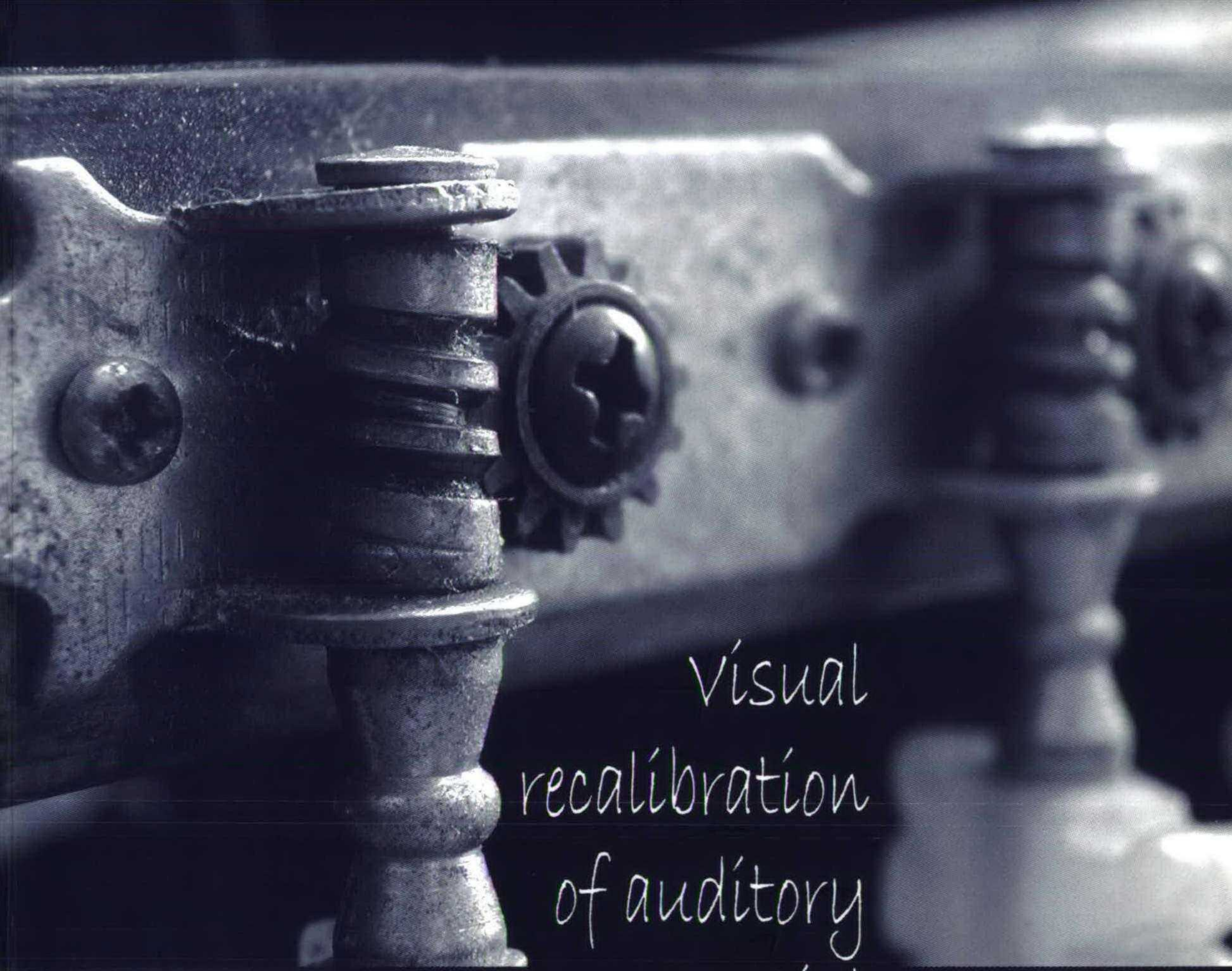
Visual recalibration of auditory spatial perception

Frissen, I.H.E.

Publication date:
2005

Citation for published version (APA):
Frissen, I. H. E. (2005). *Visual recalibration of auditory spatial perception: the aftereffects of ventriloquism*. [s.n.].



Visual recalibration of auditory spatial perception

Visual recalibration of auditory spatial perception: The aftereffects of ventriloquism

Ilja Frissen

ERRATUM

There are errors in the text as it is presented before you in this book. Normally this should almost go without saying but I feel there are some which might cause confusion and impede navigating through the thesis. They concern the numbering of one paragraph and its subparts in the introduction and several references in later chapters to earlier parts.

Headers

- page 19 must be §1.2.7
- Page 20 must be §1.2.7.1
- Page 21 must be §1.2.7.2

References

- Page 22 2nd paragr. should be §1.2.1
- Page 23 1st paragr. should be §1.2.4
- Page 80 1st paragr. should be §1.2.7.1
- Page 98 1st paragr. should be §1.2.7
- Page 106 last paragr. should be §1.3

There is also an anachronism in the text which I happily admit to. A thoroughly revised version of chapter 6 has now been accepted for publication, and the footnote on page 65 should therefore now be:

> Bertelson, P., Frissen, I., Vroomen, J., & de Gelder, B. (in press). The aftereffects of ventriloquism: Patterns of spatial generalization. *Perception & Psychophysics.*

Visual recalibration of auditory spatial perception

The aftereffects of ventriloquism

Visual recalibration of auditory spatial perception
The aftereffects of ventriloquism

Proefschrift

ter verkijging van de graad van doctor aan de Universiteit van Tilburg,
op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten,

in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit op
vrijdag 17 juni 2005
om 14:15 uur

door

Ilja Hermanus Elisabeth Frissen

geboren op 25 juli 1973 te Sittard

Promotores:
Prof. dr. B.L.M.F. de Gelder
Prof. dr. P. Bertelson

Copromotor:
Dr. J.H.M. Vroomen

Contents

Chapter 1

Introduction

1.1 Introduction

Our everyday experience compels us to believe that we observe the world as it is. What we become aware of, however, is not a direct one-to-one reflection of what is out there but rather a construction of that world based on *computation* and *inference* by the brain on all kinds of sensory information collected by the several different senses (e.g., vision, hearing, touch). Basically, from the sensory data estimates are made of the properties of an external object, which are necessary for such tasks as determining its location and identity. These estimates are, however, inherently ambiguous and noisy. Noise can come from the sensory machinery itself and/or from the stimulus (Pouget, Deneve, & Duhamel, 2002). The brain can deal with the ambiguity by collecting more information from within and across the sensory modalities and with the noise by integrating redundant information (Ernst & Bülthoff, 2004).

Collecting information from *across* the sensory modalities seems a sensible strategy given that most real-life situations produce correlated sensory inputs to the different modalities. Consider the pair of senses which are the focus of the present work, the auditory and the visual system. Both process similar information from events in the external world. For instance, they both process speech. Seeing someone speak not only provides auditory but also visual speech information provided through the movements of the lips, face and body. There is ample evidence that the perceptual system profits from this informational redundancy (e.g., Campbell, Dodd, & Burnham, 1998). Another example is one of the most widely studied cases of *crossmodal* integration, that of auditory-visual *spatial* perception. Like speech, spatial information can be obtained through the auditory and the visual systems (as well as others, of course, such as proprioception). In order to gain optimally from this redundancy the representations of auditory and visual space should be coordinated (King, Doubell, & Skaliora, 2004). This coordination is presumably achieved and maintained by systematically cross-checking between the two modalities.

One influential approach to studying auditory-visual spatial perception is to create a spatial conflict between the two senses and assess how the perceptual system deals with this. The system shows a number of ways of dealing, which can be categorized as being online (i.e., immediate effects) and offline (i.e., aftereffects). What is common to all these is that the perceived location of the discordant stimuli are shifted toward each other in order to reduce the registered spatial conflict. The processes put into play by an auditory-visual spatial conflict are collectively referred to as *ventriloquism*, after the performing ventriloquist who creates the illusion that the speech they produce comes from a puppet.

The aftereffects of ventriloquism are the object of study in this thesis. Several of its aspects have already been addressed in previous research and are reviewed below. Here we are interested in a number of important but as of yet largely outstanding questions. Briefly, what is the *extent* of the changes induced by exposure to a ventriloquism situation and what is the *time course* of these changes.

In the rest of this chapter follows an introduction to the *ventriloquism effect* (§1.2), and some of the neural substrates implicated in auditory and auditory-visual spatial processing (§1.3). Finally, an overview of the thesis is given (§1.4).

1.2 Visual influences on auditory localization: the ventriloquism effect

The research presented in this thesis in concerned with the interaction between auditory and visual spatial perception. There are many instances in the literature that report a biasing influence of a visual input on auditory localization. When the two modalities are presented with spatially incongruent inputs observers typically find the sound to be closer to the visual source than when no such visual distracter is present. At the same time the visual input can be attracted towards the location of the sound. In other words the apparent locations of the sound and light move towards each other. Exposure to a ventriloquism situation has several behavioral effects, which can be categorized as being either online (i.e., immediate) or offline (i.e., aftereffects).

This section is not intended as a review of the complete ventriloquism literature. Several of these already exist, and the interested reader is referred to these (Bertelson, 1998, 1999; Bertelson & de Gelder, 2004; Welch & Warren, 1980). Rather, the immediate effects are discussed in general and only the main issues in this literature are given, supplemented by the most recent trends and findings. Since it is also the topic of the present work, a more complete coverage is given of the studies demonstrating *aftereffects of ventriloquism.*

1.2.1 The immediate effects of ventriloquism

Two online effects of exposure to a ventriloquism situation have received experimental attention. The first is *spatial fusion*, in which an observer experiences the two (spatially discordant) inputs as coming from the same location. A recent, and the most elaborate, study to date is by Godfroy, Roumes, and Dauchy (2003). They examined the spatial limits of fusion across several spatial locations and with the discrepancy in the horizontal or vertical directions. A pink noise burst and a light flash were presented in synchrony in a large number of spatial arrangements. Auditory locations were arranged in a 3 by 3 array centered around the straight-ahead position (i.e., 0°) and placed 20° apart. Visual signals could be delivered at many locations relative to the auditory one (0°, ±2.5°, ±5°, ±7.5°, ±10°, ±12.5°, ±15°, ±17.5°, and ±20°, with the plus and minus sign indicating whether the visual stimulus was to the left or right of the auditory stimulus or below or above it, respectively). The participant's task was to judge whether they originated from the same location or not. The two main findings were that the fusion area is more extended in the vertical direction (overall mean ~22°) than in the horizontal (~13°). Also, the eccentricity of the location contributes in that the (horizontal) fusion area changes as a function of the location of the auditory-visual pair along the azimuth, with more fusion occurring in the periphery.

The second online effect, and by now the standard paradigm (Bertelson & de Gelder, 2004) is that of *immediate crossmodal bias.* In a typical study (e.g., Bertelson & Radeau, 1981; Radeau &Bertelson, 1987) a spatially discrepant auditory-visual pair is presented and the participant's task is to localize (e.g., point at) the location of the auditory input while under the instruction to ignore the visual input, or vice versa (i.e., a selective localization task). The main finding is that there is a small, but reliable, shift in localization in the direction of the irrelevant distracter. Although the shift is only a proportion of the total spatial discrepancy, it is apparently enough to bring the discrepancy below detection threshold (Bertelson, 1999). The effect is stronger for auditory localization towards the visual input than for the reverse case (for some positive results of an auditory bias on visual localization, see Bertelson & Radeau, 1981; Radeau & Bertelson, 1987; Warren, Welch, & McCarthy, 1981).

There have been many experimental demonstrations of this effect (in particular that of a visual bias of auditory location). The earliest are those by Klemm, (1909), Thomas (1941), Witkin et al. (1952). Another, often cited study, is by Jackson (1953), who reports two experiments. In the first, five electric bells and torch bulb were arranged in a semicircular horizontal array (at straight ahead, and 45° and 22.5° to the left and right of that) and occluded from vision by a cloth, which allowed one to see a bell only when lit by the corresponding light bulb. An additional freely moveable bell, mounted on a rail below and behind the setup, produced the actual auditory target. First the five auditory target locations were probed ten times. The participant indicated the apparent location, while being blindfolded, by identifying the bell (a through e). Next the blindfold was removed and the participant was given the instruction that now together with the bell there was going to be a light, which could be at the location of the bell or at a different location. The task was to identify the location both of the sound and the light, and to give a confidence rating of the sound location judgment. All combinations of sound and light locations were tested twice for a total of 50 trials. The results showed that when the bell and light came from different positions, sound localization performance decreased significantly because participants reported the location of the light and not that of the sound. This effect was largest for the smallest spatial discrepancy (22.5°) and dropped off quickly for the larger ones.

A second very similar experiment was conducted using *steam kettle whistles* because according to Jackson (1953):

> [...] it was considered that a subject, who saw a puff of steam rising from what he knew to be a steam kettle whistle, and who at the same time heard a whistling sound, had much stronger evidence for supposing that the two phenomena were connected. (p. 55)

This statement reflects an idea that has persisted for some time in the older but also more recent (e.g., Goldstein, 1996) literature and can still be found today. The idea is that there needs to be a "realistic" relationship between the auditory and the visual stimuli in order for ventriloquism to occur. But what are the necessary conditions for ventriloquism to occur?

Conditions for ventriloquism

The most important determining factors for ventriloquism can be described in terms of the physical properties of the stimuli (also referred to as *structural* or *sensory factors*; Bertelson, 1999; Radeau & Bertelson, 1977; Welch & Warren, 1980), and are effective irrespective of the observer's knowledge of the situation. The single most important one is the relative timing of the auditory and visual inputs. For instance, both Thomas (1941) and Radeau and Bertelson (1987), using different rates of presentation in trains of auditory and visual stimuli, found that synchronization is a powerful determinant of bias. The role of synchronization has been reaffirmed in subsequent studies (e.g., Bertelson, Vroomen, & de Gelder, 1997; Lewald & Guski, 2003; Slutsky & Recanzone, 2001). Spatial proximity is the other important determinant of interaction (Bertelson, 1999). Temporal and spatial proximity have also been typified in the Gestalt terms *common fate* and *proximity*, respectively (e.g., Radeau, 1994).

More perceptual factors contributed too, such as stimulus saliency. A continuous sound is attracted by an intermittent light flash (Thomas, 1941). Also, Radeau (1985) reports that the relative intensities of the stimuli partially determines the visual bias on auditory location. Finally, since visual spatial acuity is highest for stimuli presented in the fovea, but decreases with increasing distance from it, the biasing capability of a visual stimulus also decreases (Hairston, Wallace, et al., 2003).

Post-perceptual factors (or *cognitive factors*, Radeau & Bertelson, 1977; Welch & Warren, 1980) may also have a role to play. Take, for instance, the "realism" of the experimental situation. When using realistic stimuli, such as steam kettle whistles (Jackson, 1953), the voice of someone speaking and the sight of the speaker's face (e.g., Witkin et al., 1952; Warren, Welch, & McCarthy, 1981), or the sight and sound of beating drums (Radeau & Bertelson, 1977, Experiment 1), the observed localization responses might originate in the observer's *knowledge of* and/or *familiarity with* the simulated situation rather than the actually perceived location. They may, therefore, have a cognitive rather than a perceptual locus (Bertelson, Vroomen, de Gelder, & Driver, 2000).

That realism is not a major determinant can be inferred from a study by Bertelson, Vroomen, Wiegeraard and de Gelder (1994). A dual task approach was taken to investigate the effects of inverting a moving face on both the McGurk effect (an auditory-visual speech interference effect; McGurk & MacDonald, 1976) and the ventriloquism effect. Inversion of the face markedly effected the McGurk effect but left the ventriloquism effect unaffected. It appears as if ventriloquism in this was only dependent on the temporal aspects of the visual display and not on whether it was a face or not. That is, it is largely independent of the identity of the visual event.

Thomas (1941) acknowledged the possible influences of "past experience factors" (p. 164) and attempted to reduce them by using meaningless stimuli (i.e., a low buzzing sound and a light flash). Despite their meaninglessness the experimental situation was still capable in producing intersensory interactions. By now there have been many demonstrations of the ventriloquism effect using meaningless stimuli, such as sound bursts and light flashes (e.g., Bertelson & Aschersleben, 1998; Bertelson & Radeau, 1981; Bertelson, Vroomen, de Gelder, & Driver, 2000b; Hairston, Wallace, Vaughan, Stein, Norris, & Schirillo, 2003; Lewald & Guski, 2003; Radeau & Bertelson, 1987; Vroomen, Bertelson, & de Gelder, 2001), strongly suggesting that particular knowledge of the stimuli is not the crucial factor (Radeau, 1992). Although a factor such as realism is not a determining factor for ventriloquism it may still have a role to play in *modulating* the eventual percept.

1.2.2 Is it perceptual?

A major question that needs to be addressed is whether the ventriloquism effect (and crossmodal effects in general) is a manifestation of genuine perceptual changes or simply the result of post-perceptual adjustments on the part of the conscious and thinking observer (Bertelson, 1998, 1999; Bertelson & de Gelder, 2004; Choe et al., 1975). Obviously the latter case is of little interest to the student of perception, and the concern is a real one.

In many of the earlier studies the experimental situation was rather transparent to the observer. Most of the relevant experimental parameters, such as stimulus source location and semantic context, were open to conscious inspection, meaning that the speakers producing the sound were visible and that stimuli were often familiar objects that produce characteristic sounds, such as steam whistles (Jackson, 1953) or door bells (Canon, 1970). There is, however, an increasing amount of evidence that the ventriloquism effect is perceptual. The most important of which are briefly mentioned here.

It can simply not be denied that, in spite of the concerns, the effect is nothing but very compelling. Most participants are simply not aware of the discrepancy. But there is also experimental evidence that helps to argue in favor of a perceptual locus of the effect. Some evidence, although admittedly the weakest, comes from studies showing significant effects even though participants were explicitly instructed to ignore the visual distracter and to concentrate on the (auditory) localization task. Obviously the instruction alone does not guarantee that participants adhered to it. It therefore still leaves room for responses strategies.

Bertelson and Aschersleben (1998) countered this problem by adapting the psychophysical staircase method to the measurement of crossmodal bias. Sounds were presented either from the left or the right of the median plane and the participants indicates the laterality by pressing one of two buttons. The apparent location of the sound is controlled through two randomly chosen psychophysical staircases (one starting from the left and one from the right) which eventually converge. For instance, when a sound is presented from the right and is correctly located as such it is move to the left (i.e., towards the median plane) by one step (i.e., in this case it becomes slightly more difficult to localize it as being from the right). At some point the participant is no longer certain and the sound is localized incorrectly. The sound location then is moved back to the right (i.e., a reversal has occurred) until it is localized correctly again (and thus another reversal occurs), etc. What is crucial here is that there are two converging and randomly chosen staircases and that at some point during the exploration it is no longer obvious to any keen observer which of the two is being tested. In the end this method prevents participants from using any response strategies and in stead forces them to rely on their perception of the stimuli. In applying this to the measurement of crossmodal bias, the critical manipulation is the addition of a centrally presented light flash that is in synchrony with the auditory token. If there is an attraction of the sound location by the light flash then the points of uncertainty, and therefore the reversals, should occur at location further from the center than when there is no flash. This is exactly what Bertelson and Aschersleben found. Since there could be no reliance on any response strategies this constitutes strong evidence for a genuine perceptual component in the ventriloquism effect. (Recently, Caclin, Soto-Faraco, Kingstone, and Spence (2002) used a very similar methodology to demonstrate a tactile bias on auditory location.)

Another strong argument in favor of a perceptual locus of ventriloquism is the fact that the perceptual system is able to *adapt* to auditory-visual spatial discrepancy. Such an adaptation results in compensatory *aftereffects*, which are generally considered evidence for changes in perceptual processing (e.g., Vroomen & de Gelder, 2004a). Aftereffects are the topic of this thesis and are dealt with in more detail in a subsequent section (§1.2.6).

1.2.3 Modality dominance

Cue substitution

A simple early notion of ventriloquism was that the apparent sound location is substituted by that of the visual location. This implies total dominance of the visual modality, which has been reported for the visuo-proprioceptive (Hay, Pick, & Ikeda, 1965) and in the auditory-visual case (Pick, Warren, & Hay, 1969). Note that such a cue substitution does not require the postulation of any interactive or integrative processing of the two sensory signals. However, subsequent work never could replicate the total dominance result but instead found only partial bias (in the order of 30% of the imposed discrepancy).

The modality precision hypothesis

Another proposed explanation of the ventriloquist effect (and other cases of intersensory bias) is that *attention* is directed to the more precise modality, which in this case is the visual one. This in turn increases the weight given to the visual input (e.g., Howard & Templeton, 1966).

The modality appropriateness hypothesis

One very influential notion is that the dominance direction is determined by the *appropriateness* of the modality for the task at hand (the "modality appropriateness hypothesis" (MAH; Welch, 1999). The hypothesis thus states that because vision is more acute in the spatial domain it will bias the less acute spatial hearing. This then of course is the basis for the ventriloquism effect, but also for a phenomenon such as the visual capture of proprioception (Rock & Victor, 1964).

