Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Acta Psychologica xxx (2004) xxx–xxx
acta psychologica
www.elsevier.com/locate/actpsy

The aftereffects of ventriloquism: Generalization across sound-frequencies

Ilja Frissen [a], Jean Vroomen [a], Béatrice de Gelder [a,b], Paul Bertelson [b,a,*]

[a] *Tilburg University, Tilburg, The Netherlands*
[b] *Laboratoire de Psychologie Expérimentale, Université Libre de Bruxelles, CP 191, 50 Av. F.D. Roosevelt, B-1050, Brussels, Belgium*

Abstract

Exposure to synchronous but spatially discordant auditory and visual inputs produces, beyond immediate cross-modal biases, adaptive recalibrations of the respective localization processes that manifest themselves in aftereffects. Such recalibrations probably play an important role in maintaining the coherence of spatial representations across the various spatial senses. The present study is part of a research program focused on the way recalibrations generalize to stimulus values different from those used for adaptation. Considering the case of sound frequency, we recently found that, in contradiction with an earlier report, auditory aftereffects generalize nearly entirely across two octaves. In this new experiment, participants were adapted to an 18° auditory–visual discordance with either 400 or 6400 Hz tones, and their subsequent sound localization was tested across this whole four-octave frequency range. Substantial aftereffects, decreasing significantly with increasing difference between test and adapter frequency, were obtained at all combinations of adapter and test frequency. Implications of these results concerning the functional site at which visual recalibration of auditory localization might take place are discussed.

* Corresponding author. Tel.: +32 2 772 85 81; fax: +32 2 650 22 09.
E-mail address: pbrtlsn@ulb.ac.be (P. Bertelson).

0001-6918/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2004.10.004

PsycINFO classification: 2320
Keywords: Sensory perception; Intersensory processes; Perceptual adaptation; Sound localization

1. Introduction

A good deal of the currently very active movement of research into multisensory perception has been focused on the interactions between localization processes in separate sense modalities, like proprioception and vision or audition and vision. It has resorted principally to conflict situations, in which two separate modalities receive incongruent inputs regarding the spatial location of one particular external event (see Bertelson & de Gelder, 2004; de Gelder & Bertelson, 2003, for recent reviews).

The auditory–visual conflict situation, in which brief sounds and visual stimuli are presented simultaneously in slightly separate locations has proved especially convenient for experimental study, because of the control it allows on the main input parameters (Bertelson, 1999). The processes put into play by this conflict have been called *ventriloquism*, because one of their most spectacular manifestations is the illusion created by performing ventriloquists that the speech they produce without visible lip movements comes from a simultaneously agitated puppet.

Work on ventriloquism has concentrated on two main behavioral consequences of exposure to the situation. One is that the apparent location of the sounds is shifted toward the simultaneous visual inputs, in spite of instructions to ignore the latter (e.g., Bertelson & Radeau, 1981). This on-line, or immediate, effect (i.e., observed in presence of the conflict situation) has been called the *visual bias of apparent auditory location*. The other consequence, generally designated as *aftereffects*, can be observed off-line, after somewhat prolonged exposure to similar incongruent auditory–visual stimulus pairs, when localization responses to singly presented stimuli in one of the modalities are displaced in the direction in which the conflicting stimuli in the other modality were during exposure (Canon, 1970; Radeau & Bertelson, 1974, 1976). The fact that exposure to intermodal conflicts produces aftereffects has been taken as showing a degree of plasticity, by which the different localization processes recalibrate input-to-percept matches in such a way as to eliminate (or at least reduce) registered incongruences. Such recalibration can be expected to play an adaptive role in the development and later maintenance of cross-modal coordination (Held, 1965; Redding & Wallace, 1997; Welch, 1978).

The present study is part of a project stemming from the notion that aftereffects offer opportunities, not available in immediate effects, for determining the *extension* of the changes induced by conflict exposure. Measuring aftereffects at several values of the target stimulus can indicate whether these changes are specific to the values used during adaptation, or rather affect a range of neighboring values. This research strategy was inaugurated by Bedford (1989), who examined the spatial extension of visuo-proprioceptive recalibration in finger placing under displaced visual feedback, and found that recalibration achieved at a particular location generalized entirely to other locations along the azimuth.

Our project focuses on the case of ventriloquism. One line (mainly unpublished so far, except for Vroomen, Bertelson, Frissen, & de Gelder (2001)) is concerned with spatial extension. The other, to which the present paper belongs, deals with extension along the dimension of sound frequency.

