson, P., Frissen, I., Vroomen, J. and de Gelder, B. (2006). The aftereffects of ventriloquism: patterns of spatial generalization, *Percept. Psychophys.* **68**, 428–436.

Bertelson, P. and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory–visual spatial discordance, *Percept. Psychophys.* **29**, 578–584.

Bertelson, P., Vroomen, J. and de Gelder, B. (2003). Visual recalibration of auditory speech identification: a McGurk aftereffect, *Psychol. Sci.* **14**, 592–597.

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I. and King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex, *Cereb. Cortex* **17**, 2172–2189.

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H. J. and Hillyard, S. A. (2007). Neural basis of the ventriloquist illusion, *Curr. Biol.* **17**, 1697–1703.

Budinger, E., Heil, P., Hess, A. and Scheich, H. (2006). Multisensory processing *via* early cortical stages: connections of the primary auditory cortical field with other sensory systems, *Neuroscience* **143**, 1065–1083.

Burge, J., Ernst, M. O. and Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching, *J. Vision* **8**, Article 20.

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D. and David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding, *Neuroreport* **10**, 2619–2623.

Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation, *J. Exper. Psychol.* **84**, 141–147.

Canon, L. K. (1971). Directed attention and maladaptive 'adaptation' to displacement of the visual field, *J. Exper. Psychol.* **88**, 403–408.

Eramudugolla, R., Kamke, M. R., Soto-Faraco, S. and Mattingley, J. B. (2011). Perceptual load influences auditory space perception in the ventriloquist aftereffect, *Cognition* **118**, 65–77.

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration, in: *Perception of the Human Body*, G. Knoblich, I. M. Thornton, M. Grosjean and M. Shiffrar (Eds), pp. 105–131. Oxford University Press, New York, NY.

Frissen, I., Vroomen, J., de Gelder, B. and Bertelson, P. (2003). The aftereffects of ventriloquism: are they sound-frequency specific?, *Acta Psychol.* **113**, 315–327.

Frissen, I., Vroomen, J., de Gelder, B. and Bertelson, P. (2005). The aftereffects of ventriloquism: generalization across sound-frequencies, *Acta Psychol.* **118**, 93–100.

Giard, M. H. and Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study, *J. Cognit. Neurosci.* **11**, 473–490.

Held, R. (1965). Plasticity in sensory-motor systems, *Sci. Amer.* **213**, 84–94.

Kayser, C., Logothetis, N. K. and Panzeri, S. (2010). Visual enhancement of the information representation in auditory cortex, *Curr. Biol.* **20**, 19–24.

Klemm, O. (1909). Localisation von Sinnineidrucken bei disparaten Nebenreizen (Localization of sensory impressions with disparate distractors), *Psychol. Stud./Psychol. Res. (Wundt)* **5**, 73–161.

Kopco, N., Lin, I. F., Shinn-Cunningham, B. G. and Groh, J. M. (2009). Reference frame of the ventriloquism aftereffect, *J. Neurosci.* **29**, 13809–13814.

Lewald, J. (2002). Rapid adaptation to auditory–visual spatial disparity, *Learn. Mem.* **9**, 268–278.

Lewald, J. and Ehrenstein, W. H. (2001). Effect of gaze direction on sound localization in rear space, *Neurosci. Res.* **39**, 253–257.

McCollough, M. C. (1965). Color adaptation of edge detectors in the human visual system, *Science* **149**, 1115–1116.

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices, *Nature* **264**, 746–748.

Middlebrooks, J. C. and Green, D. M. (1991). Sound localization by human listeners, *Ann. Rev. Psychol.* **42**, 135–159.

Passamonti, C., Frissen, I. and Ladavas, E. (2009). Visual recalibration of auditory spatial perception: two separate neural circuits for perceptual learning, *Eur. J. Neurosci.* **30**, 1141–1150.

Radeau, M. (1973). The locus of adaptation to auditory–visual conflict, *Perception* **2**, 327–332.

Radeau, M. and Bertelson, P. (1974). The after-effects of ventriloquism, *Qtly J. Exper. Psychol.* **26**, 63–71.

Radeau, M. and Bertelson, P. (1976). The effect of a textured visual field on modality dominance in a ventriloquism situation, *Percept. Psychophys.* **20**, 227–235.

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect, *Proc. Nat. Acad. Sci. USA* **95**, 869–875.

Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time, *Hear. Res.* **258**, 89–99.

van Linden, S., Vroomen, J., de Gelder, B. and Bertelson, P. (2004). Visual recalibration of auditory speech *versus* selective speech adaptation: different build-up courses, in: *Proc. 8th Int. Conf. Spoken Language Processing (ICSLP)*, Jeju, Jeju Island, Korea.

Vroomen, J., van Linden, S., Keetels, M., de Gelder, B. and Bertelson, P. (2004). Selective adaptation and recalibration of auditory speech information by lipread information: dissipation, *Speech Commun.* **44**, 55–61.

Welch, R. B. (1978). *Perceptual Modification: Adapting to Altered Sensory Environments*. Academic Press, New York, NY.

Wozny, D. R. and Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal discrepancy, *J. Neurosci.* **31**, 4607–4612.