Under MAH, however, the reverse is expected when the task requires processing in the temporal domain. Because the temporal resolution of hearing is more acute that that of vision, whenever there is a temporal conflict between the two senses it will be resolved in the favor of hearing. Recently a number of articles have been published under the title of *temporal ventriloquism* which show just that (Aschersleben & Bertelson, 2003; Bertelson & Aschersleben, 2003; Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Vroomen & de Gelder, 2004b). The term generally refers to an auditory bias on the apparent *temporal occurrence* of a visual event. For instance, visual temporal order judgment (participants judge which one of two lights came on first) can improve due to strategically presented auditory inputs (Morein-Zamir et al.). That is, one sound presented somewhat before and the other after the onsets of the lights apparently pulls or 'ventriloquizes' the two lights apart in time, making the judgment of their relative onset easier. Another type of demonstration comes from Vroomen and de Gelder (2004b, see also Fendrich & Corballis, 2001) who report an auditory effect on a particular visual illusion, the Flash-Lag effect (FLE; Nijhawan, 1994). In the FLE a flash appears to lag behind a moving stimulus even though they are presented in the same physical location. A sound at a certain time relative to the flash can shift the temporal occurrence of that flash apparently shifts toward the occurrence of that sound. When, for instance, a sound was presented 100 ms before the flash, the latter appeared sooner in time and therefore closer in space to the moving stimulus. Consequently, the FLE is reduced.

It has now also been shown that the perceptual system is also able to *adapt* to auditory-visual temporal asynchronies (Recanzone, 2003; Vroomen, Keetels, de Gelder, and Bertelson, 2004; Fujisaki, Shimojo, Kashino, & Nishida, 2004).

Shifts in visual localization and cue reliability

One problem for MAH are several reports (already mentioned in §1.2.1) of significant auditory bias on visual localization (Bertelson & Radeau, 1981; Radeau & Bertelson, 1987; Warren, Welch, & McCarthy, 1981) and of visual aftereffects (Canon, 1970; Radeau, 1973, 1974; Radeau & Bertelson, 1969, 1974, 1976; Lewald, 2002).

Another problem for MAH is that it suggests that the biasing capacity is determined by an *intrinsic* quality of the particular modality (e.g., it is a fact that hearing is more acute than vision) and that therefore the *direction* of the dominance relation is more or less fixed.

Recent work has shown this not to be the case. It seems that the *reliability* of the perceptual estimate of the stimuli themselves is very important in determining direction. Visual capture (the strong bias of vision on the haptic perception), for instance, can be reversed into haptic-capture when the visual information is made less reliable by adding noise to the visual signal (Ernst & Banks, 2002). This strong effect of stimulus reliability results from the way the perceptual system integrates information from across the senses. It apparently does so by using maximum-likelihood estimation to combine the different inputs. Ernst and Banks demonstrated this by measuring the variances associated with the visual and haptic estimations of height. They then used these to construct a maximum-likelihood integrator, which turned out to behave very similar to humans in a visuo-haptic task. Thus visual dominance occurs when the variance in the visual estimation is lower than that for the haptic estimation.

Can the visual bias of auditory location be reversed into an auditory bias of visual location by making the visual signal more noisy? Two recent studies, also using the Bayesian integration logic, show that it can (Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003). The experimental paradigm of both studies was basically the same. They used a two-interval, two-alternative forced-choice (2I-2AFC) paradigm. Two stimuli (auditory only, visual only, or auditory visual) were presented consecutively, with a "standard" always from the central location and a comparison from one of several horizontal positions. The participant's task was to judge which one of the two was the most to the left. Visual stimuli were either blurred (Alais & Burr) or made noisy (Battaglia et al.) to various degrees, in order to manipulate their reliability. The results show that in both cases the visual dominance over auditory localization decreases to zero (Battaglia et al.) or even reverses (Alais & Burr) at the lowest levels of reliability.

Finally, adding a spatially coincident auditory stimulus can also *improve* visual localization (Hairston, Laurienti, Mishra, Burdette, and Wallace, 2003). Whereas such an addition has already been shown to be favorable to, for instance, detection (e.g., Frassinetti, Bolognini, & Ladavas, 2002), and orientation behavior (Stein, Meredith, Huneycutt, & McDade, 1989), it has proved difficult to show that it also enhances visual localization simply because of its high acuity. To cure this, participants wore low-plus lenses that rendered them effectively myopic, decreasing acuity and hence, reliability. Under these circumstances visual localization improved (i.e., variability in localization decreased) significantly by virtue of the presence of an auditory stimulus.

1.2.4 The ventriloquist with dynamic stimuli

Another trend is to look at integrative processes using dynamic, or moving, stimuli (see Soto-Faraco, Kingstone, & Spence, 2003 and Soto-Faraco & Kingstone, 2004, for reviews).

Soto-Faraco, Spence, and Kingstone (2004) report an effect they called "dynamic capture". Participants were to judge the direction of an apparent motion displays in, for instance, the auditory modality while at the same time there was a secondary visual apparent motion display. The direction of the secondary visual motion was either the same as (congruent) or opposite (incongruent) that of the auditory motion. From the results it was clear that there is a strong congruency effect, for incongruent visual motion the judgment of auditory motion dropped to chance level. This was taken to mean the visual motion can affect the perceived direction of the auditory motion, in a capture like fashion. That is, the sounds were perceived to move in the direction of the incongruent visual motion.

A more elaborate demonstration of a visual influence on auditory motion perception was reported by Vroomen and de Gelder (2003). They demonstrated that the contingent auditory motion aftereffect (CAMA; Dong, Swindale, & Cynader, 1999), an auditory analog of the visual contingent color aftereffect (McCollough, 1965), is markedly influenced by visual motion information. In the CAMA observers adapt for 10 minutes to rightward-moving sound with a falling pitch alternated with a leftward-moving sound with a rising pitch. After exposure a stationary sound with a rising pitch is perceived as moving rightward, whereas at the same time a stationary sound with a falling pitch is perceived as moving leftward (Dong et al.). Vroomen and de Gelder added visual motion to the CAMA paradigm. The critical condition was one where the visual motion was opposite the auditory motion. The results showed that the CAMA changed according to the *visual* and not the auditory motion, and that the auditory motion aftereffect was effectively cancelled by the incongruent visual motion.

1.2.5 Is it pre-attentive?

The ventriloquism effect is thought to be a genuinely perceptual effect and not the product of any post-perceptual processes (although their influence cannot be denied of course). A largely unanswered question, however, is whether there is a role for the direction of attention, which is neither perceptual nor cognitive (Bertelson et al., 2000b). Several studies have addressed this issue and all seem to indicate that attention is not necessary for obtaining the ventriloquist effect and that ventriloquism in fact *precedes* attentive processes (i.e., is pre-attentive).

Two studies have shown that the effect is independent of the direction of visual spatial attention. The first examined the role of *endogenous* visual spatial attention (Bertelson et al., 2000b). The participants pointed at the apparent location of a sound while ignoring a concurrent visual distracter that was randomly presented on either the left or the right site of a computer screen. Spatial attention was manipulated by having the participant monitor for occasional small changes. In one condition this change occurred in the central fixation location and in another condition within the visual distracters themselves. If attention were to play a role a larger bias was expected in the latter condition. The results, however, clearly showed that, although bias was obtained, it did not differ as a function of attentional condition. The second study examined the role of *exogenous* visual spatial attention (Vroomen, Bertelson, & de Gelder, 2001b). Sound localization now was assessed using the already mentioned psychophysical staircase method by Bertelson and Aschersleben (1998). Exogenous spatial attention was manipulated by means of "singletons". Visual spatial attention is attracted to that item in an array of items that is different by a particular feature (Treisman & Gelade, 1980), such as its color. In this case the singleton was defined by its size, it was smaller than the rest of the items in the display. The results showed that sound localization shifted *away* from the singleton. That is, in the direction opposite that expected when exogenous spatial attention is a determining factor. Experimental controls showed that the singleton was however effective in attracting spatial attention.

Another useful approach in obtaining information regarding the role of attention is by studying ventriloquism in patients with unilateral visual neglect, a syndrome caused by local brain damage, which involves a reduced capacity to report visual stimuli in the visual hemifield contralateral to the lesion (Bisiach & Vallar, 1988). Could a visual stimulus in this "neglected" field still bias auditory localization? Bertelson, Pavani, Ladavas, Vroomen, and de Gelder (2000) examined the ventriloquist effect in left hemifield neglect patients using two experimental tasks, to describe the visual display and to point at the apparent location of the sound. Sounds came from the left, centre, or right. Four different visual conditions were used, no visual stimulus (control condition), a single distracter in the right (non-affected) or left (neglected side) visual field, or two simultaneous squares in both the left and right field. The important point to make is that patients showed a leftward shift in sound localization when a single visual distracter was presented in the neglected left visual field. In other words, an *undetected* visual stimulus can still bias sound localization, although the absolute size of the bias was smaller than in healthy controls.

Two psychophysical studies showed that exogenous auditory spatial attention can be drawn to the illusory location of a ventriloquized sound (Spence & Driver, 2000; Vroomen, Bertelson, de Gelder, 2001a), strongly suggesting that ventriloquism occurred *before* spatial attentional processes come into play. In the study by Vroomen et al. the task was to judge the elevation of an auditory target that was delivered in the left or the right periphery. This lateral position then was irrelevant for the task itself. Before the auditory targets was presented there was either an auditory, visual, or auditory-visual cue to that side. The auditory-visual cue consisted in a tone in the straight ahead location synchronized with a light flash in the periphery, thus creating a ventriloquism situation. Whereas the visual cue had no facilitatory effect (as measured with response times), the auditory and auditory-visual cues did. In the latter case, the effect presumably resulted from the attraction of the apparent location of the tone towards the flash.

In line with this are two psychophysiological studies using electroencephalography (Colin, Radeau, Soquet, Dachy, & Deltenre, 2002; Stekelenburg, Vroomen, & de Gelder, 2004), and more in particular the "mismatch negativity" (MMN, Näätänen, 1992) paradigm. The MMN is thought to reflect pre-attentive processes and is characterized by a change in the EEG brought about by an occasional deviant in an otherwise homogenous sequence of auditory stimuli. It is, for instance, elicited by a change in the physical location of a sound source (e.g., Paavilainen, Karlsson, Reinikainen, & Näätänen, 1989). Stekelenburg et al. found that an *illusory* sound location change, by means of the ventriloquism effect, also elicits an MMN that closely resembles the MMN evoked by an *actual* location change. Colin et al., in principle, showed the same result, except in a reverse manner, by *preventing* an auditory location change through the use of the ventriloquism effect and consequently eliminating the location change MMN.

Taken together it can be concluded that the visual bias of auditory location takes place at a stage *before* attentional selection that is concerned with spatial scene analysis (Vroomen et al, 2001a).

1.2.6 Recalibration and aftereffects

The present thesis deals exclusively with the aftereffects of ventriloquism. These can be observed off-line after somewhat prolonged exposure to a ventriloquism situation, and consist in post-exposure shifts in auditory localization (Canon, 1970, 1971; Kalil & Freedman, 1967; Lewald, 2002; Radeau & Bertelson, 1969, 1974, 1976, 1977, 1978, Radeau, 1973, 1974, 1992; Recanzone, 1998), and in visual localization (Canon, 1970; Radeau, 1973, Radeau & Bertelson, 1974, 1976; see Lewald, 2002 for a partial result).

The shifts are compensatory in that they effectively reduce the registered intersensory discrepancy. Thus, sound localization shifts toward the visual input and visual localization shifts toward the auditory input. The aftereffects are generally considered to be consequences of a *recalibration* of input-to-percept matches (Held, 1965; Welch, 1978), which serves the maintenance of the coordination between the auditory and visual spatial senses. Causes for disturbances in this coordination are developmental in nature (Held, 1965), such as head growth, but also post developmental, such as sensory drift and noise.

As a dependent measure, aftereffects are also generally considered to be a better index of perceptual processes than bias. Aftereffects are measured by taking the difference in responses on *unimodal* (i.e., either purely auditory or visual) a localization task, before and after exposure to an auditory-visual spatial discordance. Since in these *unimodal* localization tests the visual stimulus is *not* present it can also not influence the response system. Thus any change in localization is ascribed to changes in perceptual processing of the stimulus.

One factor long held to be crucial for recalibration to occur is *reafference* (Held, 1961), from which it followed that the main condition for the occurrence of recalibration was exposure to rearranged reafferent stimulation. Reafferent in this case meaning self-produced arm movements (while observing it through laterally displacing prisms). Held & Hein (1958), for instance, found significant recalibration only when the participant produced the arm movement, but none when the experimenter moved the participant's arm (i.e., passive movement). An obvious way of testing the hypothesis is to test whether recalibration can be obtained using completely exafferent stimuli.

The study by Radeau and Bertelson (1974) is one such an instance. The experiment followed the classical pretest-exposure-posttest design. In the pre and posttest participants engaged in a unimodal (either auditory or visual) localization task, where they were required to point at the apparent location of the target. During exposure participants observed, for a total of 20 minutes (four blocks of five minutes), a constant auditory-visual spatial discrepancy (the visual stimulus was offset to the right or to the left by 15° by means of a wedge prism). Aftereffect were calculated by taking the difference in localization before and after exposure and counted as positive when the shift is in the direction of the of the location of the other modality (i.e., when sound localization shifts in the direction previously occupied by the visual input, and vice versa).

In this case aftereffects were about 4° for auditory localization, and 1.7° for visual localization, both significant. The experiment then is a clear demonstration that recalibration can be obtained without reafference. Besides being of theoretical importance aftereffects of ventriloquism are of course testament to the plasticity of auditory and visual spatial perception.

If not reafference, what then is the mechanism underlying recalibration? Wallach (1968) proposed recalibration is based on "*informational discrepancy*" (see Epstein, 1975, for his similar concept of "*recalibration by pairing*"). Wallach formed the concept based on his work on vision. Visual depth is determined by means of a number of sensory cues, such as eye convergence and divergence, retinal disparity, and several pictorial cues (Coren, Ward, & Enns, 1994). Cues that determine the same perceptual parameter are referred to as "paired cues", which under normal conditions are consistent with each other. Paired cues can be made to be inconsistent by imposing an artificial distortion on one of them, making them "produce different values for their common parameter" (Wallach, 1968, p. 210). The visual system resolves the discrepancy by recalibrating one or both of the cues.

Table 1.1. *Overview of aftereffect studies in humans (in alphabetical order).*

	Author(s)	Exposure Trials**	Exposure Duration	Discrepancy (°)	Mean AE (°)	Mean AE (%)***
1	Bermant & Welch, 1976	18	-	10, 20, 30	ns	-
2	Canon, 1970	-	20 min	11	2.25	20
3	Canon, 1971	-	10 min	17	7.16	42
4	Frissen et al, 2003*	2400	20 min	9	1.87	21
5	Frissen et al, 2005	8 x 60	8 x 1 min	18	2.50	15
6	Held, 1955	Continuous		22	10.0	45
7	Kalil & Freedman, 1967	-	15 min	15	3.00	20
8	Lewald, 2002	1800	17 min	20	3.76	19
9	Radeau, 1973	4 x 120	4 x 5 min	15	3.46	23
10	Radeau, 1992	210	1.75 min	15	2.10	14
11	Radeau & Bertelson, 1969	90		11	2.20	20
12	Radeau & Bertelson, 1974	4 x 120	4 x 5 min	15	4.11	27
13	Radeau & Bertelson, 1976			15	2.07	14
14	Radeau & Bertelson, 1977	Cont. video	12 x 1 min	20	2.68	13
15	Radeau & Bertelson, 1978	Cont. video	12 x 1 min	20	2.82	14
16	Recanzone, 1998	2500	20 to 30 min	8	7.08	89
17	Zwiers et al, 2003	Continuous	2 to 3 days	-	-	-
				Mean	**3.80**	**26.4**
				Median	**2.82**	**20.0**

* Experiments 2 and 3 (Chapter 2)

** 'x' denotes that exposure was distributed over a number of blocks, with a subset of post tests interspersed. The first digit indicates the number of such blocks and the second the number of trials per block. The same goes for Exposure duration except that the second digit stands for the number of minutes the block lasted.

*** Mean aftereffect as the proportion of the imposed discrepancy.

Canon (1970, 1971) proposed a model of adaptation to auditory-visual spatial conflict consisting in two main parts. The first is that in order for adaptation to occur there needs to be "intermodality inconsistency of input". This is basically Wallach's (1968) concept of informational discrepancy. The second is that the direction of attention determines the locus of adaptation. The modality that is attended to during exposure does not yield to adaptation, or conversely, adaptation (and hence aftereffects) occurs only in the non-attended modality. A similar directing role of attention has been argued for by Kelso, Cook, Olson, and Epstein (1975), for the case of visuo-proprioceptive spatial conflict.

Evidence in favor of the model, and in particular the directing role of attention, comes from two studies conducted by Canon himself (1970, 1971). The basic design of both was again the classic pretest-adaptation-posttest. On pre and posttests participants pointed at a number of targets scattered on the azimuth by means of a pivot pointer. Two versions of pre and posttest were run with the difference being whether the stimulus displacing devices were in place or not. Thus the experiments consisted in five phases, pretest without devices (version I), pretest with devices (version II), adaptation, posttest with devices (II), and posttest without devices (I). The devices were a pseudophone and a prism going in opposite directions (1970), or only the prism (1971). Comparison of pre and posttest I gives an index of the level of adaptation obtained, that is, the reduction in localization error caused by the devices. Comparison of pre and posttest II gives an index of the aftereffects of this adaptation. The adaptation phase consisted in exposure to auditory-visual spatial discrepancy of 22° for 20 minutes (1970) or 16.7° for 10 minutes (1971), during which the participant actively pointed at a target. The target depended on the particular condition, two of which were common to both experiments. In the first the participants pointed at the visual stimulus. In the second there was no spatial discrepancy but the auditory and visual stimuli alternated randomly and the participant points at both. In another condition the target is the auditory stimulus (Canon, 1970). The main result that arose from both experiments were more or less as predicted by the model. When, during exposure, attention was on the visual stimulus, adaptation and aftereffects were significant only for auditory localization. When attention was on the auditory stimulus adaptation and aftereffects were apparent in *both* modalities, although they were smaller for the auditory one. This is not completely in line with the model since it predicts no effects in the attended modality (in this case the auditory one), but clearly there are.

1.2.5 Review of the aftereffect literature

Here we review the few published studies (summarized in table 1.1) on the aftereffects of ventriloquism and related ones, except for the ones by Frissen et al, which are dealt with in the present thesis, and the studies by Canon (1970, 1971) and Radeau and Bertelson (1974), which have already been discussed in the previous section.

Before addressing the individual studies, consider the general picture given by table 1.1, which present the overall aftereffects obtained. A number of observations can be made. First, the mean absolute magnitude of aftereffects is 3.8° (sd = 2.4°), ranging from 1.9° to as large as 10°. In relative terms, aftereffects are approximately a quarter of the imposed discrepancy, which is comparable the case of visuo-proprioception recalibration (Welch, 1978, 1986). Second, compensation is never complete. Recall that recalibration serves the maintenance of coordination between the auditory and visual spatial senses. The experimentally created discrepancies are very likely much larger than those created by drift and noise which the perceptual system typically has to deal with. The incomplete compensation then is due to the limits of recalibration. Third, there is only one study that was not able to obtain significant aftereffects (i.e., Bermant & Welch, 1976). The reason for this is relatively straightforward, the exposure period was to short.

The rest of the review is thematic and consists in two major parts. First, what have been the methodological issues involved in running an aftereffects experiment? Second, what was the theoretical impetus for the particular experiments and what were the results?

1.2.4.1 Methodological considerations and concerns in studying the aftereffects of ventriloquism

There are a number of methodological considerations and concerns in running an experiment on the aftereffects of ventriloquism worth pointing out. The first is how to measure sound localization performance. The second is how to present the auditory and visual inputs. Finally, what should the participant do during the exposure phase?

By far the most common method for measuring sound localization performance is having participants simply point with their hand at the apparent location of the sound source (Canon, 1970, 1971; Radeau, 1973, 1974, 1992; Radeau & Bertelson, 1969, 1974, 1976, 1977, 1978). It is also the preferred method in the present work. Similarly, swivel pointers, a metal rod that could be rotated in the horizontal plane, have been used (Canon, 1970, 1971; Lewald, 2002). Participants in Recanzone's (1998) study were required to point their head in the direction of the sound source. Besides these "absolute" methods a small number of studies have employed "relative" localization methods. For instance, Kalil and Freedman (1967) and Radeau (1973) used the subjective auditory straight ahead. Participants manipulated the direction of a sound source until they judged it to be in the straight-ahead, or median, position. Another method is to present a visual reference point followed by a test tone, which the participant judges to be to the left or right of the reference (e.g., Chapter 7; Recanzone, 1998). An advantage of the relative methods is that there is no directed motor response necessary on the part of the participant, and consequently there is less additional noise due to the use of an effecter system.