In a first study, we examined how specific auditory recalibration is to the frequency of the sounds used during adaptation (Frissen, Vroomen, de Gelder, & Bertelson, 2003). In three separate experiments, participants were exposed to pure tones at either 750 or 3000 Hz synchronized with light flashes 9° to their left or right, and unimodal localization of the same tones was measured pre- and post-exposure. In each of the three experiments, the aftereffects generalized quasi totally (with only small non-significant reductions) across the two-octave distance. Forcing attention during exposure either to the visual or to the auditory modality in no way affected the results.

These results are in sharp contrast with ones reported earlier by Recanzone (1998), in which no generalization whatsoever occurred across the same two frequencies. These data, however, were based on a single direction of adaptation for just three participants. Large individual differences are not uncommon in intersensory recalibration (Redding & Wallace, 1997), and we mentioned that several of our own participants showed similarly no generalization. We suggested that the apparent contradiction might be explained by the small scope of Recanzone's investigation.

One important reason for being interested in the extension across sound frequencies, which was already discussed by Frissen et al. (2003), was the information it could provide about the site in the functional architecture at which recalibration takes place. It is well-known that of the two major sound localization mechanisms, the one based on interaural time differences (ITD) is sensitive mainly to low frequencies and the other one, based on interaural intensity differences (ILD) is sensitive mainly to high frequencies (Blauert, 1997; Cohen & Knudsen, 1999). [1] Finding (as did Recanzone) that adaptation does not generalize between a low tone (in this case 750 Hz) and a high one (3000 Hz) would have suggested that recalibration occurs at the level of the just mentioned two peripheral processes, while the opposite finding would point to a more central site, reached after the outputs from these processes have been integrated. Frissen et al.'s (2003) results of course supported the second possibility.

The theoretical importance of the issue made it desirable to extend the exploration of the generalization pattern beyond the two particular frequencies solely considered in both the Recanzone (1998) and Frissen et al. (2003) studies. In the following experiment, generalization was examined across a four-octave range from 400 to 6400 Hz. Aftereffects were measured at one-octave intervals across that range.

There were three main motives for the extension. First, in the preceding study (Frissen et al., 2003) a tendency toward smaller aftereffects at the non-adaptation frequency was observed in all three experiments, but without reaching significance.

[1] Other cues also contribute to sound localization, most notably monaural spectral cues from the pinnae. These, however, are particularly important for localizing along the vertical axis (see Middlebrooks & Green, 1991).

Considering generalization over a wider range provided a better opportunity for any frequency specificity to manifest itself. Second, measuring aftereffects at four distances from each adaptation frequency made it possible to determine the shape of any possible gradient. Finally, it avoided possible problems resulting from the use of 3000 Hz as one of the adapting frequencies in both the Recanzone (1998) and Frissen et al. (2003) studies. There is some evidence for an overall drop in tone localization performance around that frequency, more precisely between 2000 and 4000 Hz, which has been interpreted as meaning that both the ITD and the ILD processes are deficient in that range (Stevens & Newman, 1936). It could thus be argued that 3000 Hz was not an adequate frequency for measuring the specific performance of the ILD process. No such concern exists for the 6400 Hz high frequency used in this new experiment.

2. Method

2.1. Participants

Fifteen students from Tilburg University (age 18–26, three male), with normal hearing and normal or corrected to normal vision, participated in four sessions each.

2.2. Apparatus and material

The testing was carried out in a dark soundproof and sound attenuated booth. The setup involved six display units and an array of push buttons. Display units, which were occluded by means of a black, acoustically transparent cloth, each consisted of a loudspeaker with an LED over its center. They were arranged along a semi circular array at 41 cm from the chin rest supporting the participant's head, at respectively 45°, 27°, and 9° left and right of the latter's median plane. The two most extreme locations were used only for presenting visual inputs. Pushbuttons, 108 in total, were arranged along another circular array, 5 cm in front of the display units. The auditory stimuli were five 400 ms pure tones, with 10 ms linear rise/fall, at frequencies of 400, 800, 1600, 3200, and 6400 Hz, and presented at 66 dB (A).