Spatial discrepancies between the auditory and the visual inputs have been achieved in a number of ways. One is to have observers walk around for extended periods of time with devices that rearrange either auditory (e.g., Held, 1955) or visual space (Zwiers, van Opstal, & Paige, 2001) while going about their daily business. Auditory rearrangement is readily achieved using so-called pseudophones. Young (1928), for instance, studied the effects of complete left-right reversal on auditory localization but did not find any systematic compensatory shifts. Complete reversal is most probably a too drastic, and "unnatural", change for the perceptual system to cope with, as was apparently also the case for complete inversion of the optical array (Stratton, 1897). Small *shifts* of the auditory (and visual) space, on the other hand, have been shown capable of bringing about changes in auditory localization. Probably the earliest example of this is the study by Held (1955). In the first of two experiment participants wore pseudophones, set to a 22° displacement, for seven hours. Other studies using pseudophonically created shifts are by Freedman, Wilson, & Rekosh (1967) and Mikaelian (1969). Zwiers et al. (2003) had their participants wear refracting lenses, for 2 to 3 days, that compressed the visual array. They did so by a factor 0.5 and only within a radius of 20°, the remaining peripheral visual field was masked. Vision, then, could directly affect sound localization only in a limited part of space. The results showed that sound localization indeed shifted in a manner corresponding to the visual distortion. One obvious problem with studies in which participants wear devices for days on end while going about their daily business is that of experimental control. There is no control over the multitude of sensory experiences which may contribute to the adaptation end points. For instance, it has been demonstrated that the tactile sense is capable of biasing auditory localization (Caclin et al., 2002) and that the perceptual system can adapt to auditory-tactile spatial conflict (Freedman & Wilson, 1967).

Most studies have been conducted entirely in the controlled environment of the laboratory Researchers used prismatic rearrangement of the visual array (Canon, 1970, 1971; Radeau, 1973, 1974; Radeau & Bertelson, 1969, 1974). A general disadvantage of using rearrangement through devices is the different types of extra distortions they create. Prisms, for instance, create color fringes and distort straight lines to curved lines especially toward the periphery (Welch, 1986). Spatial disparity has also been created by simply presenting auditory and the visual inputs in physically different locations. Stimuli are typically produced by means of arrays of LEDs and loudspeakers (e.g., the present thesis; Lewald, 2002; Radeau & Bertelson, 1976; Recanzone, 1998).

There is a concern with respect to the task during the exposure phase. A number of studies required participants to localize the auditory or visual inputs during exposures (e.g., Bermant & Welch, 1976; Canon, 1970, 1971; Radeau & Bertelson, 1976). A serious problem with this is that the participant may engage in motor response learning, and therefore any aftereffect need not be a reflection of a perceptual change but rather a learned stereotypical response tendency. An elegant remedy for this is Radeau and Bertelson's (1974) bimodal monitoring task. During exposure the participant monitors both the auditory and the visual inputs for occasional decreases in either of their intensities, and presses a button to indicate such an occasion. In the present work a variation of this is used, namely a unimodal monitoring task, and in no case were participants required to do any pointing during the exposure phase. In most cases the participant monitors the display for occasional omissions of the visual stimulus and in one case for a decrease in sound intensity (i.e., Experiment 3 in Chapter 2).

1.2.4.2 Theoretical impetuses and results

The theoretical impetus for several of the older studies was to test Held's (1961) reafference hypothesis by showing that adaptation could occur with strictly exafferent stimulation (Canon, 1970; Kalil & Freedman, 1967; Radeau & Bertelson, 1969, 1974). All studies found clear aftereffects ranging from 2.2° to 4.1°, thereby falsifying the reafference hypothesis.

Other studies aimed at answering more specific questions regarding the aftereffects themselves. Several studies were interested in the role of so-called sensory, or *structural factors* (Radeau & Bertelson, 1977; Welch & Warren, 1980). That is, the ones related to the physical attributes of the conflict situation. So, for instance, the aims of the study by Bermant and Welch (1976) were, first to look at the effects of different magnitudes of spatial discrepancy (ranging from 10° to 30°) on aftereffects (and immediate effects), and second, to look at the effects of eye fixation during exposure. Unfortunately, as was already mentioned, the experimental procedure was not sufficient for obtaining recalibration and these interesting questions remain unanswered.

Radeau (1973) was interested in determining the functional locus of recalibration. Theoretically, recalibration can occur at any locus between the eye and the finger and the ear and the finger, some of which are common to both sets, such as the head-body joint (or, articulation). Radeau hypothesized that, from an economical point of view, the locus is somewhere above the neck, in the eye-head and the ear-head articulation. Two methods were used to determine the size of the aftereffects. In the first, participants pointed at the location of the auditory and visual targets, before and after adaptation. In the second, participants moved the target until it appeared straight-ahead, also before and after adaptation. Under the hypothesis there should be no difference in the magnitude of the aftereffects, which indeed turned out to be the case.

Radeau and Bertelson (1977, Experiment 2) looked at the effect of desynchronization of a voice and its corresponding face during exposure on recalibration. Thus, auditory and visual inputs were either synchronous in one condition or sound lagged by 350 ms in another. Desynchronization lead to a significant reduction of the size of the aftereffects, which is, of course, its effect on the immediate effects (see §1.3.1).

We have seen that on the level of immediate effects the interaction between auditory and visual localization is not completely dominated by the visual modality and that the auditory one is also capable of exerting bias. Likewise, on the level of aftereffects, shifts in visual localization have been found in several studies (Canon, 1970, Lewald, 2002; Radeau, 1973, 1974; Radeau & Bertelson, 1969, 1974, 1976). Visual shifts are generally smaller and found less systematically than the auditory ones. According to Radeau and Bertelson (1976) the failure to find visual shifts in some studies might be related to the "degree of visual articulation of the visual field" (p. 228). In those experiments were visual shifts were found, the visual input was a simple point of light in an otherwise totally dark visual field, and in those where they were not found, participants had, for instance, a full view of the apparatus as in Canon's (1970) experiment. In the latter case a rich visual framework is available to the observer, which is necessary for stabilizing visual localization (Matin, 1972). Radeau and Bertelson tested this directly by manipulating the visual background during the exposure phase. In one condition this was totally black (i.e., dark) and in another it was textured. The texture was provided by a strip of white paper, across the back of the experimental setup, with vertical black stripes of .5 cm at 4.5 cm intervals, which was dimly illuminated. As expected significant visual aftereffects (mean = 1.32°) were only found when there was no visual background. In a second experiment the presence of a textured background did not affect the size of the auditory aftereffects, which in both conditions were on average 1.66°. In both these experiments participants pointed, during exposure, at either the visual (experiment 1) or the auditory (experiment 2) input of the auditory-visual spatial conflict, which was necessary to determine the effects of a textured background on aftereffects as well as immediate effects (another aim of this particular study). In a third experiment the effects of the pointing instructions were controlled for by having participants perform a bimodal monitoring task that required no pointing whatsoever but the detection of an occasional drop in the intensity of the auditory or visual input. The results confirmed those of the other two experiments, significant visual aftereffects were only obtained in a completely dark visual field (mean = 1.57°) and not with a textured one. The auditory aftereffects were again of equal magnitude irrespective of whether there was a textured (mean = 3.32°) or a dark (mean = 3.15°) background.

Several studies have looked at the role of *cognitive factors* in adaptation. The efforts by Canon (1970, 1971) to show the importance of the direction of attention during adaptation have already been discussed (§1.3.3). Radeau and Bertelson (1974) tested, besides the reafference hypothesis, whether a priori knowledge of the spatial relation between the auditory and the visual inputs during exposure affect adaptation. Participants were told that the origin of the auditory and the visual inputs during the exposure was either the same or different. In a control condition no reference was made to the relation of the inputs. All three conditions yielded significant aftereffects in both auditory and visual localization, although aftereffects did vary with condition. Auditory shifts were equally large (and largest) in the two experimental conditions, and smallest in the control condition.

In two later studies, Radeau and Bertelson (1977, 1978) manipulated the degree of realism of the spatially discordant stimuli. The "realistic" conditions featured the sight and sound of hands playing bongos (1977 Experiment 1; 1978) or the moving face and voice of a male speaker (1977 Experiment 2). They are realistic in the sense that they simulate situations known to produce naturally correlated auditory and visual stimuli. For the non-realistic counterparts the same auditory stimuli were used. They were also used to modulate the light in the visual display, producing a complicated pattern of blots of diffuse light appearing in rhythm with the auditory stimulus. The results showed that realism, as it was operationalized here, was irrelevant to adaptation. Aftereffects in either condition were of equal magnitude.

That cognitive influences are not a necessary condition for recalibration is also clear from those studies that found significant aftereffects using meaningless stimuli such as sound bursts and light flashes. Work by Radeau and Bertelson (1969, 1974; Radeau, 1973, 1992; see also Bermant & Welch, 1976) are good examples of this. For instance, Radeau (1992), arguing for the "cognitive impenetrability" of auditory-visual interaction, pitted a situation that involved mainly sensory factors against a one that involved mainly conceptual factors. The former was basically the one used in, for instance, the Radeau and Bertelson (1974) study. Stimuli were sound bursts and light flashes and thus completely meaningless. The latter was similar to the one used by Weerts and Thurlow (1971), and consisted in again sound bursts but as visual input the (unsynchronized) illumination of a dummy loudspeaker. In a second part of the experiment the same conditions were used to test their effects on the immediate bias. The results were rather straightforward, only the sensory condition produced significant aftereffects (and bias). In short, the results provide little evidence for an influence of cognitive factors (as manipulated here) on the resolution of auditory-visual spatial discrepancy.

More recently studies have been undertaken to determine the specificity of the aftereffects. Recanzone (1998) reports on three experiments demonstrating aftereffects of ventriloquism. The first and second were replications of the basic phenomenon but with head pointing and a relative localization task, respectively. The latter consisted in the localization of a tone relative to a visual reference location. Adaptation was established by 20 to 30 minutes exposure to an auditory-visual spatial discrepancy of 8° while monitoring for occasional drops in intensity in the auditory input. As expected, both showed significant aftereffects. In the head pointing experiment, aftereffects even reach 89% of the size of the imposed discrepancy. The third experiment on the other hand allowed some new insights. Aftereffects were assessed for tones at the frequency used during exposure and at other frequencies. The rational behind this was to link aftereffects to cortical structures, such as the primary auditory cortex (AI) which are known to be involved in auditory localization and organized in a frequency specific manner. From work with macaques it was established that cells respond to 750 Hz tones but not to 3000 Hz ones, and vice versa. Finding that aftereffects of ventriloquism do not generalize would implicate AI as one of the neural substrates for recalibration. This is exactly what Recanzone found and concluded. The results, however, were based on a very small sample (only three participants) and on a single direction of discordance (to the right) only. Given the theoretical importance a replication of results seemed in order using at least a larger sample size and both directions of discordance. These have now been provided by Lewald (2002 Experiments 3 and 4). He replicated the frequency generalization results of Recanzone using a somewhat larger sample size and a somewhat different methodology. Seven participants adapted to a 20° auditory-visual spatial discrepancy for 17 minutes. The auditory stimulus was either a 1000 Hz or a 4000 Hz tone and aftereffects were measured at either the same or the other frequency, but no generalization was found.

Later work (Woods & Recanzone, 2004) also demonstrated aftereffects in macaque monkeys. This time there was evidence for substantial generalization, in the order of 45%, from adaptation with a 4000 Hz tone to a 1000 Hz one. Because this was apparently only found in a single location (i.e., straight ahead) it was not recognized by the authors as evidence for transfer. This "spatial restriction", however, is more likely due to the type of task used (2AFC sound lateralization) than to any perceptual effects.

In any case, the importance of finding aftereffects is that the parallel in human and the macaques behavior justifies the use of the latter as a model for exploring the neuronal mechanisms of multisensory perception, which is, for obvious reasons, much more difficult in humans. Some of the main lessons already learned about physiological implementation of multisensory integration as it applies to the auditory-visual case is discussed in the next section.

1.3 Brief neurophysiology of auditory and auditory-visual spatial processing

What brain areas are responsible for auditory spatial processing and the integration and combination of information from across the different senses? A great deal of work has been done to answer these questions and we present only a very brief overview of some of the midbrain and cortical sites known to be involved in auditory and auditory-visual spatial processes. More authoritative and complete treatments are available in Philips and Brugge (1985), for sound localization, and Calvert, Spence, and Stein (2004), and Spence and Driver (2004), for multisensory processes.

1.3.1 Neurophysiology of auditory localization

Since the spatial locus of a sound cannot be directly represented on the cochlear, sound localization is necessarily a computational task. The brain does this based on the signals arriving at the two ears (binaural cues) and signals created by each individual ear (monaural, or spectral cues). Both mid and forebrain structures are involved in this task (Cohen and Knudsen, 1999). The midbrain structures are, the *cochlear nucleus*, the *medial* and *lateral superior olives* (MLO and

LSO), the *inferior colliculus* (IC), the external nucleus of the IC (ICx), and the *superior colliculus* (SC, or its avian analogue, the *optic tectum*). Some of the cortical structures that are involved in auditory spatial processing are the *auditory thalamus*, the *primary auditory cortex* (AI), the association areas, the *anterior ectosylvian sulcus* (AES), and the *posterior parietal cortex* (PPC).

The most important localization cues are those obtained from differences between the two ears as a function of the horizontal position of the sound source. These come in two varieties, interaural time differences (ITD), mostly sensitive to low sound-frequencies and interaural level differences (ILD), mostly sensitive to high sound frequencies (Blauert, 1997). Within the midbrain these two cues are processed separately in the MSO and LSO respectively, and are integrated in the next stage, the IC.

At the same time, there is evidence that the brain retains separate representations of the two interaural cues at a cortical level (e.g., Brugge, Dubrovsky, Aitkin, & Anderson, 1969; Schröger, 1996; Ungan, Yagcioglu, Goksoy, 2001). There are also behavioral data that are in line with having separate neural structures for ITD and ILD processing. Wright and Fitzgerald (2001) showed that ITD and ILD discrimination learning have different time courses. Although both cues showed rapid initial learning, a later slow improvement phase was found for the ILD cue only (Wright & Fitzgerald). In addition, the PPC has also been shown to be involved in the processing of, at least, ITDs (Lewald, Foltys, & Töpper, 2002).

1.3.2 Midbrain structures for multisensory spatial perception

The most widely studied structure, and also a favorite model (Wallace, 2004) of auditory-visual integration, and multisensory integration in general, is the *superior colliculus* (SC), which forms part of the top of the midbrain beneath the posterior part of the cerebral cortex (King, 2004). It has been extensively studied in the cat (e.g., Stein, Jiang, Stanford, 2004; Stein & Meredith, 1993), the ferret (e.g., King, Doubell, & Skaliora, 2004), and to a lesser extent in primates (Jay & Sparks, 1984, 1987). Besides these mammalian models, a lot of work has been done with owls (e.g., Gutfreund & Knudsen, 2004; Konishi, 2000; Luksch, Gauger, & Wagner, 2000), in the avian analogue of the SC, the *optic tectum* (OT).

The following section presents some of the key features of the SC and/or OT that are most common to all animals models, and relevant for auditory and visual spatial perception.

Maps of space

The SC has topographical and overlapping representations of auditory and visual space. This organization allows an efficient way of integrating and coordinating spatial information from the two senses (the SC also has a map of tactile space, but that is not dealt with here).

The auditory and visual maps are derived in fundamentally different ways. Whereas the visual map is more or less a direct derivative of the topographic projections from the retina, the synthesis of the auditory map requires substantial computational effort by the nervous system. It needs to integrate information from across a number of different types auditory localization cues, such as spectral and interaural cues.

The visual map is formed in the superficial layers of the SC/OT, with a larger region devoted to frontal space as compared to more peripheral locations. The auditory map is formed in the deeper layers of the mammalian SC and in the *external nucleus of the avian inferior colliculus*

(ICx). In the owl, for instance, the midbrain auditory localization pathway consists in three structures, the *central nucleus of the inferior colliculus* (ICc), the ICx, and the OT. Auditory localization cues are initially processed in frequency specific channels which end up in the ICc. In the next stage the various auditory spatial cues are combined to produce a frequency a-specific auditory space map in the ICx. Although auditory spatial RFs can be very broad here, the majority are spatially tuned, responding only to a specific sound direction (King et al, 2004). From the ICx the auditory map is projected to the OT in a topographic manner. It is in the OT that the auditory and visual maps merge into a multimodal map.

Response enhancement and depression

A large proportion of multisensory SC cells respond most vigorously to a combination of auditory and visual stimuli, more so than to the most effective of one of these stimuli alone. This has been characterized as a response enhancement, which reflects an interaction in the multisensory SC neurons (e.g., Stein & Meredith, 1993). Enhancement occurs in particular when auditory and visual stimuli are spatially coherent, signifying they arose from the same external event. When they are not spatially coherent a significant response depression occurs.

These physiological effects have a direct reflection in behavior. In one study (Stein et al., 1989), for instance, cats were trained to orient to and approach a visual (a dim light) target while ignoring an auditory distracter (low intensity noise burst). The visual target was presented either alone or together with the auditory one. The auditory distracter was presented either from the same location as the visual target or from a different one. Visual detection and localization performance was markedly improved by auditory stimulus, but only when it was spatially coherent. When it was not spatially coherent detection and localization performance dropped below that of when the visual stimulus was presented alone. This pattern of performance is very similar to the pattern of responses found on the neuronal level. Similar results have also been reported in humans (e.g., Bolognini, Frassinetti, Serino, & Ladavas, 2005).

Vision instructs the auditory space map

It is important that auditory and visual receptive fields (RFs) across the two maps are in register, if they are to be of any behavioral relevance, and it is the visual map in the superficial layers of the SC/OT that provides the spatial template (Gutfreund & Knudsen, 2004; King et al., 2004) against which auditory tuning is matched. When, for instance, owls are raised wearing horizontally displacing prisms, the auditory space map in the ICx shifts in accordance with the optical displacement.

Between the auditory and the visual maps there is a point-to-point correspondence. This has been elegantly demonstrated by Hyde and Knudsen (2002), who studied the effects of a spatially restricted lesion (corresponding to visual frontal space) in the owl's OT, on the auditory space map in the ICx. The results showed that the lesion eliminates adaptive shifts in the auditory map but only for that region representing frontal space. The rest of the auditory map was unaffected and continued to adjust adaptively. Activity in the OT, then, calibrates the auditory space map in a location-specific manner. Similar findings have been reported for the ferret (King, Schnupp, & Thompson, 1998).

1.3.3 Cortical structures for multisensory spatial perception

Multisensory areas are not restricted to the midbrain, but can also be found in the cortex. Here we discuss several of these areas known to be involved in the processing of auditory and visual spatial information.

Anterior ectosylvian sulcus (AES) and lateral suprasylvian sulcus (rLS)

The (feline) AES has a multisensory region on the borders of its three modality specific (auditory, visual, and somatosensory) areas. This particular multisensory region behaves in a manner similar to the SC (Stein, Stanford, Wallace, Vaughan, & Jiang, 2004). Its auditory and visual RFs, for instance, have a tight spatial register, and AES neurons show response enhancements in a similar fashion. In fact, the only difference between AES and SC seems to be the proportion of multisensory neurons and the incidence and magnitude of multisensory depression (Stein, Stanford, et al., 2004). This suggests that multisensory integration is carried out in parallel, and according to the same principles, in the SC and the AES. It is, however, likely that there are functional differences between the two structures. Whereas the SC is responsible for spatial orienting behavior, the AES is more likely to be involved in higher order tasks such as stimulus identification (Calvert, 2001; Stein, et al., 2004).

Interestingly, the *modality specific* parts of the AES play a significant part in multisensory processing in the SC (Wallace, 2004). Wallace and Stein (1994) found that deactivating the AES (using cryogenic blockade, which allows the affected brain region to return to its original status) eliminates the characteristic response enhancement in SC multisensory neurons. At the same time, modality specific responses of SC neurons remained unaffected.

A similar effect has been found for deactivation of the rostral aspect of the LS (rLS), and for simultaneous deactivation of AES and rLS (Jiang, Wallace, Jiang, Vaughan, & Stein, 2001). Deactivation of these structures and their effect on SC multisensory neurons has corresponding effects on overt orienting behavior. The normal improvement of orientation performance due multisensory stimulation is significantly degraded as a function of cortical deactivation (Jiang et al., 2001; Wilkinson, Meredith, & Stein, 1996).

The cortical-midbrain interactions are likely candidates routes for higher level modulations of orienting and localization behavior (Wilkinson et al, 1996). The current knowledge of the role of these cortical structures in multisensory integration is, however, very limited (Calvert, 2001), and therefore any such specific conclusions are tentative at best.

Posterior parietal cortex (PPC)

The PPC is heavily involved in representing sensory targets that are going to be the object of future motor actions and is therefore an important link between perception and action (Cohen & Andersen, 2004). It consists of at least three areas, each responsible for coding for a specific movement plan. The *lateral intraparietal area* (LIP) codes for saccades to sensory targets, and has a very large population of spatially tuned neurons that respond to both auditory and visual stimuli (Mazzoni, Bracewell, Barash, & Andersen, 1996). The *parietal reach region* (PRR) codes for reaches, and the *anterior intraparietal area* (AIP) codes for grasping objects.

The PPC supports several reference frames for the different spatial senses. These senses, however, also all have different initial reference frames. The visual system codes locations in an eye-centered frame, the auditory system in a head-centered frame, and the somatosensory system in a body-centered frame. This, of course, makes good functional sense. Eye movements, for instance, are most easily coded based on the current and the intended direction of the eyes. Similarly, arm movements are most easily coded based on the current and the intended position of the arm. Having different reference frames becomes a problem, however, when information from one is to be used to guide motor actions in another. For instance, how can an arm movement, coded in a body-centered frame, be directed to a sound source that is coded in a head-centered frame?

The solution is to recode, or transform, the reference frames into a common one, and the PPC plays a major role in this. The common frame turns out to be eye-centered, apparently because this is the frame for the visual system which is the keenest of all spatial senses when it comes to spatial perception (Cohen & Andersen, 2004), and it is further modulated by current eye, limb, and body position signals. The common frame has been found in two parts of the PPC, in the LIP (e.g., Andersen, Bracewell, Barash, Gnadt, & Fogassi, 1990), as well as the PRR (Batista, Buneo, Snyder, & Andersen, 1999). In both these areas 42-44% of the cells that were recorded coded in an eye-centered reference frame, 33% (LIP) or 45% (PRR) in a head-centered reference frame (the target was auditory), and the remainder in an frame that coded in a fashion intermediate between eye and head-centered.

1.4 Overview of the thesis

The aim of the present work is to further extent our knowledge of visual recalibration of auditory spatial perception by studying a number of issues as of yet (almost) untouched. The thesis consists in two parts each addressing a major theme. The first, is concerned with the *extension* of changes that are induced by exposure to a ventriloquism situation. The second is concerned with the *time course* of recalibration.

1.4.1 Extension of visual recalibration of auditory localization

At the root of the work in this part of the thesis is the notion that aftereffects provide information, not available in immediate effects, regarding the *extent* of the changes induced by exposure to conflict situations.

The corresponding methodology is basically that of stimulus generalization, which has been extensively studied in research on classical conditioning (e.g., Hovland, 1937) and learning in general. The general approach is to probe for the occurrence of a conditioned response (or in our case aftereffects) at several stimulus values along a certain (physical) dimension, such as sound-frequency or space, after selective exposure to only a very limited part of that dimensions (e.g., a single sound-frequency or spatial location). Typically, in classical conditioning studies, the learned response is largest at the trained stimulus value and tends to diminish for more distant values on the dimension of interest. The resulting function is called the generalization gradient.

In research on multisensory recalibration the stimulus generalization strategy has been used sporadically and only in the study of visuo-proprioception, or "prism adaptation" (Welch, 1978). In a typical study participants watch their hand through laterally displacing optic prisms which create a discrepancy between the seen and felt location of the hand. Exposure is restricted such that participants only see their (displaced) hand in a particular location in space. They very quickly adapt to this in a manner such as to reduce the registered discrepancy, presumably by recalibrating the felt position of the hand (Harris, 1963). Tests of, among others, proprioceptive localization are conducted in that location but also in other locations that were not involved in the exposure period. Bedford (1989), further developed the generalization paradigm to explicitly study how the system recalibrates the mapping between visual location and placing of a finger when only a minimum amount of information is made available about the spatial relationship between the two. The visual input was more strictly controlled than in previous work, and consisted in a small LED that lit up only when the finger pointed exactly at the experimentally intended position. Her results showed that adaptation achieved in one particular location transferred entirely to all other locations along the azimuth. In other words, there was complete generalization across the azimuth. Subsequent work has however produced different patterns of spatial transfer, with the largest aftereffects at the adaptation location and going down with increasing distance from that location (Field, Shipley & Cunningham, 1999; Ghahramani, Wolpert & Jordan, 1996).

The present work applied the stimulus generalization strategy to the case of the ventriloquism aftereffect. Two physical dimensions were explored, those of sound-frequency and space.

Frequency specificity

Chapters 2 through 5 are dedicated to the generalization of the aftereffects of ventriloquism across sound-frequencies. One reason to be interested in this is the information it could provide about the functional and physiological locus of recalibration. Some of the auditory localization pathways are frequency tuned whereas others are not (Cohen & Knudsen, 1999). Patterns of generalization across sound-frequencies can provide valuable clues as to the pathways that are involved in recalibration. A second reason is a more specific version of the first. Generalization is informative about the respective roles in recalibration of the two main sound localization processes based respectively on interaural time differences (ITD) and on interaural intensity differences (ILD). These two mechanisms operate in different frequency domains, low frequencies for the ITD and high frequencies for the ILD. Finding generalization across the two frequencies domains would be a strong indication that the locus of adaptation is beyond that of these peripheral localization mechanisms.

Spatial generalization from local remapping

Work on owls shows that there exists a point-to-point relation between the auditory and the visual space map (Hyde & Knudsen, 2002). Assuming this also to be the case in humans than it might be expected that changes brought about by ventriloquism in a limited part of auditory space does not affect more distant regions. Some very recent evidence collected in humans seems to contradict this. The study by Zwiers et al. (2003; mentioned above) shows that sound localization in regions beyond that of the direct influence of visual information is also affected. In other words, recalibration of auditory space extents (generalizes) beyond regions of auditory space that were directly affected.

Chapters 6 explores generalization across space using a ventriloquist situation allowing more experimental control than the lenses study by Zwiers et al. (2003). Exposure to the spatial conflict is restricted to a single location in space and aftereffects are assessed for that location and more distant ones (e.g., Vroomen, Bertelson, Frissen, de Gelder, 2001).

Relative localization: The effects of having participants point at the target

Whereas all of the work up to this point is done by having participants point at the apparent location of the sound sources, Chapter 7 recaps two studies that use a different localization paradigm that does not involve any directed motor activity. Instead, participants localize the sound relative to a visual reference point. The first study relates to frequency specificity and the second to spatial generalization. Demonstrating aftereffects without a directed motor response might also allow inferences as to the involvement of certain cortical structures in recalibration.

1.4.2 Time course of Recalibration: Acquisition and Retention

The second, although admittedly much smaller part of the present work is on the time course of recalibration, which is described in Chapter 8. It explores two questions, one of *acquisition* and one of *retention.* Acquisition refers to how fast recalibration builds up and when asymptote is reached. Retention refers to how long, given no other stimulation, recalibration is retained, or conversely, how fast it dissipates.

Dissipation has been studied in a large number of different kinds of aftereffects. As for the build up, the (spontaneous) dissipation of an aftereffect is informative on the nature of the underlying mechanisms. Very fast recuperation to baseline suggests the involvement of peripheral mechanisms, whereas very long retention times points to a locus that is much more central and involves cognitive mechanisms. Dissipation functions can also assist in distinguishing between different perceptual mechanisms, such as, for instance, speech adaptation and speech recalibration. Vroomen, van Linden, Keetels, de Gelder, and Bertelson (2004) found that recalibration was dissociable from speech adaptation effect (Eimas & Corbit, 1973) as it could be shown that the dissipation functions of the two effects had distinctly different time courses.

In addition, to date the amount of exposure administered in all studies has been arbitrarily picked by the investigators. It varies from as much as 2500 (Recanzone, 1998) to as little as 18 (Bermant & Welch, 1976) exposure trials. Large differences in exposure might very well lead to differences in adaptation end-points (e.g., Lewald, 2002) and conclusions drawn from one state may not necessarily apply to another. Therefore, the systematic study of the amount of exposure needed to obtain asymptote is clearly of practical and theoretical importance.

Thus in the second part of this thesis a number of experiments are described that have made a beginning at studying both the acquisition and dissipation of the aftereffects of ventriloquism. So far, there have been no previous reports on dissipation, and there are only two brief ones on acquisition (Bertelson, 1993; Radeau & Bertelson, 1976). The present work on acquisition follows the Bertelson study closely in an attempt to replicated those results, while the work on dissipation is necessarily entirely novel.

Chapter 2

The Aftereffects of Ventriloquism: Are They Sound-Frequency Specific? [1]

---

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

2.1 Abstract

Exposing different sense modalities (like sight, hearing or touch) to repeated simultaneous but spatially discordant stimulations generally causes *recalibration* of localization processes in one or both of the involved modalities, which is manifested through *aftereffects*. These provide opportunities for determining the *extent* of the changes induced by the exposure. Taking the so-called *ventriloquism* situation, in which synchronized sounds and light flashes are delivered in different locations, we examine if auditory recalibration produced by exposing tones of one frequency to attraction by discordant light flashes generalizes to different frequencies. Contrary to an earlier report, generalization was obtained across two octaves. This result did not depend on which modality attention was forced on through catch trials during exposure. Implications concerning the functional site of recalibration are briefly discussed.

2.2 Introduction

The visual and the auditory system maintain coordinated representations of external space. The coordination is presumably achieved and maintained by systematically cross-checking between the two modalities. Research on audio-visual spatial coordination, like that on other cases of intermodal coordination, has been mostly based on conflict situations, in which discordant informations regarding the location of potentially the same event is fed simultaneously in the two modalities. Exposure to such spatial discordance produces both online immediate biases and offline aftereffects.

Presenting an observer with synchronous but spatially discrepant auditory and visual information creates a percept in which sound is located nearer to the location of the visual input (Bermant & Welch, 1976; Bertelson & Radeau, 1981; Klemm, 1909; Radeau & Bertelson, 1987). This visual bias of auditory location is generally known as *the ventriloquist effect* (Bertelson, 1999). The effect involves a genuinely perceptual component and cannot be reduced to post perceptual corrections (Bertelson & Aschersleben, 1998; Bertelson & Radeau, 1981). Although demonstrations have often been based on quasi-realistic situations (e.g., steam kettles and whistling noises as in Jackson, 1953, or speech and the moving face of the talker as in Witkin, Wapner & Leventhal, 1952), these are by no means necessary, as biases are easily obtained with stripped-down stimuli such as sound bursts and light flashes (Bertelson & Radeau, 1981; Bertelson, Vroomen, de Gelder & Driver, 2000b; Choe, Welch, Gilford & Juola, 1975; Radeau & Bertelson, 1987; Vroomen, Bertelson & de Gelder, 2001b).

Exposure to the ventriloquism situation also leads to compensatory *aftereffects*, consisting in post exposure shifts in auditory localization (Canon 1970; Radeau 1973; Radeau and Bertelson 1974, 1976, 1977, 1978; Recanzone, 1998), and sometimes also in visual localization (e.g., Radeau 1973; Radeau and Bertelson, 1974, 1976 Experiment 1). It is generally agreed that aftereffects reflect a recalibration process that results in a reduction of the perceived discrepancy, and would play an important role in achieving and maintaining a coherent intersensory representation of space (Held, 1965; Welch, 1978).

At the root of our present research program is the notion that aftereffects may provide information, not available in immediate effects, regarding the *extent* of the changes induced by exposure to conflict situations. Probing for the occurrence of aftereffects at several stimulus values can tell us whether these changes are specific to the values used in the conflict situation or instead generalize to a range of neighboring stimuli. This research strategy was inaugurated by Bedford (1989) who examined the extension of visuo-proprioceptive spatial recalibration in a task consisting of learning new mappings between visual location and placing of a finger. She found that adaptation achieved in one particular location transferred entirely to all other locations along the azimuth. Other work on visuo-proprioceptive or visuo-motor remapping has however produced a different pattern of spatial transfer, with aftereffects largest at the adaptation location and going down with increasing distance from that location (Field, Shipley & Cunningham, 1999; Ghahramani, Wolpert & Jordan, 1996). A similarly decreasing pattern of spatial generalization has also been reported by Vroomen, Bertelson, Frissen, and de Gelder (2001) for the after-effects of ventriloquism.

In the present study, we considered another extension-related question, namely whether the visual recalibration of perceived sound location is specific to the spectral characteristics of the sounds used during adaptation. One reason for being interested in that question was its possible relation to the distinction between the two basic sound localization mechanisms based on respectively interaural level differences (ILD) and interaural time differences (ITD). ILD is known to be the main cue for localizing sound frequencies above 1.5 kHz and ITD for lower frequencies (Cohen & Knudsen 1999; Blauert, 1997). Finding that adaptation does not generalize from one of these two frequency ranges to the other one would suggest that it affects specifically these somewhat peripheral mechanisms, rather than one located at a more central site.

Results briefly mentioned by Recanzone (1998) seemed to support the latter suggestion, although the relation to ILD and ITD was not considered in the paper. Three participants were exposed to either 750 Hz or 3000 Hz tone bursts synchronized with a light flash 8 deg to the right. Strong auditory aftereffects (7.08 deg overall) occurred when the test stimuli's frequency was the same as that of the adapters, but none at all with test stimuli at the other frequency. Given the importance of their possible implications, these results, based on a single direction of adaptation and only three participants, clearly needed some verification. Here, the generalization across the same two sound frequencies of the visually induced recalibration of apparent sound location was examined, for both leftward and rightward adaptation, with pure tones and light flashes as material, and a more substantial number of participants. Three patterns of results were in principle possible. First, no transfer whatsoever across sound frequencies, as reported by Recanzone (1998). Second, complete transfer, as Bedford (1989) found for visuo-proprioceptive re-mapping. Third, partial transfer, meaning an aftereffect that is maximal around the adapting frequency and goes down with increasing distance from that frequency, as in Field et al. (1999) and Ghahramani et al. (1996), also for the visuo-proprioceptive case.

2.3 Experiment 1

Following the classical pre-tests–adaptation–posttests paradigm, participants in this experiment localized tones at two frequencies, 750 and 3000 Hz, before and after repeated exposure to one of these tones synchronized with a spatially discordant light flash, 9 deg to its left or its right. This design provided, for each adaptation frequency, separate measures of aftereffects at the adaptation frequency and at the other one.

2.3.1 Method

Participants.

Eighteen students from Tilburg University (age 18-30, twelve 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, at separations of at least one hour.

Apparatus and material.

The testing was carried out in a completely dark soundproof booth. The setup consisted of three display units which were occluded by means of a black, acoustically transparent cloth and an array of push buttons. Display units, each of which consisted of a loudspeaker with an LED over its center, were arranged along a semi-circular array at 85 cm from the participant's head, at eye level and at –9, 0, and +9 deg distance from centre in the horizontal plane. Pushbuttons, 108 in total, were arranged along another semi-circular array 20 cm in front of the display units, and 30 cm below them.

The auditory stimuli were 200 ms pure tones (with 5 ms linear rise/fall envelopes) at either 750 Hz or 3000 Hz, presented at 70 dB(A).

Procedure

Each session consisted of 120 auditory pretests, followed by 60 bimodal exposure trials, and then by a 120 posttests. On pre- and posttests, the sound was presented 20 times on each of the three loudspeakers at each of the two frequencies (750 Hz and 3000 Hz). All combinations of speaker location and sound frequency were presented in random order, with 3 sec inter-trial intervals. The participant was instructed to always push the button corresponding to the apparent horizontal location of the presented sound. Bimodal exposure trials all consisted of the presentation of the 200 ms sound in the central loudspeaker in synchrony with the 200 ms lighting up of the LED of the adjacent display unit, to the left or to the right, depending on the session. To impose attention to the visual distracter, two catch trials, on which the sound was presented without accompanying light flash, were interspersed at random among the exposure trials, and participants were instructed to report their occurrence verbally. Half the participants (n=9) had all their bimodal exposure trials with the 750 Hz tone, and the other half with the 3000 Hz tone. Each participant was adapted to the left on one session and to the right on the other session, in balanced order.

2.3.2 Results

Performance on catch trials was flawless.

Separately for each speaker location and test frequency, responses that were more than 2.5 standard deviations from the mean were excluded from the analysis. They represented 1.6% of the data. Aftereffects were calculated by subtracting mean reported locations on pretests from those on posttests. Aftereffects were counted as positive when they went toward the visual distracter.

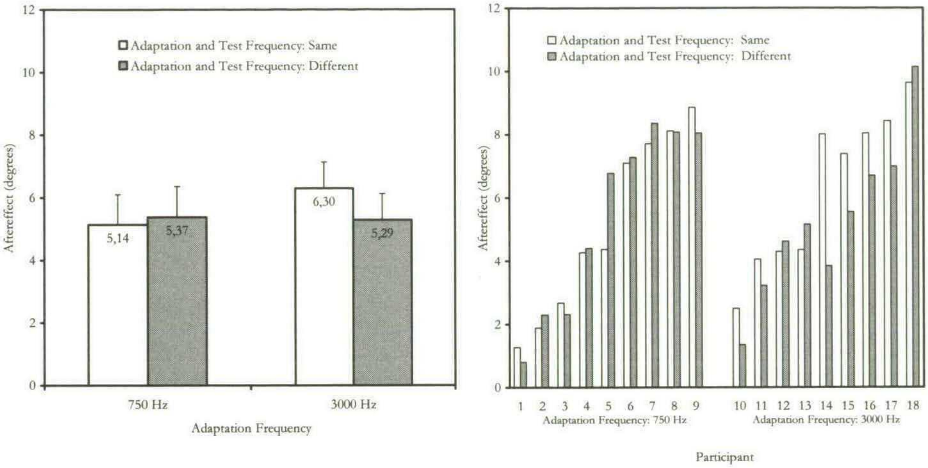


Figure 2.1 Experiment 1. Left panel: Mean aftereffects per combination of adaptation frequency and test frequency (with standard errors). Separate groups of nine participants were adapted at each frequency. Right panel: Individual results. In each group, participants were ordered according to overall effect size.

Aftereffects for each combination of adaptation frequency and test frequency (Fig. 2.1, left panel) were all substantial, with small differences between those measured with same frequency tests and with different frequency tests. They were submitted to an Adaptation Frequency (750 Hz vs. 3000 Hz) x Test Frequency (same as in adaptation vs different) x Direction of Adaptation repeated measures ANOVA. The overall aftereffect, across all conditions, was highly significant, $F(1, 16) = 78.50, p < .001$. None of the main effects, Adaptation Frequency ($F < 1$), Test frequency ($F(1, 16) = 1.73, p = .21$) nor Direction ($F(1, 16) = 3.48, p = .067$) was significant, neither was any of the interactions. Among these, the one between Adaptation Frequency and Test Frequency was however close to significance, $F(1, 16) = 4.43, p = .052$. This presumably reflects the existence across the two Adaptation Frequencies of Test Frequency effects which, although small, go in opposite directions. By t-tests, none of these two effects reached significance (3000 Hz: $t(8) = 2.01, p = .079$; 750 Hz: $t < 1$).

Individual results (Fig 2.1, right panel) show considerable variations in mean aftereffect, but the differences between "same" and "different" were all small and displayed no systematic tendency: 10 of 18 participants had a larger aftereffect in the "same" condition. This deviation from equi-frequency is of course non-significant ($p = .41$).

2.3.3 Discussion

Measuring recalibration with a tone whose frequency is different from the one used during exposure had no significant effect on the size of the aftereffects. Judging from this result, visual recalibration of auditory location is not specific to sound frequency. This conclusion, of course, contrasts sharply with that of Recanzone (1998). There are however several differences between the procedures of the two studies that might have contributed to the discordant outcomes. The main one is that Recanzone used a larger number of exposure trials (2500, against 60 here), and one may wonder whether prolonged exposure does not bring about a “fine tuning” process, by which recalibration would become more specific of the particular sound frequency at which it was produced. Other differences concern the locations at which exposure was conducted, a single central location like here or a range of locations (nine in Recanzone’s case), and the allocation of conditions to the participants. Recanzone’s participants were adapted in a single direction, but at each of the two sound frequencies, while each of ours was adapted, on separate sessions, in the two directions, but at a single frequency. A new experiment was run in which these various differences were minimized.

2.4 Experiment 2

This new experiment was run with the same setup and the same design as Experiment 1, but with a procedure modified in such a way as to bring it closer to that of Recanzone (1998). The modifications concerned the number of adaptation trials (2400 instead of 60 in Experiment 1), the number of locations at which adaptation was conducted (five instead of a single one) and the conditions under which each participant worked (now a single direction of adaptation, but both frequencies, on separate sessions).

2.4.1 Method

Participants

Fourteen new participants from the same student population (age 19-32, eleven female) participated in two sessions each.

Apparatus and material.

The setup and stimuli were the same as in Experiment 1, except that the number of display units was increased to seven, located at -27, -18, -9, 0, 9, 18 and 27 deg from centre.

Procedure.

Each of the two sessions consisted of 200 auditory pretests, followed by 2400 exposure trials, and then 200 posttests. Pre and posttests were now initiated by the participant by pressing a button 20 cm in front of her/him with the pointing hand, a procedure ensuring that all pointing movements started from a constant position. The sound was presented 20 times on each of the five central loudspeakers (at –18, -9, 0, +9, and +18 deg), and at each of the two frequencies, all in randomized order.

On bimodal exposure trials, the sound was presented in randomized order across the five speaker locations, and the light flash always on the left adjacent unit for half the participants, and to the right for the other half. An average of 25 catch trials, on which the light flash was omitted and which the participants had to count, were interspersed at random locations among exposure trials. As in Experiment 1, their role was to force attention to the visual distracters. The participant reported his final count at the end of the adaptation phase, immediately before the post-tests. Each participant was adapted with one frequency on one session and with the other frequency on the other session, in balanced order.

2.4.2 Results

Catch trial detection, estimated from participants' reported counts, ranged from 82% to 100%. Application of the same principles as in Experiment 1 resulted in the exclusion of 1.4% of outlying responses.