2.3. Procedure

The experiment was run in four counterbalanced sessions, one for each condition (adaptation to the left or to the right, with either a 400 Hz or a 6400 Hz adapter). A session consisted of one block of 160 pre-test trials, followed by eight adaptation-post-tests blocks. On both pre-tests and post-tests, a sound at one of the five frequencies was presented from one of the four loudspeakers, and the participant was instructed to press the pushbutton corresponding to the apparent direction of its source. The participant initiated the presentation by pressing a centrally located button, a procedure that ensured a constant starting position for pointing movements.

Each adaptation-post-tests block started with 60 presentations of the condition's particular adapter bimodal pair, with the sounds delivered in sets of five, equally distributed across the four central loudspeakers. The synchronous light flash was delivered on either the left or right immediately neighboring display unit (depending on the session), thus producing an 18° leftward or rightward discordance. Adaptation trials followed each other at a one-per-sec speed. They were followed by 20 post-tests trials, one with the sound at each frequency delivered on each loudspeaker, in randomized order.

3. Results

Aftereffects were calculated by subtracting mean reported locations on pre-tests from those on post-tests, and were counted as positive when they went toward the visual distracter. Mean values pooled over speaker locations and directions of adaptation are shown in Fig. 1 as functions of test frequency.

At both adaptation frequencies, aftereffects go down slightly with increasing distance between test and adaptation frequency, but substantial values (in the order of 70% of those at zero distances) were still observed at the maximum four-octave distances.

In a 2 (adapter frequency) $\times$ 5 (test frequency) $\times$ 2 (direction of adaptation) repeated measures ANOVA, none of the main effects was significant, but two interactions were. One is the adapter frequency by test frequency interaction, $F(4, 56) =$

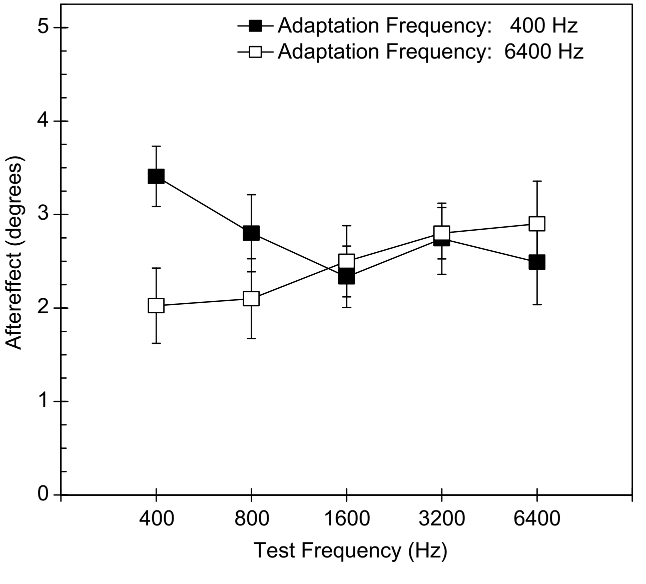

Fig. 1. Mean aftereffects (with standard errors), averaged across test locations and directions of adaptation, as functions of test frequency and with adapter frequency as parameter.

4.55, $p < .01$, which reflects the opposite trends of the two functions. The other is the interaction between test frequency and direction of adaptation, $F(4,56) = 3.97$, $p < .01$. The latter interaction was apparently putting the pooling of data across directions of adaptation into question. For that reason, t-tests were applied to mean aftereffects for the two directions for each combination of adaptation and test frequency. None of the obtained (uncorrected) p values was significant (all >.23), meaning that the pooling was effectively permissible.

This put us finally in a position to measure the statistical significance of the effect of main interest, that of the distance from adapter to test frequency. The data were submitted to a new 2 (adapter frequency) × 5 (distance) repeated measure ANOVA. Only the main effect of distance was significant, $F(4,56) = 5.00$, $p < .01$. Trend analysis of this effect produced a significant linear component, $F(1,14) = 11.71$, $p < .005$, and a non-significant interaction with adapter ($F < 1$). A quadratic trend whose existence was suggested by the curves at the largest distances was found non-significant, $F(1,14) = 3.16$, $p = .097$. Regression analysis was used to determine the slopes of the two functions. They were −.19°/octave for adaptation at 400 Hz and −.25°/octave at 6400 Hz.

Separate corrected t-test showed that at each distance aftereffects were significantly superior to zero (all p's < .005).