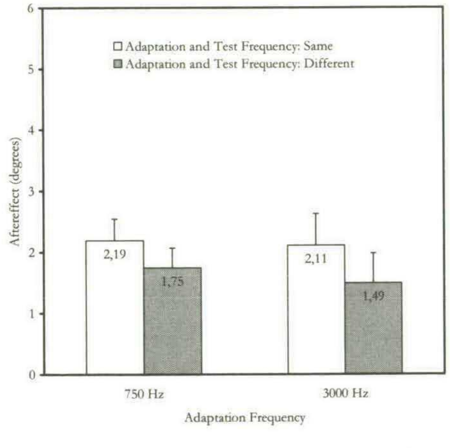


Figure 2.2 Experiment 2. Mean aftereffects per combination of adaptation frequency and test frequency (with standard errors). Separate groups of seven participants were adapted in each direction.

As seen in Fig 2.2, substantial aftereffects were obtained in each of the four conditions. Differences between conditions with same frequency tests and with different frequency tests, although still small, now went for both adaptation frequencies in the direction of stronger aftereffects in the "same" condition. In the Adaptation Frequency (750 Hz vs. 3000 Hz) x Test Frequency (Test Frequency same as Adaptation frequency vs different) x Direction repeated measures ANOVA, the overall aftereffect was highly_significant, $F\,(1, 13) = 31.26, p < .001$, but neither Test Frequency ($F < 1$), Adaptation Frequency ($F < 1$), Direction ($F\,(1, 13) = 1.94, p =$ .19) nor any of their interactions was significant. This was in particular the case for the Adaptation Frequency x Test Frequency interaction that was nearly significant in Experiment 1, but now fell clearly short of significance ($F\,(1, 13) = 2.84, p = .118$).

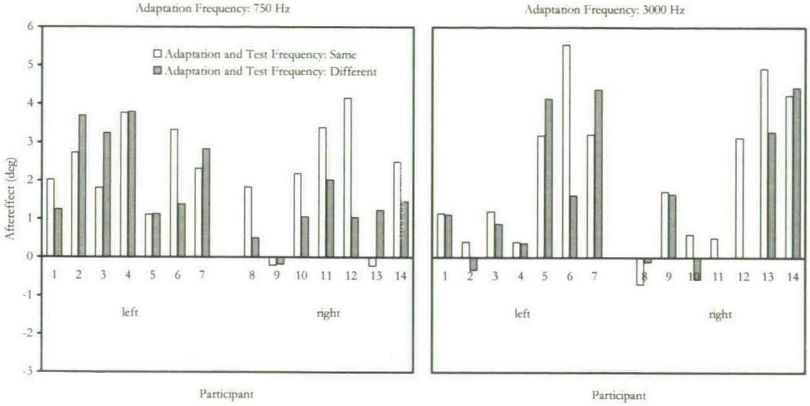


Figure 2.3 Experiment 2: Individual results. In each panel, participants from each group were ordered according to their overall effect size across adaptation frequencies.

At the level of individual participants, Fig. 2.3 shows that the differences between aftereffects measured at the adaptation frequency and at the other frequency showed again large variations. Seven out of the 14 participants had larger aftereffects in the same condition for adaptation frequency 750 Hz, and 10 out of 14 for adaptation frequency 3000 Hz, both deviations from equi-probability non-significant by sign-test.

2.4.3 Discussion

Just as in Experiment 1, aftereffects measured at either of the two adaptation frequencies were only slightly and non-significantly larger than those measured at the other frequency. The changes we introduced in our procedure to make it more similar to Recanzone's, at the levels of number of adaptation trials, number of exposure locations and different conditions each participant worked with, failed to eliminate the contrast between the respective results. However, one difference whose possible influence has not been considered so far concerns the modality to which attention was directed by catch trials during exposure. In the two preceding experiments, it was the visual modality, whereas Recanzone had his participants monitor the auditory signals for occasional amplitude reductions. Could it be the case that increased attention to the auditory stimulus during adaptation makes recalibration more specific of the spectral characteristic of that stimulus? To check on this possibility, the task of Experiment 2 was in the next experiment administered with Recanzone's auditory catch trials substituted for our visual ones.

2.5 Experiment 3

Experiment 3 was in all details identical to Experiment 2, except that during exposure participants performed an auditory monitoring task instead of a visual one.

2.5.1 Method

Fourteen new participants from the same student population (age 18-24, eleven female) took part in two sessions each. Apparatus, material and procedure were as in Experiment 2, except that catch trials (again an average of 25 interspersed among exposure trials) now consisted of the attenuation of the tone by 10 dB. Participants were again instructed to count the number of catch trials.

2.5.2 Results

Individual performance on catch trials ranged from 76% to 100%. One percent of outlying responses were discarded.

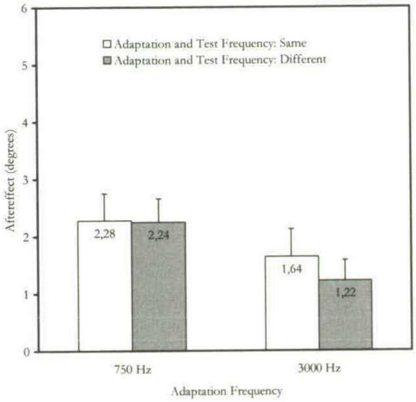


Figure 2.4 Experiment 3: Mean aftereffects per combination of adaptation frequency and test frequency (with standard errors). Separate groups of seven participants were adapted in each direction.

Fig 2.4 shows that substantial aftereffects were obtained in each of the four conditions. They are now stronger at both adaptation frequencies in the condition with "same" test frequency, but none of the two differences was significant by t-test, both ts < 1. In the Adaptation Frequency (750 Hz vs. 3000 Hz) x Test Frequency (Test Frequency same as Adaptation frequency vs different) x Direction of Adaptation (Left vs. Right) repeated measures ANOVA, the overall aftereffect, across all conditions, was highly significant, $F\ (1, 13) = 49.82$, $p < .001$, but neither Adaptation Frequency ($F\ (1, 13) = 2.46$, $p = .143$), Test Frequency ($F < 1$, supporting above mentioned *t*-test results), Direction ($F < 1$), nor any of their interactions (the three-way interaction $F\ (1, 13) = 2.72$, $p = .125$, all two-way Fs < 1) were significant. *T*-tests applied to differences between conditions with testing frequency same as adaptation frequency vs. different were both non-significant.

This pattern of results is thus nearly identical to the one obtained in Experiment 2. To check on the absence of difference between the two results, a 2 (Experiment 2 or 3) x 2 (Adaptation Frequency) x 2 (Test Frequency same or different) ANOVA was carried out. Except for the overall effect, F (1, 26) = 116.61, $p < .001$, none of the main effects nor of their interactions was significant (all Fs < 1).

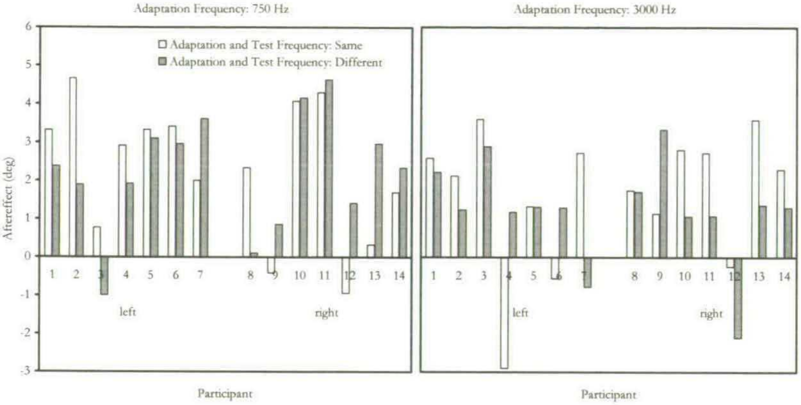


Figure 2.5 Experiment 3: Individual results. In each panel, participants from each group were ordered according to their overall effect size across adaptation frequencies.

At the level of individual participants (Fig. 2.5) there were again large variations. Seven out of 14 participants had larger aftereffects at adaptation frequency for adaptation frequency 750 Hz, and 9 out of 14 for adaptation frequency 3000 Hz, both deviations from equi-probability non-significant by sign-test.

2.5.3 Discussion

Replacing the visual catch trials of Experiment 2 by auditory ones had no detectable effect on the pattern of extension, which for the two adaptation frequencies remained strong generalization to the other frequency, with only a small and non-significant reduction. The use of an auditory monitoring task could thus not the be the reason behind Recanzone's no-generalization result.

2.6 General discussion

In each of the preceding three experiments, exposure to spatially discrepant tones and light flashes resulted in strong and highly significant recalibrations of apparent sound location, and measuring these recalibrations at tone frequencies distant by two octaves from the one used during adaptation, rather than at the same frequency, produced only small and non-significant reductions of recalibration size. These practically identical results were obtained across wide variations in experimental parameters such as number of adaptation trials, adaptation at a single or several locations, attention orientation to one or the other of the involved two sensory modalities, and range of conditions each participant was exposed to.

Our findings are thus clearly inconsistent with the notion of spatial recalibration as specific to the sound frequency used during exposure. To that extent, they are in sharp contrast to those of Recanzone (1998) who obtained no generalization whatever between exactly the same two tone frequencies. The changes introduced in our procedure across the successive experiments were meant essentially at identifying the sources of the contrasting findings. In Experiment 3, these changes culminated in having all the main parameters practically identical to those in Recanzone's experiment. The only remaining explanation of the contradiction must reside in the small scope (three participants) of Recanzone's investigation. Some of the present participants also displayed small or inexistent degrees of transfer, like participants 8 and 12 in Experiment 2 and participant 8 in Experiment 3 (this one for adaptation at 750 Hz only).

As explained in the Introduction, the main theoretical motivation for the present study was the information it might provide regarding the possible roles of the basic ILD and ITD sound localization mechanisms in recalibration. Of the two frequencies we used, one, 3000 Hz, belongs to the range in which ILD dominates, and the other one, 750 Hz, belongs to the ITD range. Confirming Recanzone's (1998) finding of strong specificity would have suggested that recalibration takes place at the level of these peripheral mechanisms. Our observation of important generalization between the two frequencies implies that, to the contrary, recalibration probably takes place at a later processing stage, after outputs from the two peripheral systems have been combined.

The present result, strong generalization with only small and non-significant reduction across frequency change, is, as already mentioned, inconsistent with one of the three patterns considered in the Introduction, strict specificity. Given that observed reductions were significant in none of the experiments, it is for the time being compatible with total generalization, but a larger range of frequency shifts must still be examined before the alternative possibility of a diminishing gradient can be completely ruled out.

Our search for an explanation of the contradiction between our results and those of Recanzone has led us to consider the possible effect of selectively attending to either modality. The comparison of the results from respectively Experiments 2 and 3 revealed no effect whatsoever on the generalization pattern. On the other hand, there was also no effect on the overall size of the aftereffects. Although it has no relation to the main objectives of our study, this result could at first sight be taken as contradicting earlier results in the literature showing a dependence of that size on which modality attention was focused on during exposure. Both Canon (1970) and Radeau (1974) measured auditory and visual aftereffects after a period of exposure to audio-visual discrepancy, during which participants pointed selectively either to the sounds or to the flashes. Auditory aftereffects were larger in the visual pointing conditions than in the auditory pointing ones, and vice-versa for visual aftereffects (There were actually no visual aftereffects after visual pointing in Canon's experiment). Focusing selective attention in one modality through pointing instructions apparently increased its relative weight in the recalibration process. Similar results were obtained by Kelso, Cook, Olson & Epstein (1975) for a case of visuo-proprioceptive conflict.

There are important differences between the methods by which attention to modality was controlled these three studies (through selective pointing) and in our own one (through stimulus or stimulus change detection), which may have played a role in bringing about the divergent outcomes. First, pointing involves other processes than just orientation of attention, most notably motor processes. On the other hand, the attention it requires is specifically attention to target location, which may not necessarily be required by the detection tasks. The whole question of the role of attention in crossmodal interaction should clearly be revisited with better consideration for the various possible ways of manipulating it.

Chapter 3

The Aftereffects of Ventriloquism: Generalization Across Sound-Frequencies [1]

---

[1] Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2005). The Aftereffects of Ventriloquism: Generalization Across Sound-Frequencies. *Acta Psychologica, 118, 93-100.*

3.1 Abstract

Exposure to synchronous but spatially discordant auditory and visual inputs produces, beyond immediate crossmodal biases, adaptive recalibrations of the respective localization processes that manifest themselves as aftereffects. The present study is part of a research program focused on the way such recalibrations generalize to stimulus values different from those used for adaptation. In contradiction with earlier reports that auditory aftereffects did not generalize across sound frequency, we recently found quasi-total generalization across two octaves. In the experiment described in this paper, participants were adapted to a 18° auditory-visual discordance with either 400 or 6400 Hz tones, and their subsequent sound localization was tested across this whole four-octave frequency span. Substantial aftereffects were obtained at all combinations of adapter and test frequency, with only small reductions with distance from adapting frequency. Implications of these results concerning the functional site at which visual recalibration of auditory localization takes place are discussed.

3.2 Introduction

The currently very active movement of research into multisensory perception resorts extensively to conflict situations, in which separate sense modalities receive incongruent inputs regarding a same aspect of the environment (see Bertelson & de Gelder, 2004, for a recent review). One conflict situation that has proved especially convenient for experimental study involves presentation of synchronous auditory and visual stimuli in slightly separate locations. 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 (Bertelson, 1999).

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 effect (i. e. observed in presence of the conflict situation) is generally called the *visual bias of apparent auditory location.* The other consequence can be observed off-line, after repeated exposure to similar incongruent audiovisual stimulus pairs, when localization responses to singly presented stimuli in one of the modalities are displaced in the direction occupied during exposure by the conflicting stimuli in the other modality (Canon, 1971; Radeau & Bertelson, 1974, 1976). Such *aftereffects* show that exposure to the spatially incongruent data produces *recalibration* of stimulus-to-percept matches in the involved modalities. Such recalibration generally reduces the perceived incongruence, and for that reason can be expected to play an adaptive role in the development and later maintenance of crossmodal coordination (Held, 1965; Redding & Wallace, 1997; Welch, 1978).

A consideration at the origin of the present line of investigation was that aftereffects also offer opportunities, not available in immediate effects (like biases), 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 extension of visual-proprioceptive spatial recalibration in finger placing under displaced visual feedback, and found that adaptation achieved at a particular location generalized entirely to other locations along the azimuth.

We have started a research program that focuses on the extension issue in the particular case of auditory aftereffects of exposure to audiovisual spatial conflict. One part (mainly unpublished so far, but see 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 deg 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 a small, non-significant, reduction) 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), and 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 concluded that the apparent contradiction might be explained by the small scope of Recanzone's investigation.

One important reason for being interested in the question of extension across sound frequencies that was 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, based on interaural time differences (ITD) and interaural level differences (ILD), the former of which is known to be sensitive mainly to frequencies below 1500 kHz and the latter to higher frequencies (Blauert, 1997; Cohen & Knudsen, 1999). Finding (as did Recanzone) that adaptation does not generalize between these two frequency ranges would have suggested that recalibration occurs at the level of these peripheral processes, while the opposite finding would point to a more central site, posterior to the integration of the peripheral processes' outputs. Frissen et al's results of course brought support for the second conclusion.

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 the Frissen et al (2003) studies. In the following experiment, generalization was examined across a four-octave range from 400 to 6400 Hz, providing a better opportunity for any form of frequency-specificity to manifest itself. Aftereffects were measured at one-octave intervals across the range, thus making it possible to determine the shape of any possible gradient.

3.3 Method

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.

Apparatus and Material

The testing was carried out in a dark soundproof 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 deg left and right of the latter's sagittal 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).

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 8 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 deg 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.

3.4 Results

Aftereffects were calculated by subtracting mean reported locations on pretests from those on posttests, and were counted as positive when they went toward the visual distracter. They were pooled over speaker locations and direction of adaptation. The results are shown in Fig. 1 as functions of distance between adapter and test frequency.

There were no differences between the aftereffects obtained at the two adaptation frequencies, neither when measured at these frequencies nor at the other ones. At both frequencies, aftereffects were slightly reduced when distance between test and adaptation frequency increased, but substantial aftereffects (in the order of 70 % of those at zero distance) were still observed at the maximum four-octave distance.

In a 2 (Adapter) x 5 (Distance) repeated measures ANOVA, the overall aftereffect of 2.61 deg. was highly significant, F (1, 14) = 91.56, $p < .001$, as was the main effect of Distance, F (4, 56) = 5.00, $p < .01$. Trend analysis of the distance effect produced a significant linear component, F (1, 14) = 11.71, $p < .005$, and a non-significant interaction with adapter ($F < 1$). Slopes were -.19 deg/octave for adaptation at 400 Hz and -.25 deg/octave at 6400 Hz. The main effect of Adapter ($F < 1$) and its interaction with Distance, F (4, 56) = 1.19, $p > .32$, were both non-significant. Separate corrected t-test showed that aftereffects at each distance were significantly bigger than zero (all p's < .005).

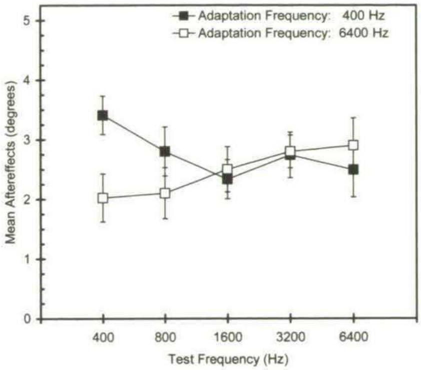


Figure 3.1 Mean aftereffects (with standard errors) averaged across speaker locations and directions of adaptation, for fifteen participants as a function of distance between adaptation and test frequency. Adapters were either a 400 or a 6400 Hz tone, and test frequencies, separated by one-octave intervals, were 400, 800, 1600, 3200, and 6400 Hz.

3.5 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 brings strong additional support for our earlier conclusion that visual recalibration of auditory location is not limited to the sound frequency presented during conflict exposure. This makes it more plausible that the no-generalization finding of Recanzone's (1998) was reflecting sampling error.

Second, aftereffects were reduced when the frequency difference between the adapter and test increased. As mentioned in the Introduction, small reductions of aftereffects measured at the non-adaptation frequency occurred in each of Frissen et al.'s (2003) experiments, leaving open the possibility of a monotonous 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 interesting implications regarding the roles of ITD and ILD in the recalibration of sound localization. At the largest distance between test and adaptation frequency that were considered here, test tone and adaptation tones must each have been localized initially via one of the two processes. The fact that substantial generalization was obtained nevertheless suggests strongly that recalibration takes place mainly at a more central site, posterior to the integration of outputs from the two peripheral processes. Another aspect of the results worth noting is the absence of any discontinuity (possibly corresponding to a boundary between the domains of operation of the two processes) in the obtained generalization gradients.

Chapter 4

The Aftereffects of Ventriloquism: Generalization Across Sound-Frequencies is Unaffected by Rise Time

4.1 Introduction

A large portion of research on multisensory perception has been concerned with the interaction between the spatial senses and in particular between the visual and the auditory system. In a typical study these two modalities are presented with spatially incongruent inputs and an assessment is made of how the perceptual system deals with the discrepancy. The processes put into play by this conflict have been called *ventriloquism*, after the performing ventriloquist who creates the illusion that the speech he produces comes from a puppet (for a review see, Bertelson, 1999).

Exposure to a ventriloquism situation has two main behavioral consequences. The first is the immediate, or on-line, *visual bias of apparent auditory location.* That is, the apparent location of a sound is shifted towards that of a synchronous but spatially discrepant visual input, despite instructions to ignore the latter (e.g., Bertelson & Radeau, 1981). The second consists of compensatory *aftereffects*, which can be observed off-line after exposure to a ventriloquism situation. These consist in post-exposure shifts in auditory localization (Canon, 1970; Radeau & Bertelson, 1974), and sometimes also in visual localization (Radeau, 1973, Radeau & Bertelson, 1974). Aftereffects are generally considered to be the manifestation of a *recalibration* of input-to-percept matches, which would result in a reduction in the registered intersensory spatial discrepancy (Held, 1965; Welch, 1978).

The research project to which the present study belongs stems from the notion that aftereffects offer information not available from immediate effects, namely the *extent* of the changes induced by the conflict situation. Here we look at the sound-frequency specificity of the aftereffects of ventriloquism. The importance of this line of investigation lies in the information it can provide as to the site in the functional architecture at which recalibration takes place. Following the classic duplex theory of sound localization there are two peripheral mechanisms for horizontal sound localization. The first is based on interaural level differences (ILDs) and operates on high frequency signals, and the second is based on ongoing phase, or interaural time differences (ITDs) and operates on low frequency signals (Blauert, 1997; Cohen & Knudsen, 1999; Hafter, 1984). There is also neuroanatomical evidence that these two types of cues are initially dealt with separately in the nervous system (Phillips & Brugge, 1985).

In two previous studies we reported evidence for a strong generalization across sound-frequencies. Participants were adapted to an auditory-visual spatial discrepancy with the auditory adapter at a particular sound-frequency and aftereffects were measured at that frequency and at other ones. In the first study (Frissen, Vroomen, de Gelder, & Bertelson, 2003) we showed that aftereffects generalized quasi totally across a two-octave distance. This held irrespective of which modality attention was allocated to during adaptation. In the second study (Frissen, Vroomen, de Gelder, & Bertelson, 2005) the exploration of the generalization pattern was extended to a four-octave range of test frequencies (between 400 and 6400 Hz) and aftereffects were measured at one-octave intervals across that range. This study again showed substantial and significant aftereffects across the whole range of test frequencies, which however went down significantly with increasing distance between adapter and test frequency. These findings strongly suggest that recalibration occurs at a site more central to the two peripheral sound localization mechanisms.

These results are in contrast with those of two studies in which no generalization was found across sound-frequencies (Lewald, 2002; Recanzone, 1998). In Recanzone's study participants were exposed to either 750 or 3000 Hz tone bursts synchronized with a light flash spatially offset by 8°. Substantial auditory aftereffects were obtained only when adapter and test frequencies were the same, but not when they were different. These results, however, are based on a single direction of adaptation and just three participants. It is not uncommon to find large individual differences in intersensory recalibration (Redding & Wallace, 1997) and several of our own participants also showed no generalization. We have therefore suggested that the results of Recanzone can be explained by the small scope of that investigation. Lewald (2002), however, recently replicated Recanzone's (1998) results using a somewhat larger sample size and both adaptation directions. Two groups of seven participants adapted to either 1000 or 4000 Hz tone bursts synchronized with a light flash spatially offset by 20°. As in Recanzone's study, substantial aftereffects were found only when adapter and test frequency were the same.

A closer look at the auditory stimuli used in our studies and Lewald's (2002) suggested an explanation for the contrasting results. Whereas our auditory stimuli had short rise/fall times (R/F) of 5 ms (Frissen et al., 2003) or 10 ms (Frissen et al., 2005), the corresponding value was 20 ms in Lewald's study. This difference takes on significance in light of the fact that interaural timing cues come in two types. As described before, one cue is the ongoing phase delays between the two ears. However, when the signal first reaches the two ears it also creates a transient *arrival time difference* simply because it reaches the ear closer to the source sooner than the farther ear. The latter cue is presumably mainly dependent on the onset of the signal and should be effective only when the onset is fast (i.e., short) enough (e.g., Rakerd & Hartmann, 1986). The dependency means that the cue should be relatively independent of the sound-frequency of the signal. It can then be argued that in our previous experiments that the dominant localization cue was the onset defined arrival time difference, and that it was this cue, and not the supposed interaural phase differences cue, that was subject to recalibration [1]. Consequently, it should not be surprising that aftereffects generalize to any tone, irrespective of its sound-frequency, that has a similar steep onset, as any potential frequency specificity of recalibration is masked. By the same logic, elimination of the onset cue should make any frequency specificity apparent.

The exact role of the stimulus onset for free-field sound localization, however, seems unclear. Most studies relevant to the issue of stimulus onset used highly constrained dichotic headphone stimulus presentations and free-field studies are rare. Apparently, under free-field conditions sound localization is only weakly dependent on onset. Krahe, Larsen, and Ronacher (2000) could not find any systematic difference in localization performance for fast versus slow onset stimuli (i.e., 2 and 18 ms, respectively), and this was the case irrespective of sound-frequency (i.e., 750, 3000, or 5000 Hz). Rakerd and Hartmann (1986) on the other hand, found that under certain circumstances the onset may be the only reliable localization cue, such as with a continuous pure tone signal in an echoic environment. An example of this is the Franssen effect (Hartmann & Rakerd, 1989) that can be created, in an echoic room, using two loudspeakers placed on either side of an observer. A sine tone is presented from one of the speakers (e.g., the left one) with an immediate onset, which then starts to decay at a certain rate

---

[1] We thank Dr. Jorg Lewald for suggesting this possibility.

immediately after its onset. At the same time and rate a similar sine tone begins to rise at the other (right) speaker, which continues for an arbitrary amount of time. Thus presented at the left speaker is a transient signal, whereas at the right speaker is presented a steady-state signal. Observers perceive such a stimulus as coming from the *left* speaker only, and are unaware that the right speaker ever was active. The explanation given by Franssen (in Hartmann & Rakerd, 1989) was that the auditory system gives a particular importance to the onset. The effect can be likened as a particular instance of the precedence effect (e.g., Litovksy, Colburn, Yost, & Guzman, 1999), wherein the signal that first reaches the observer's ear completely dominates later signals that are created by, for instance, reflections from the walls.

The present experiment therefore is a direct test of whether the fast onset of the stimuli in our earlier studies contributed to the strong generalization across sound-frequencies. Auditory stimuli were created (at 400 and 6400 Hz) that had either short (5 ms) or long (50 ms) rise/fall times (R/F). (The 50 ms onset is even longer than Lewald's (2002) 20 ms onset.) The resulting four stimuli all served as adapter and test stimuli in a completely crossed design.

4.2 Method

Participants

Eighteen students from Tilburg University (age range 18-26, 7 female), all naïve as to the purpose of the Experiment, with normal hearing and normal or corrected to normal vision, each participated in two two-hour sessions, at least one day apart.

Apparatus and Stimuli

The testing was carried out in a dark, soundproof and semi-reverberant 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 a red 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 only used for presenting visual inputs. Push buttons, 108 in total, were arranged along another circular array, 5 cm in front of the display units.

The adapter/test stimuli were 200 ms pure tones at 400 or 6400 Hz, with either a short (5 ms) or a long (50 ms) linear rise/fall time. Presentation was at 66 dB(A). Visual stimuli were 200 ms light flashes from one of the LEDs, always in synchrony with the auditory stimulus.

Procedure

Eight conditions were created, one for each combination of adapter tone and direction of the visual distracter. Per session four of the eight conditions were administered in balanced order with the restriction that all adapters had the same rise/fall times. Half the participants first completed the conditions with the fast onsets and the other with the slow onsets. Between conditions there was a break of 15 minutes that was spent outside the laboratory.

Each separate run consisted of one block of pre-test trials, followed by 6 adaptation-post-tests blocks. On the pre-tests, each of the four test tones was presented six times from each of the four loudspeakers, for a total of 96 randomized localization trials. The participant was instructed to always press the pushbutton corresponding to the apparent direction of its source. The participant initiated the presentation by pressing a separate button located 20 cm straight in front of the participant, a procedure that ensured a constant starting position for pointing movements.

Adaptation-post-tests blocks started with 60 presentations of the condition's particular bimodal adapter 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-second speed, for a total of 1 minute of exposure. To ensure that the visual distracter was attended to during exposure, it was occasionally omitted (3.3%). The participant was required to press the same button as for initiating the localization trials whenever such an omission occurred.

The series of adaptation trials were followed by 16 quasi-randomized post-tests trials, one for each test tone delivered on each loudspeaker. Randomization was such that each test tone was presented on each of the 16 possible sequential positions in the posttest. The procedure here was the same as on the pre-test trials.

4.3 Results

One participant was excluded from further analysis because of his poor performance on the visual catch trials (Z-score < -4, relative to overall group performance). The mean percent correct for the remaining set of participants was 94% (SD = 2.9) and ranged from 89 to 99%.

Table 4.1 Mean aftereffects (in degrees, with standard errors) for each combination of adapter and test frequencies and rise/fall times (R/F).

Adapter		Frequency Same		Frequency Different	
Frequency	R/F	R/F Same	R/F Different	R/F Same	R/F Different
400 Hz	5 ms	2.68 (0.48)	2.63 (0.43)	1.75 (0.58)	2.07 (0.58)
	50 ms	3.11 (0.40)	2.74 (0.46)	1.90 (0.35)	1.78 (0.43)
6400 Hz	5 ms	2.60 (0.25)	3.40 (0.38)	2.25 (0.38)	2.45 (0.35)
	50 ms	2.54 (0.61)	2.13 (0.49)	1.65 (0.42)	1.44 (0.42)
	Mean	2.73 (0.25)		1.91 (0.24)	

Note. All aftereffects were larger then zero, all *p*'s < .008.

Aftereffects were calculated, following the usual convention, by subtracting mean reported locations on pre-tests from those on post-tests, and counted as positive when they went toward the visual distracter. They were then averaged across the four test locations and across directions of the visual distracter [2]. The resulting means are shown in Table 4.1.

Of main interest here was whether changing the strong onset cue (i.e., 5 ms) to a very weak one (i.e., 50 ms) would reveal a larger, or even complete, decrease in aftereffects for test frequencies that were different from that of the adapter. The corresponding values can be calculated from the table. When the onset was 5 ms, aftereffects, on average, decreased from 2.64° when adapter and test frequency were the same to 2.00° when they were not, a decrease to 76%. The corresponding values for the 50 ms onset tones were 2.83° and 1.78°, respectively, a decrease to 63%.

The aftereffects were submitted to a 2 (Adapter Frequency: 400 vs. 6400) x 2 (Test Frequency: same vs. different from adapter) x 2 (Adapter Rise Time: 5 vs. 50 ms) x 2 (Test Rise Time: same vs. different from adapter) repeated measures ANOVA. Except for the overall effect, $F(1,16) = 96.12, p < .001$, only the main effect of Test Frequency was significant, $F(1,16) = 57.42, p < .001$. There was a marginal interaction between Adapter Rise Time and the Test Rise Time, $F(1,16) = 4.14, p = .059$.

The bottom row of Table 1 shows the values that correspond to the significant main effect of Test Frequency. For the different frequency conditions (M = 1.91°), aftereffects were in the order of 70% of the frequency same conditions (M = 2.73°).

4.4 Discussion

The present experiment shows that the a steep onset of the sound stimulus is not responsible for the substantial generalization of the aftereffects of ventriloquism across sound-frequencies. When both the adapter and test stimuli had short rise times the aftereffect for the different frequencies was 76% of that for the same frequency, and the corresponding value for long rise times was only somewhat, but non-significantly, smaller at 63%.

We are still left with the contrasting findings of Recanzone (1998) and Lewald (2002). We have already argued that a reason for Recanzone's no-generalization result could be the small scope of his sample (only 3 participants and a single adaptation direction). This, however, cannot be the explanation for Lewald's results since he used a (somewhat) larger sample size and both adaptation directions. Across a series of five experiments involving a total of 78 participants we have studied exposure duration (Frissen et al, 2003, exp. 2), the modality to which attention was allocated during exposure (Frissen et al., 2003, exp. 3), and the rise/fall times of the auditory stimuli (present study) as possible sources for the discrepancy. However, none of these factors appeared to be responsible for the discrepancy.

Our own studies have consistently shown substantial generalization across sound-frequencies. Small generalization decrements were apparent only when the distance between the adapter and test frequency was as large as four octaves (Overall, aftereffects for the different frequency were in the order of 70% of those for the same frequency in both the Frissen et al., 2005, and the present study). Recently, Woods and Recanzone (2004) also found generalization (in the order of 50%) across a two-octave distance in macaques monkeys, although it was not recognized as such by the authors. Together they strongly support the conclusion that the

---

[2] Previous studies have shown these factors not to be of importance.

functional locus of the recalibration is central to the peripheral interaural auditory localization mechanisms.

Such a central locus makes good functional sense. The strength of *cross-modal integration* lies in the reduction of the consequences of modality-specific variability (de Gelder & Bertelson, 2003; Bertelson & de Gelder, 2004). Such variability reduction also occurs on a modality specific level through the combination of cues that provide (redundant) information about an object's particular feature (e.g., its location). The integration of cues from low-level auditory mechanisms like ITDs and ILDs into a unified auditory directional percept is an example of such *intra-modal integration*. Likewise, in the visual system several depth cues are combined in order to increase sensitivity to distance (e.g., Epstein, 1975). Cross-modal integration presumably takes place after intra-modal integration, and therefore the influence of the visual information on auditory localization takes effect *after* the formation of the auditory (spatial) percept. Such an organization is clearly more economical than one where visual information needs to be combined with all separate lower level localization mechanisms.

Chapter 5

Aftereffects Generalize to Non-Octave Intervals

5.1 Introduction

The series of experiments reported in Chapters 2 through 4 have consistently shown that the aftereffects of ventriloquism generalize to a large extent across sound-frequencies. A number of potential modulating factors were investigated across the series. However, aftereffects remained unaffected by the factors under investigation, such as exposure duration or the modality to which attention was allocated during exposure (Chapter 2), the test frequency range (Chapter 3) and the rise/fall times of the auditory stimuli (Chapter 4). A major conclusion drawn from all this was that the locus of visual recalibration of auditory localization is beyond the level of the peripheral, frequency specific, sound localization mechanisms. We argued instead that vision asserts its influence only *after* intra-sensory integration, which is necessary to establish the auditory spatial percept, has been completed (e.g., Bertelson & de Gelder, 2004).

The present set of experiments examines yet another factor that might affect generalization across sound-frequencies. In our previous studies, but also those by Lewald (2002) and Recanzone (1998), test frequencies were always at octave intervals to the adapter frequency. Octaves are fundamental to music perception and to the auditory system in general (Dowling & Harwood, 1986). Perhaps the "harmonic" relation between the tones might have contributed to the extensive generalization.

There is some evidence in the classical conditioning literature that supports such an intuition. Humphreys (1939) measured the amount of sensory generalization of the galvanic skin response across sound-frequencies with a 1967 Hz tone as the (conditioned) training stimulus. Interestingly, there was a *larger* amount of generalization to a test tone that was an octave below (i.e., 984 Hz) the training frequency than to a test tone that was closer to the training tone (i.e., 1000 Hz). This "octave effect" was replicated in rats by Blackwell and Schlosberg (1943), which means that the effect is not due to any musical training. Moreover, the authors argue that it was improbable that the effect could be attributed to any learning due to sounds heard in the rats' "normal life" environment. They conclude that the sensitivity to octave intervals is based on some fundamental organization of the auditory neural pathways (Blackwell & Schlosberg, 1943). Octave generalization has, recently, also been demonstrated in rhesus monkeys (Wright, Rivera, Hulse, Shyan, & Neiworth, 2000).

If these presumed octave sensitive neural mechanisms are involved in the aftereffects of ventriloquism, then this could have one of two consequences for the interpretation of the generalization results found so far. First, in the strong version, generalization was entirely based on the octave sensitive mechanisms. This would mean that generalization across sound-frequencies was in fact generalization across octaves. When tested we would expect to find no aftereffects at non-octave intervals and very steep drop offs around the octave intervals. Such a finding would have major consequences for the locus of adaptation. Second, in the weak version, octave sensitive mechanisms only *contributed* to the generalization. In this case we expect smaller aftereffects at intermediate frequencies.

This chapter presents the results of two experiments designed to explore these possibilities. Both follow the same procedure as that used by Frissen et al. (2005, i.e., Chapter 3) except for the test tone frequencies used. The first experiment examines generalization to test frequencies that either are or are not at an octave interval with the adapter. The second experiment examines generalization within a one-octave interval.

5.2 EXPERIMENT 1: Generalization across and between octaves

The first experiment was designed to examine whether generalization across sound-frequencies is specific to stimuli that are at octave intervals of the adapter frequency. The adapter was always a 400 Hz tone, and aftereffects were measured at that frequency, and at one and two octave intervals plus at three additional frequencies that were harmonically unrelated.

5.2.1 Method

Participants

Fifteen first year students of Tilburg University participated in the study. All were naïve as to the purpose of the experiment. They had normal hearing and normal or corrected to normal vision. Each completed two sessions, one for each direction of auditory-visual discordance.

Apparatus and Material

Testing was carried out in the same room and with the same equipment as described in Chapter 2. Briefly, the setup consisted of a semi circular array of six, from vision occluded, display units and an array of 108 push buttons. The participant's head was fixed by a height-adjustable chin rest.

Auditory stimuli were 400 ms pure tones with 10 ms linear rise/fall times presented at 66 dB(A). So-called "harmonics" were tones at 400, 800, and 1600 Hz, and the so-called "non-harmonics" were tones at 520, 1040, and 2080 Hz. The frequency of the latter tones was 1.3 times that of the corresponding harmonics, which ensured that they were harmonically nor musically related (Burns & Ward, 1982). The adapter was always the 400 Hz tone.

Procedure

The experiment was run in two counterbalanced sessions, one for each direction of the visual distracter. A session consisted of one block of 168 pre-test trials, followed by 7 adaptation-post-tests blocks. On both pre-tests and post-tests, a sound at one of the six test frequencies was presented from one of the four loudspeakers. The participants task was to press the pushbutton corresponding to the apparent direction of the sound's 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 bimodal adapter pair, with the sound 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), producing an 18 deg auditory-visual spatial discordance. Adaptation trials followed each other at a one-per-sec speed. They were followed by 24 randomized post-tests trials, one for each test frequency from each loudspeaker.

5.2.2 Results

Aftereffects were calculated by subtracting mean reported locations on pre-tests from those on post-tests. They were counted as positive when they went in the direction of the visual distracter. They were then pooled across the four test locations and direction of the visual distracter.

The solid squares in Figure 1 show the mean aftereffects of Experiment 1 as a function of test frequency, plotted separately for each harmonic type. The open circles represent the results of Experiment 2, to which we turn later. The dotted vertical line indicates the adapter frequency.

A number of observations can be made. First, substantial aftereffects were obtained irrespective of test frequency. Second, aftereffects were largest at the adapter frequency (4.1 degrees), then dropped and leveled off quickly to values between 3.0 and 3.5 degrees for the other frequencies. Third, aftereffects at non-harmonic frequencies appeared to be somewhat smaller than the aftereffects at harmonic frequencies.

These observations were confirmed by statistical analysis. Separate corrected t-test showed that all aftereffects were significantly greater than zero (all *p*'s < .001). The aftereffects were submitted to a 2 (Direction of Visual Distracter: to the left vs. to the right) x 3 (Test Frequency: first, second, third) x 2 (Harmonicity) repeated measures ANOVA. Except for the overall effect, F(1, 14) = 81.76, p < .0001, only the main effect of Harmonicity was significant, F(1, 14) = 5.46, p = .035.

Despite the non-significance of the interaction between Test Frequency and Harmonicity, inspection of Figure 1 suggests that the main effect of Harmonicity is due not to a decrease for the non-harmonics but to an overall larger effect for the 400 Hz test tone (i.e., the adapter frequency). This possibility was tested by reevaluating the main effect of Harmonicity with the exclusion of the 400 Hz aftereffect. The difference between the pooled aftereffects of the 800 and 1600 Hz harmonic tones (M = 3.38°) and of the three non-harmonic tones (M = 3.07°) was very small and did not reach significance in a one-tailed t-test, *t*(14) = 1.55, p = .071.

5.2.3 Discussion

The first experiment examined to what extent the aftereffects of ventriloquism generalize to frequencies that were at non-octave intervals to the adapter frequency. Analysis revealed a statistically significant main effect of harmonicity, reflecting smaller aftereffects at non-octave intervals. Although the statistical effect was mainly ascribed to the larger aftereffects at the adapter frequency itself (i.e., at 400 Hz), a general pattern emerged where aftereffects at non-octave frequencies were always smaller than those at octave intervals.

The effect of harmonicity then validates the notion that the auditory system's special sensitivity to octaves (Dowling & Harwood, 1986) plays a role. Nevertheless, the actual effect of harmonicity was very moderate at best (the overall difference between aftereffects at octave and at non-octave intervals was approximately .6°). In other words, there was still substantial generalization to non-octave intervals, in the order of 85%. Therefore it is concluded that, although harmonicity seems to play a role it is unable to completely explain the generalization across frequencies results from previous studies.

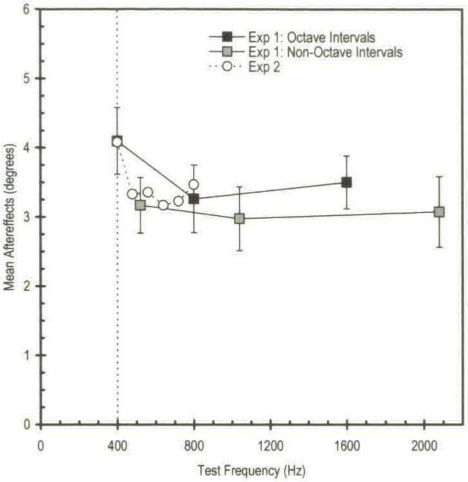


Figure 5.1 Mean aftereffects found in Experiment 1 and 2.

5.3 EXPERIMENT 2: Generalization within one-octave

Inspection of the curves of figure 5.1 shows that for as far as there is a decrease in aftereffects, most of it occurs within the first octave interval, between 400 and 800 Hz. This second experiment was designed to examine more closely the steepness of the generalization gradient within this range. All was the same as for Experiment 1 except for the test frequencies used. The adapter was again the 400 Hz tone and aftereffects were determined for a number of test frequencies ranging from 400 to its octave 800 Hz in 80 Hz steps.

5.3.1 Method

Participants

Five new students from the same pool participated in the study. All were naïve as to the purpose of the experiment. They had normal hearing and normal or corrected to normal vision. Each completed two sessions, one for each direction of auditory-visual discordance.

Stimuli

Auditory stimuli were again 400 ms pure tones with 10 ms linear rise/fall times presented at 66 dB(A). The adapter frequency was again 400 Hz and the test frequencies were 400, 480, 560, 640, 720, and 800 Hz.