4. Discussion

Two main findings emerged from this experiment. First, substantial and significant aftereffects were observed across the whole four-octave range of test frequencies. This result further supports our earlier conclusion (Frissen et al., 2003) that visual recalibration of auditory location is not limited to the sound frequency presented during conflict exposure. It also re-inforces the suggestion that the no-generalization result of Recanzone (1998) was reflecting some sampling error.

Second, aftereffects were reduced significantly with increasing difference between adapter and test frequency. As mentioned in Section 1, the small effects of that difference observed in Frissen et al.'s (2003) experiments were all non-significant, leaving open the possibility of a uniform generalization across all frequencies, reminiscent of Bedford's (1989) results in her finger placing tasks. This extreme possibility presumably need not be considered any longer.

Our new, more extensive picture of generalization across frequencies has important implications regarding the possibility that the peripheral ITD and ILD mechanisms play specific roles in the visual recalibration of perceived sound location. The higher adapter frequency used in this experiment (6400 Hz) is, unlike the 3000 Hz one used by both Frissen et al. (2003) and Recanzone (1998), well beyond the range (2000–4000 Hz) in which the role of ILD has been questioned, while the lower adapter frequency (400 Hz) is clearly in the range of principal dependence on ITD. Thus, at the largest differences between test and adapter frequency that were considered here, test and adapter tones must each have been localized essentially via one of the two peripheral processes. The fact that substantial generalization was nevertheless obtained suggests strongly that recalibration takes place mainly at a more *central* site, i.e., as already mentioned, one subsequent to the integration of the data from the peripheral processes. Another aspect of the results with similar implications is the absence in the obtained generalization gradients of any discontinuity, possibly corresponding to a boundary between the domains of operation of the two processes.

The preceding conclusion regarding the processing stage at which visual recalibration takes place makes good functional sense. In earlier papers from our group, it was proposed that the main evolutionary advantage of cross-modal integration was its effect in reducing the consequences of modality-specific variability (de Gelder & Bertelson, 2003; Bertelson & de Gelder, 2004). This kind of advantage may be derived not only from cross-modal integration, but equally from intra-modal integration, of which the combination of outputs from ILD and ITD into an integrated representation of sound location (on the horizontal axis) would be a good example. Two separate integration operations, the intra-modal integration of ILD and ITD outputs within audition, and the cross-modal integration of auditory and visual data into an amodal impression of event location, would thus be involved in the phenomena at the focus of the present study. In terms of this dual integration view, the main implication of our findings is that cross-modal integration operates after, and on the combined outcomes of, intra-modal integration. This operating sequence is obviously a more economical one than having the visual data combined separately with the component outputs from ILD and ITD. Whether the same kind of sequence occurs in all cases of cross-modal recalibration is a question for future research.

Acknowledgments

Thanks are due to the Editor, Albert Postma, and two anonymous reviewers, for insightful and constructive comments on a first version of the paper. The last author's participation was partially supported by the Belgian National Fund for Collective Fundamental Research (Contract 10759/2001 to Régine Kolinsky and Paul Bertelson).

References

Bedford, F. L. (1989). Constraints on learning new mappings between perceptual dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 232–248.

Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Aschersleben, T. Bachman, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–362). Amsterdam: Elsevier Science Publishers.

Bertelson, P., & de Gelder, B. (2004). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141–177). Oxford: Oxford University Press.

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

Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. Cambridge, MA: MIT Press.

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

Cohen, Y. E., & Knudsen, E. I. (1999). Maps versus clusters: Different representations of auditory space in midbrain and forebrain. *Trends in Neurosciences, 22*, 128–135.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences, 7*, 460–467.

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

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

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42*, 135–159.

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *Quarterly Journal of Experimental Psychology, 26*, 63–71.

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

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: The ventriloquism aftereffect. *Proceedings of the National Academy of Sciences of the USA, 95*, 869–875.

Redding, G. M., & Wallace, B. (1997). *Adaptive spatial alignment*. Hillsdale, NJ: Lawrence Erlbaum.

Stevens, S. S., & Newman, E. B. (1936). The localization of actual sources of sound. *American Journal of Psychology, 48*, 297–306.

Vroomen, J., Bertelson, P., Frissen, I., & de Gelder, B. (2001). Short-term adaptation to auditory–visual discordance: The ventriloquism aftereffect. *Abstracts of the Psychonomic Society, 6*, 31–32.

Welch, R. B. (1978). *Perceptual modification: Adapting to altered sensory environments*. New York: Academic Press.