5.3.1 Results

The results are represented by the open circles in Figure 1. A number of observations can be made. First, substantial aftereffects were obtained at all the test frequencies. Second, the mean aftereffects at all the intermediate intervals did not differ from those at 800 Hz. Third, the size of the aftereffects at 400 Hz (M = 4.08°) and 800 Hz (M = 3.47°) were very similar to those in Experiment 1 (4.10° and 3.26°, respectively). But of primary interest here is the finding that aftereffects were again largest at the adapter frequency and drop off immediately within the first test interval (i.e., between 400 and 480 Hz).

The first observation was confirmed by separate uncorrected t-tests, all p's < .05. Given the small number of participants no correction was applied to prevent too much loss of statistical power. A 2 (Direction of Visual Distracter) x 6 (Test Frequency) repeated measures ANOVA revealed that, except for the overall effect, $F(1, 4) = 34.24$, $p < .005$, none of the terms were significant, all F's < 1.

5.3.3 Discussion

The second experiment examined the generalization of aftereffects within a one-octave range from the adapter frequency to more finely determine the tuning of the initial large drop-off in the size of aftereffects apparent in Experiment 1. What was immediately clear, as can be seen nicely in Figure 1, is that the results are in complete agreement with those of the first experiment. More pertinent here, however, is that the largest drop in aftereffects is in the first interval from the adapter frequency. Most of the decrement in this case occurs within 20% of the original adapter frequency (at an approximate slope of 4°/octave) and then levels off from there.

Several interesting questions arise from this experiment. For instance, could the drop-off be even steeper? To determine this, aftereffects for test frequencies even closer to the adapter than used here need to be looked at. Also, since test frequencies were always above that of the adapter frequency, one can wonder whether the drop-off is similarly steep for frequencies *below* that of the adapter.

5.4 General Discussion

In the Introduction it was suggested that neural mechanisms that are specifically sensitive to octave intervals (Blackwell & Schlosberg, 1943; Woodworth & Schlosberg, 1954) could either completely explain (strong version) or at least have contributed to (weak version) the substantial generalization across sound frequencies found in the previous studies (chapter 2 through 4).

Both experiments show substantial generalization to test stimuli at non-octave intervals to the adapter frequency, and the first Experiment revealed a significant effect of harmonicity in that aftereffects were somewhat smaller for test frequencies at non-octave intervals. Taken together the results do not support the strong version of the hypothesis that generalization is based entirely on octave sensitive neural mechanisms since in that case aftereffects at non-octave intervals should have been null. The results, on the other hand, do support the weaker notion that octave sensitivity *contributed* to the results.

In addition, the second experiment showed that most of the (overall small) decrement in aftereffect with increasing distance in frequency from the adapter occurs already very close (within 20%) to the original adapter frequency and seems to levels off from there. The functions then describe a negatively accelerated generalization decrement, which is typical for generalization gradients in general (Shepard, 1987).

One potential concern was that the speakers used in the experiments might have produced extraneous energy at non-intended frequencies and that generalization is in fact artefactual and simply based on such impurities. Post-hoc spectral analyses (detailed in §9.2.1.1) of the stimuli as they reach the position of the ear, however, did not support this.

Chapter 6

The Aftereffects of Ventriloquism: Patterns of Spatial Generalization from Local Recalibration [1]

---

[1] Bertelson, P., Frissen, I., Vroomen, J., & de Gelder, B. (2004). *The aftereffects of ventriloquism: Spatial Generalization from Local Recalibration.* Manuscript submitted for publication.

6.1 Abstract

We examined how visual recalibration of apparent sound location obtained at a particular location generalizes to untrained locations. Participants pointed toward the origin of tone bursts scattered along the azimuth, before and after repeated exposure to bursts in one particular location, synchronized with point flashes of light a constant distance to their left/right. Adapter tones were presented straight ahead in Experiment 1, and in the left or right periphery in Experiment 2. With both arrangements, generalization followed different patterns depending on whether it was measured on the visual distracter's side of the auditory adapter, or on the opposite side. On the distracter side, recalibration generalized following a descending gradient, while practically no generalization was observed on the other side. This dependence of generalization patterns on the direction of the discordance imposed during adaptation has apparently never been reported, perhaps because the experimental designs in use did not allow its detection.

6.2 Introduction

A substantial part of behavioral research on crossmodal interaction has focused on relations between auditory and visual processing of related inputs. As for the equally well-documented case of the interactions between proprioception and vision, studied mostly with prismatic rearrangement, the main approach has been based on the imposition of experimental conflict between the information provided in the two modalities concerning one aspect, or dimension, of the inputs (Howard, 1982; Welch, 1978).

The most extensively studied auditory-visual conflict is the one concerning spatial location. When auditory and visual stimuli such as tone bursts and light flashes are presented synchronously but in different locations, the apparent location of the auditory stimulus is typically shifted in the direction of the visual stimulus (Bermant & Welch, 1976; Bertelson & Aschersleben, 1998; Bertelson & Radeau, 1981; Bertelson, Vroomen & de Gelder, 1997; Bertelson, Vroomen, de Gelder & Driver; 2000; Bertelson, Pavani, Ladavas, Vroomen & de Gelder, 2000; Hairston, Wallace, Vaughan, Stein, Norris, & Schirillo, 2003; Radeau, 1992; Radeau & Bertelson, 1987; Thomas, 1941; Vroomen, Bertelson & de Gelder, 2001b; for reviews see Bertelson, 1999; Bertelson & de Gelder, 2004; Welch, 1999; Welch & Warren, 1980). This so-called *visual bias of perceived sound location* generally represents only a fraction of the sound-flash distance. It can nevertheless be sufficient to bring the perceived discrepancy below the detection threshold, which could explain the illusion created by performing ventriloquists that the speech they produce without visible articulation comes from a synchronously agitated dummy. The term *ventriloquism* has, for that reason, come to be used to designate collectively all manifestations of auditory-visual spatial interaction (Bertelson, 1999; Howard & Templeton, 1966). When it was measured also, the converse effect, the auditory bias of visual location, was small but nevertheless reached significance in some studies (Bertelson & Radeau, 1981; Radeau & Bertelson, 1987; Warren, Welch, & McCarthy, 1981; see Radeau, 1985, for a negative result).

Apart from the immediate, or on-line, effects represented by crossmodal biases, ventriloquism also manifests itself through off-line aftereffects (AEs), by which the apparent location of test sounds, presented unimodally after a period of exposure to spatially incongruent sound-flash pairs, is displaced, in relation to pre-exposure tests, in the direction of the preceding visual competitors (Canon, 1970; Frissen, Vroomen, de Gelder & Bertelson, 2003, 2005; Lewald, 2002; Radeau, 1973, 1992; Radeau & Bertelson, 1969, 1974, 1976, 1977, 1978; Recanzone, 1998; Zwiers, Van Opstal & Paige, 2003). The converse AEs, post exposure shifts of visual localization in the opposite direction, i.e. toward the auditory competitors, can also be obtained, but are generally of smaller amplitude (Radeau & Bertelson, 1969, 1974, 1976; and see Lewald, 2002, Exp. 1, for a partial replication).

The occurrence of AEs has generally been attributed to a process of perceptual learning, by which the correspondences between stimuli and resulting percepts are *recalibrated,* in both modalities, or at least in one of them, in a way that reduces the existing incongruence. Such recalibrations should play a role in the development or in the later maintenance of crossmodal coordination (e.g. de Gelder & Bertelson, 2003; Held, 1961; Welch, 1978).

Another reason for being interested in AEs is the information they can provide regarding the extent of the interactions caused by intermodal conflicts. Measuring AEs at several stimulus values after exposure to conflict at a particular one can tell us whether interactions involve only the stimuli present during exposure, or also a range of stimuli along the discordance dimension. This research strategy was first put forward by Bedford (1989). Taking the case of the conflict between the seen and felt location of a body part (the traditional object of prism adaptation studies), she considered the possibility that recalibrations might affect whole perceptual dimensions rather than just the particular stimulus (or stimuli) involved in the conflict. She measured finger pointing to visual targets before and after a period spent pointing to prismatically displaced targets. On each exposure trial, the participant received feedback (lighting of an LED) whenever the responding finger entered a critical area around the target. In the main experiment, recalibration achieved at one of three different azimuthal locations was found to generalize entirely across a whole range of test locations (52.5° on either side of straight ahead). Thus, AEs from adaptation carried out at a particular location generalized entirely across a 105° range.

Bedford's finding was consistent with those from earlier studies carried out with different main objectives, but in which there were similarly no significant differences between AEs observed at trained and untrained locations (Baily, 1972; Harris, 1963; Hay, Langdon & Pick, 1971). Quite different generalization patterns were obtained however in two later studies. Ghahramani, Wolpert, and Jordan (1996) had participants point at visual targets at various two-dimensional locations in a horizontal area, with re-mappings imposed, as in Bedford's experiments by visual feedback. To judge from the vector field graphs in the paper, post-exposure shifts tended to be largest at the exposure location and to go down with distance from that location. Field, Shipley, and Cunningham (1999) exposed participants with a task in which they tried to intercept with an unseen finger a falling ball that was visible through laterally displacing prisms for a variable segment of its trajectory. With vision of the ball through a sufficiently long segment, generalization followed a typical diminishing gradient pattern, with a peak AE at exposure location and rapid reductions on both sides of that location. No convincing explanation has been proposed for those diverging data, and in fact the tasks that were used differed on so many dimensions that no easily testable hypothesis presents itself.

The generalization paradigm has rarely been applied to aftereffects of ventriloquism. Four recent studies have considered generalization along the dimension of sound frequency, with again rather diverging results. In two studies (Lewald, 2002; Recanzone, 1998) no generalization was observed over distances of the order of two octaves, while total or near-total generalization was reported in the other two (Frissen, Vroomen, de Gelder & Bertelson, 2003, 2005) over up to four octaves. Reasons for these differences are currently being tested and they will not be discussed here.

In the present study, we examine, for the first time in the ventriloquism literature, generalization across space. Participants pointed to the apparent location of sound bursts delivered in several azimuthal locations, before and after exposure to a series of identical sound bursts in one particular location, each accompanied by a synchronous point flash of light, a constant angular distance to either its left or its right. Adaptation was conducted with the sound in the participant's median plane in Experiment 1, and in two peripheral locations, respectively in the left and the right half space, in Experiment 2. The focus of the study was the kind of spatial generalization pattern that would obtain in these situations.

Since the study was focused on generalization, its feasibility was conditional on obtaining the usual basic adaptation at the exposure location. To obtain it, we simply resorted to procedures, which in our earlier work, as well as that of colleagues, had proved capable of bringing it about. It must on the other hand be clear that we were not trying to answer any particular question regarding the conditions of occurrence of the basic effect. For instance, no particular measures were taken to insure that the participants attended to the visual distracters, earlier results having shown that such attention was not necessary to obtain either visual bias (Bertelson, et al., 2000b) or auditory AEs (Frissen et al., 2003). Had the expected adaptation not occurred, we would simply have had to reconsider our procedures.

Three patterns of generalization were considered possible: a) no generalization, that is recalibration restricted to the locus of adaptation, b) uniform generalization across the azimuth (as found by Bedford, 1989), or c) generalization following a decreasing gradient on both sides of adaptation locus (as found by Field et al., 1999).

6.3 Experimental setup

The testing was carried out in a dark, semi-reverberant and soundproof booth, 4.6 m in length, 2.4 m wide and 2.2 m high. Participants sat in front of a table with their head restrained by a fixed chinrest at ~40 cm above the tabletop. The setup involved seven display units, for presentation of auditory and visual stimuli, and an array of push buttons, to be used by participants in auditory localization tests. The display units, which were hidden behind a black, acoustically transparent cloth, were arranged along a semi-circular array on the same vertical plane as the chin rest and at a 42 cm radial distance from it, one straight-ahead (0°) and the others at 17.5°, 35° and 52.5° left and right of straight-ahead. Each of them consisted of a loudspeaker (Philips, 30 watt wide frequency Box 410, Ø = 9 cm) with a red LED (Ø = 1 cm) over its center. The pushbuttons, a total of 108, were arranged on the table top, along another semi-circular array, at 1° intervals, and 5 cm in front of the display units (thus at 37 cm horizontal distance from the chinrest).

Auditory stimuli, used in pre- and post-tests as well as in bimodal exposure trials, consisted of a single 200 ms burst of 750 Hz pure tone, with 5 ms linear rise/fall envelopes, presented at 66 dB(A). Speakers had a characteristic that was flat within 12 dB between .3 and 15 kHz, with approximately 5 dB/octave roll-off. The reverberation time (measured in the booth with all the experimental equipment in place) for our 750 Hz tone was less than .33 s. The flashes also lasted 200 ms, and their luminance was set at 28 cd/m2, as measured from a continuous light at a distance of 1 m. When delivered (on bimodal exposure trials only) they were clearly visible through the occluding cloth.

6.4 Experiment 1: Adaptation in central location

In this first experiment, adaptation was carried out with the sounds coming from straight-ahead, and the discordant visual stimuli on either of the next display units, 17.5° to the left or the right. Its effects were measured through pre- and post-exposure localization tests with auditory targets at seven equidistant locations, respectively straight ahead (0°), and 17, 5°, 35° and 52, 5° left and right.

6.4.1 Method

Participants

Sixteen students from Tilburg University (age 18-25, twelve 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.

Procedure

Each of the two sessions was ran throughout with the distracter flashes either left or right of the sounds, in balanced order. A session began with 98 auditory pre-tests, 14 from each of the seven loudspeakers, in randomized order. These were initiated by the participant, by pressing a button located in the median plane, 20 cm in front of her/him, and the sound followed after 500 msec. This procedure ensures a constant starting position for all pointing movements. Instructions were always to press the response button closest to the apparent direction of the sound, and no stress whatever was put on response speed. The whole series of pretests usually lasted 5 to 7 min. The session continued with seven adaptation-post-test blocks. Each of these blocks involved 60 bimodal exposure trials, followed by 14 post-tests. On a bimodal exposure trial, the

200 ms sound was delivered in the median loudspeaker simultaneously with a flash in the next display unit, 17.5° to the left or to the right, depending on the session. Participants were told to look at the location in which flashes were delivered. It follows from the preceding description that exposure trials were always identical for the seven blocks of each session. They followed each other at one sec intervals, so that each exposure phase lasted just above one minute (exactly 60, 2 sec). There was a 3.5 sec interval between the end of each exposure phase and the start of the following set of posttests. The 14 post-tests, two from each loudspeaker, in randomized order, were self-paced just like the pre-tests, so that the total duration of a post-tests phase varied between participants. It typically lasted 40 to 60 sec. The following exposure phase was started when the participant again pressed the median button.

6.4.2 Results

Responses were filtered for outliers by discarding, separately for each sound test location, values lying more than 2.5 standard deviations from the mean. These represented 1.3 % of the data. AEs were then calculated by subtracting mean reported locations on pretests (14 values per participant and per sound location) from those on post-tests (2 values per participant and sound location for each block x 7 blocks, making 14 values again). AEs were counted as positive when they went toward the visual distracter.

Figure 1 shows, separately for each direction of discordance (visual distracter to the left vs. to the right), mean AEs measured at the different locations. Two main points of interest emerge.

First, both generalization functions have a peak in the vicinity of the adaptation location (here, straight-ahead). In the figure, the two peaks occur in fact at different points, at the straight-ahead location for leftward discordance, but at the next location to the right of straight-ahead for rightward discordance. This aspect of the results should actually not detain us, for a paired t-test applied to participants' individual peak locations fell short of significance, $t(15) = 1.20$, $p = .25$. Mean peak locations (leftward discordance: 11.9° left; rightward: 1.1° right) are thus not significantly different. Also, both mean locations were not significantly different from the median location, leftward discordance, $t\ (15) = -1.72$, $p = .11$, and rightward, $t\ (15) = .15$, $p = .88$.

Second, the two curves follow asymmetrical courses on the two sides of their respective peaks, and this asymmetry varies with direction of discordance. For leftward discordance, substantial AEs are found for tests carried out in the left half space, and practically none in the right half space. The opposite asymmetry occurs for the rightward discordance. Thus generalization occurs mainly, if not only, in the direction in which sounds were attracted during the preceding exposure.

The data were first submitted to a 2 (Discordance Direction: leftward vs. rightward) x 7 (Test Location) repeated measures ANOVA. The main effect of Discordance Direction (henceforth DD) was non-significant, $F < 1$, but its interaction with Test Location was highly significant, $F(6, 90) = 7.58, p < .0001$, reflecting the opposite asymmetries of the generalization patterns obtained under the two DDs. The main effect of Test Location was also significant, $F(6, 90) = 3.32, p < .01$, but given the strong interaction with DD, the fact has no meaningful implication.

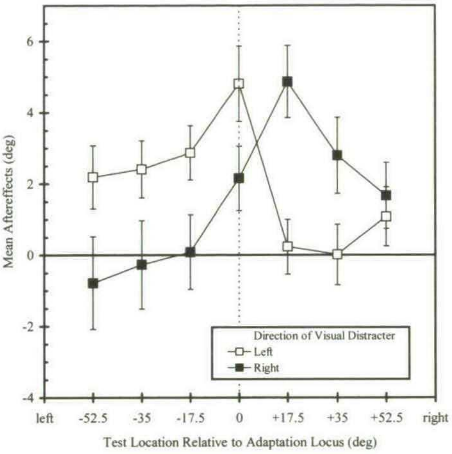


Figure 1. Experiment 1: Adaptation in central location. Mean aftereffects (and standard errors) as functions of test location. During exposure the sounds were always presented at 0° (dotted line) and the visual distracters at 17.5° to the left or to the right depending on the discordance direction.

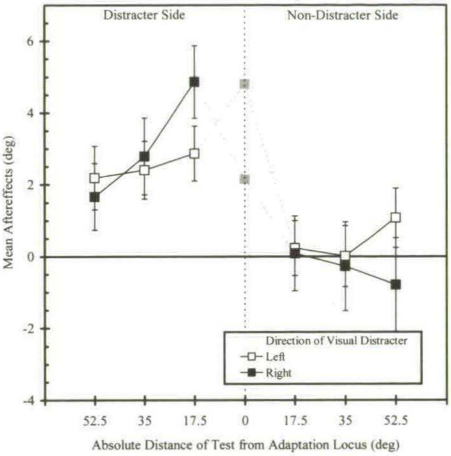


Figure 2. Experiment 1: Adaptation in central location. Same data as in Figure 1, re-plotted as functions of distance from central adaptation locus, measured along distracter side of that locus (left hand part), or on non-distracter side (right hand part). Aftereffects at adaptation locus, which belong to neither category and were not included in the ANOVA, are shown in gray.

To further examine the dependence of the results on DD, the AEs were considered and analyzed in terms of their being measured on the side of the auditory adapter on which the visual distracter was delivered during exposure (henceforth "distracter side"), or on the opposite side ("non-distracter side"). In Figure 2, the mean AEs obtained at generalization (non-central) locations are shown as functions of their distance from display center, on respectively the non-distracter side (on the left) and the distracter side (on the right). The AEs obtained in the central location, which belong to neither category, were not included in the new analysis. They are shown in gray in the figure.

The new ANOVA was a 2 (Test Side: distracter vs. non-distracter) x 2 (DD) x 3 (Distance) repeated measures one. Distances were entered in absolute values (respectively 17.5°, 35.0° and 52.5°). The main effects of Test Side, $F(1, 15) = 10.13, p < .01$, and of Distance, $F(2, 30) = 3.40, p < .05$, were significant, but that of DD ($F < 1$) was not. Among interactions, those between Distance and respectively DD, $F(2, 30) = 3.71, p < .05$, and Test Side, $F(2, 30) = 4.10$,

$p < .05$, were both significant, while that between Test Side and DD was not, nor was the second order interaction (both *F*'s < 1). The Distance by Test Side interaction corresponds to the fact, visible on the figure, that the effect of distance is smaller on the non-distracter side than on the distracter one. The Distance by DD interaction reflects the fact, also visible on the figure, that the effect of distance is present on both Test Sides for rightward discordance, and smaller (or even inverted, on the non-distracter side) for leftward discordance.

6.4.3 Discussion

As expected, exposure to the present form of auditory-visual conflict has produced significant auditory AEs at, or in the vicinity of exposure location. The condition was thus obtained for examining spatial generalization. Here, a surprising pattern occurred. Whatever the direction of discordance, different patterns of generalization occurred on the two sides of adapter sound location. In the two cases, substantial generalization was found on the distracter side, and practically none on the non-distracter side. On the distracter side, generalization followed a gradient that diminished with increasing distance from center, hence presumably from adaptation location. On the non-distracter side, no generalization was found, meaning either that none occurred, or possibly one that went down with distance so rapidly that it had vanished at the first location at which it was examined, 17.5° from adaptation location.

6.5 Experiment 2: Adaptation in the periphery

In Experiment 1 adaptation was induced in the median plane. In this second experiment, spatial generalization was examined for adaptation at more peripheral locations, with the target sound 35° on either side of center. This arrangement allowed the consideration of test locations across a wider angle (87.5°), from adaptation location toward center and beyond. Our main purpose was to test the generality of the findings of Experiment 1 with respect to the generalization gradients, in particular their dependence on DD.

6.5.1 Method

Participants

Fourteen new students from the same pool (age 18-27, eight female), all naïve again as to the purpose of the Experiment, and with normal hearing and normal or corrected to normal vision, participated in four sessions each.

Procedure

One of the four sessions was devoted to each combination of adapter sound location (left vs. right periphery) and DD (visual distracter to left vs. right of adapter sound). Just as in Experiment 1, each session began with 98 auditory pre-tests, 14 from each of the seven loudspeakers, in randomized order, and continued with seven adaptation-post-test blocks. Each of these blocks consisted of 60 bimodal exposure trials, with the sound, depending on the session, at 35° left or right of straight ahead, and the flash from the next display unit, 17.5° to its left or right, followed again immediately by 14 auditory post-tests, two from each of the seven loudspeakers, in randomized order. Other aspects of the procedure, like the pointing instructions and the key-pressing initiation of test trials, was the same as in Experiment 1.

We have seen that generalization was, for each adaptation location, measured on one side at five locations extending from the adapter to the center and then beyond, but on the other side at the single remaining location only. Given our interest in the shape of generalization gradients, the data from these single locations were for each condition excluded from the analysis. The corresponding test trials can thus be considered as fillers.

6.5.2 Results

Outlying responses, amounting to 1.2% of the data, were discarded.

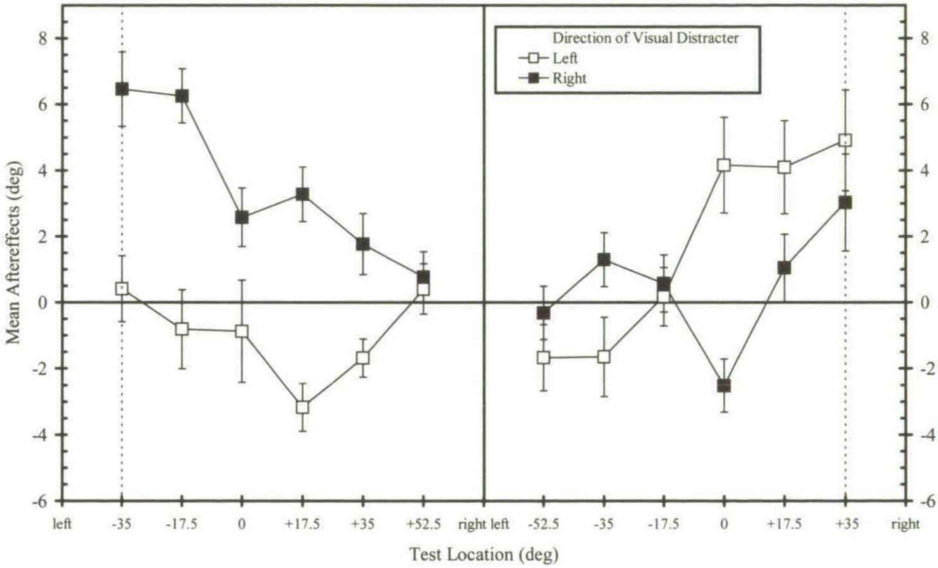


Figure 3. Experiment 2: Adaptation in the periphery. Mean aftereffects (and standard errors) as functions of test location. During exposure, sounds (dotted vertical lines) were presented at 35° left (Left panel) or right (right panel).

Mean AEs per sound test location, computed again by subtracting mean pointed locations on pre-tests from those on post-tests, are shown, separately for the four conditions, in Figure 3. The left panel shows the data for adaptation in the left half space and the right panel for the right half space.

AEs have again clear maxima at adaptation locations, and go down when measured at other locations. On the other hand, both the peaks and the generalization gradients depend on DD. This dependence is strongest for adaptation in the left half space. When the visual distracter is delivered on the right side of the adapter sound, a substantial peak (more than 6°) obtains at adaptation location and AEs go down with increasing distance from peak location, following a quasi-monotonic gradient. With the distracter on the opposite side (i.e., to the left), the starting peak is practically at zero level, and AEs become increasingly negative at the next three locations, before going back to starting level. In the right half space, a similar (though less accentuated) pattern is obtained. AEs are higher at both the adaptation location and the next two locations for the condition with the distracter toward the center (i.e., to the left). With the distracter away from center (i.e., the right), the starting peak is also lower and a final rebound (similar to the one obtained in the left half space) occurs again at the more distant locations.

The differences in peak values between conditions with the adapter on center's side and on the other side were tested with paired t-tests. The difference is significant for adaptation in the left half space, $t\,(13) = 3.59, p < .01$, and not for adaptation in the right hemisphere, $t < 1$.

In order to better illustrate the role of DD, the data were re-grouped in Figure 4, with generalization on the distracter side and on the non-distracter side in separate panels. The similarity of the curves in each panel is very apparent, as are the differences between the panels.

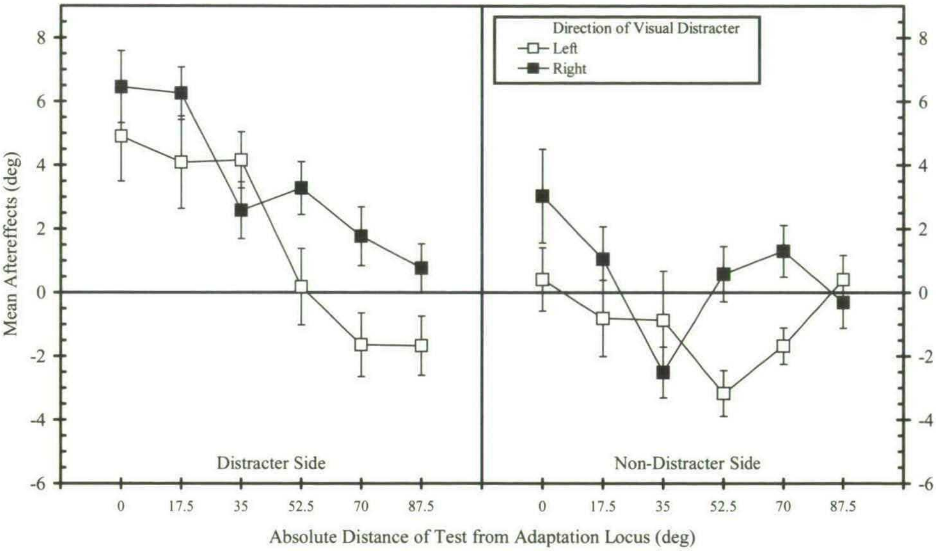


Figure 4. Experiment 2: Adaptation in the periphery. Same data as in Figure 3, re-plotted as functions of absolute distance from adaptation locus. Left panel: adaptation toward center; right panel: adaptation away from center.

The AEs were submitted to a 2 (DD: leftward vs. rightward) x 2 (Test Side: distracter vs. non-distracter) x 6 (Distance) repeated measures ANOVA. Distance was again entered in absolute values (0°, 17.5°, 35°, 52.5°, 70°, and 87.5°). The main effect of Test Side was significant, $F\,(1, 13) = 14.64, p < .01$, and its interaction with DD was not, $F < 1$. The main effect of Distance was highly significant, $F\,(5, 65) = 22.9, p < .001$, but this factor's interactions with DD, $F\,(5, 65) = 4.01, p < .01$, and with Test Side, $F\,(5, 65) = 4.93, p < .01$, were also significant. The main effect of DD fell narrowly short of significance, F (1, 13) = 3.98, p =. 067, and the second-order interaction, F (5, 65) = 1.36, p = .25 was also non-significant.

The effects of distance were further explored through trend analyses carried out separately on the distracter and on the non-distracter side. On the distracter side, the linear component was highly significant, F (1, 13) = 30.9, $p < .0001$, and all higher-order ones were non-significant. The linear trend's interaction with DD (left vs. right) was non-significant, $F < 1$. On the non-distracter side, there was no significant linear component, $F < 1$, only a quadratic one, F (1, 13) = 15.1, $p < .002$. The latter reflects the rebound in both functions. Its interaction with DD was also non-significant, $F < 1$.

6.5.3 Discussion

This new experiment has brought several results calling for comments.

After exposure at each of the two adaptation loci, more generalization occurred when the visual distracter had been presented on the central side of the auditory adapter than on the lateral side. Since the generalization was in both cases measured at locations extending from adaptation locus toward center and beyond, the effect means that, just as in Experiment 1, generalization was stronger in the direction of the former visual distracter than in the opposite direction.

The larger set of locations over which generalization was now measured has provided a picture of generalization more complete than the one obtained in Experiment 1. On the distracter side, that picture is clearly of a gradient diminishing monotonically with distance from adaptation locus. On the other hand, in the two conditions with testing on the non-distracter side, AEs rebounded upward at the largest distances from adaptation location. These rebounds were probably responsible for the significant non-linear trends found specifically in these conditions. There is for the time being no obvious explanation for that particular aspect of the results.

Finally, in the present experiment, DD affected not only the pattern of generalization, but also the adaptation occurring at the adaptation locations themselves. At both these locations, the AE was larger in the condition with the distracter on the central side. This difference might suggests that part of the present influence of DD on generalization may be a consequence of the adaptation occurring at adaptation locus, higher adaptation peaks producing more generalization. This relation could of course not account for the whole of the DD effect, since in Experiment 1 DD created different generalization patterns from the same peak adaptations. Moreover, the effect on adaptation peak only reached significance at one of the two adaptation loci (in the left half space). No strong conclusions can be drawn concerning this particular issue until the data have received a more general confirmation.

6.6 General discussion

As presented in the Introduction, the main purpose of the study was to determine the pattern of spatial generalization following visual recalibration of perceived sound location through ventriloquism. Three patterns were considered possible: uniform extension to all locations, decreasing gradients, or no generalization. The uniform generalization pattern was clearly not found. AEs always peaked around the location at which sounds had been delivered during exposure (straight ahead in Experiment 1 and 35° left or right in Experiment 2), and went down when measured away from that location.

The novel, and completely unexpected, outcome of the study is however that different patterns were observed on respectively the side of the visual distracter, and the opposite side. This *discordance direction (DD) effect* was most clearly demonstrated in Experiment 1, with the auditory adapter straight ahead and the visual distracter on either its left or its right. In both cases, substantial generalization, going down with increasing distance from adaptation locus, was found on the distracter side, and no, or very little, generalization on the non-distracter side. The same effect of DD was observed in Experiment 2. With the auditory adapters in the periphery, generalization could only be measured at locations extending toward the center (and beyond). Clear generalization only occurred in the conditions in which the distracter had itself been located on the same central side of the adapter sound. In these conditions, a clear monotonically decreasing gradient was obtained. When the distracter had been presented on the non-central side, practically no generalization was observed.

Our attribution of the contrasting generalization pattern to auditory-visual discordance has been put into question. Could the effect, we were asked, not simply be due to the fact that the participants were fixating the visual distracter? If the apparent direction of sound sources moved along with the direction of gaze, an effect mimicking the visual bias of sound location would be produced, and be subsequently consolidated as an AE. This proposal actually runs into several difficulties.

First, much recent work has focused on the effects of eye position on sound localization, and for horizontal auditory localization the most frequent result has been a shift *away* from fixation (e. g., Bohlander, 1984; Lewald, 1997, 1998), thus the opposite from the direction necessary to account for visual bias and also for the dependence of its AEs on the position of the visual distracter during exposure. Second, if the manifestations of ventriloquism were simple consequences of fixation on the visual distracter, they would occur equally with synchronized and desynchronized bimodal presentations. Actually, desynchronizing visual and auditory inputs has been shown to eliminate (Bertelson & Aschersleben, 1998) or at least strongly reduce (Bertelson, Vroomen & de Gelder, 1997; Choe, Welch, Gilford & Juola, 1975; Radeau & Bertelson, 1987; Thomas, 1941; Warren, Welch & McCarthy, 1981) the visual bias of perceived auditory location, and also to reduce significantly auditory AEs (Radeau & Bertelson, 1977). Finally, a result that has often been quoted as showing the dependence of adaptation on gaze direction is one reported by Weerts and Thurlow (1971), for an exposure condition in which the participants just looked at a non-changing visual target some distance (20°) to one side, while hearing clicks originating from straight ahead. The authors reported a small but statistically significant mean AE in the direction of the fixated location. This result has however not been replicated by Radeau and Bertelson (1977, Experiment 2) whose participants monitored for occasional lightings an LED, located 20° to one side of a speech-emitting loudspeaker, and displayed no AE whatsoever in speech localization post-tests.

Another critical comment we have received was that the DD effect might have been created by sound reverberations occurring in our experimental booth. The problem with that notion is that in each of our exposure conditions the total acoustic input, main component plus its eventual reverberations, was the same irrespective of the side, left or right, on which the visual distracter was delivered. It seems thus unlikely that very different generalization patterns would be found if the experiment was run in a fully anechoic environment.

So, it appears that the DD effect really tells us something about the spatial extension of auditory recalibration. This recalibration would apply differently in the two halves of the discordance dimension, which, given that the gaze was in all probability directed toward the visual distracter for most of the exposure phases, means the two half visual fields. In the half toward which the target sound is moved, the shift would extend to other locations along the dimension, with a strength that decreases with distance. In the other half, there would be no generalization (or, as already mentioned, one that decrease too fast to be still visible at the smallest 17.5° distance considered in our study).

An obvious question is whether this generalization pattern reflects a specific constraint of auditory localization recalibration, or one that applies also in other modalities. As we have seen in the Introduction, the only other case of spatial recalibration for which patterns of generalization have been examined is the visual recalibration of proprioception. Unfortunately, nearly all the studies in which this was done were carried out with a single DD, so that the only available source of information about possible DD effects was the comparison between AEs obtained at corresponding eccentricities in the two half spaces. These AEs can have been influenced, beyond DD, by specific characteristics of the locations at which they had been measured, like for instance local susceptibility to recalibration. For example, in Field et al.'s (1999) experiments with the falling ball task, which were run with rightward prismatic displacement only, symmetrical gradients occurred on the two sides of exposure location. This result might suggest that DD was not an effective factor in that situation. There is however a possibility that a real DD effect, producing in the case stronger generalization in the right half space, happened to be counterbalanced by a higher susceptibility to recalibration of points in the left half space. To take the apparently opposite kind of results, Gharamani et al. (1996) showed some results that, although the authors did not mention it, might have been related to DD. In one of their conditions, which involved remapping of pointing toward the body along the sagittal axis, generalization (as judged from their Figure 7b) occurred only on the re-mapping side of adaptation locus. However, in the absence of data for remapping in the opposite direction, the result could here also reflect some local differences in susceptibility to recalibration as well as a DD effect. The only way to effectively rule out contamination of generalization results by local factors is thus to carry out all recalibrations, as was done in the present study, in two opposite directions. The effects of changing DD can then be measured in exactly the same locations. Until such controls have been applied, our question regarding DD effects in other cases than auditory recalibration receives no answer.

Chapter 7

Evidence for Generalization Across Sound-Frequency and Spatial Generalization Without a Directed Motor Response

7.1 Introduction

In all of the experiments in the preceding chapters sound localization was assessed by having participants point at the apparent location of the sound source. In the present chapter two experiments are described that use a localization task that did not involve such a pointing response. There were three motives for this. First, in Chapter 1 (§1.3.4.1) it was argued that as long as no pointing response was required *during the exposure period* there is no reason to expect from it any inadvertent effects (e.g., Radeau & Bertelson, 1974). Nevertheless, a direct demonstration that aftereffects can be measured without any pointing at all was felt necessary. Second, it allowed a replication with a different method of previous results. Third, the task allows us to assess the involvement of the posterior partietal cortex (PPC; Recanzone, 1998). Recall from §1.4.3 that the PPC is heavily involved in representing sensory targets that are going to be the object of future motor actions such as reaching for it. If the PPC (or better, its subpart the parietal reach region) is a major site for recalibration than sidestepping this structure through the use of a non-pointing localization task (which in our case is of course very similar to reaching) should substantially affect the results.

The major methodological change that is introduced is the manner in which participants localize the sound source. Instead of pointing with their hands they indicate its location *relative to a preceding visual reference point.* On each localization trial a 750 ms LED flash is given in a certain spatial position followed by the auditory target. The participants engage in a 2AFC task and indicate whether the sound was to the left or to the right of the reference location. The trials of interest are those for when the target was presented in the same location as the reference. Theoretically, the proportion of, say, right responses is 50%. The measure of interest here is whether this proportion shifts as a result of exposure to the auditory-visual discrepancy. If the direction of discordance was to the right we expect an increase in 'right' response and a decrease when it was to the left.

Two Experiments were run that applied this relative localization method. The first Experiment was concerned with generalization across sound-frequencies while the second was concerned with spatial generalization.

7.2 Experiment 1: Generalization across sound-frequencies

In this experiment participants were adapted to an auditory-visual spatial discrepancy of 9° with either a 750 or a 3000 Hz tone, and tested on sound localization with either tone.

7.2.1 Method

Participants

Twelve first year students from Tilburg University participated in this study. They were naïve to the purpose of the study and all had corrected, or corrected to normal vision and normal hearing.

Apparatus and material.

The testing was carried out in a dark soundproof booth. The setup consisted of three display units which were occluded by means of a black, acoustically transparent cloth. Display units, each of which consisted of a loudspeaker with a red LED over its center, were arranged along a semi-circular array at 85 cm from the participant's head, at eye level and at –9°, 0°, and +9° distance from centre in the horizontal plane.

The auditory stimuli were 200 ms pure tones (with 5 ms linear rise/fall envelopes) at either 750 Hz or 3000 Hz.

Procedure

The experiment was run in two consecutive sessions, one for each direction of the visual distracter. A session started with an exposure phase which consisted of 60 presentations of the auditory-visual spatial discrepancy at a rate of 1 presentation per second. The sound was always presented from the central location and the visual distracter from the adjacent unit to the left or to the right of the sound, depending on the session, thus creating a 9° spatial discrepancy. To ensure that participants attended the display an occasional catch trial (10%) was introduced. On such a trial the LED was omitted and the participants were to immediately and verbally report on such an occurrence. Next followed 120 randomized posttest localization trials, 20 repetitions of each test frequency per speaker location. In order to prevent spontaneous decay of the aftereffect a brief 'top-up' or readaptation of six presentations of the discrepancy was given before each posttest.

Posttest trials were self paced and started 1500 ms after the top-up and consisted of first a 750 ms flash of the LED at one of the three test locations, followed after another 750 ms by the test tone. The participant indicated the location of the test sound as being either to the left or to the right of the visual reference by pushing one of two dedicated buttons. Unknown to the participant the sound was always presented from the same location as the reference.

Half the participants were adapted to the 750 Hz tone and the other half adapted to the 3000 Hz tone.

7.2.2 Results

The proportions of responses that were congruent with the direction of discordance were determined. Aftereffects, shown in table 7.1, were calculated by subtracting .5 from these proportions, which represents the expected pretest performance. Values larger than zero are representative of a shift in auditory localization toward the visual distracter.

Table 7.1 Experiment 1: Mean aftereffects (standard error) for each combination of adapter and test frequency.

		Test Frequency (Hz)	
		750	3000
Adapter	750	.34 (.05)	.35 (.05)
Frequency (Hz)	3000	.30 (.03)	.22 (.05)

From the table two observations can be made. First, substantial aftereffects were obtained across all conditions. Second, and more important, aftereffects were of equal magnitude irrespective of whether the adapter and test frequency were the same (mean = .27) or not (mean

= .33). In other words there was (complete) generalization of aftereffects across sound-frequencies.

These observations were confirmed by statistical analysis. The values were submitted to a 2 (Adapter Frequency: 750 vs. 3000 Hz) x 2 (Test Frequency: Same vs. Different) repeated measures ANOVA with Adapter Frequency as a between subjects factor. Only the overall effect reached significance, $F(1, 10) = 783.68$, $p < .0001$. The main effects and their interaction were not significant, all p's > .16.

7.2.3 Discussion

Two (related) points can be made about the present results. First, they replicate the findings of chapters 2 through 5. Aftereffects were obtained irrespective of whether the adapter and test frequencies were the same or not. Thus there was again significant generalization of recalibration across sound-frequencies. The theoretical relevance of this finding has already been addressed several times in the mentioned chapters.

Second, the results were obtained using a localization task that is completely different from that used in the previous studies. Participant indicated, in a 2AFC task, the apparent location of the sound source relative to a visual reference. Thus, throughout the experiment no pointing was required at all. As was argued in §7.1, there is no a priori reason to expect any effects of exposure on pointing since during the exposure period the participant does not engage in any pointing activity, and even if were the case, no explicit performance feedback was available. The present experiment explicitly demonstrates this reasoning to be valid, and therefore that the patterns of generalization found thus far can not be attributed to the particular response mode used.

7.3 Experiment 2: Generalization from local recalibration

When studying the spatial extension of recalibration (Chapter 6) we discovered a remarkable asymmetry in the generalization functions. Briefly, generalization was found only on the side where the distracter was during adaptation. No aftereffects, however, were found on the opposite side of the locus of adaptation. In addition, aftereffects decreased when distance from the location of adaptation increased.

Here we tried to replicate these findings using a version of the relative localization method. There were some procedural differences with Experiment 1, which are pointed out here. First, a pretest was performed to serve as a reference against which posttest performance was compared. Second, on pre and posttest two additional locations per reference location were probed. These were the two adjacent speaker positions on either side of the reference. This was mainly done to present the participant with a task that was not always at threshold level as was the case in Experiment 1 (i.e., tests here were always performed on the reference location only). When the test sound was presented on one of the adjacent locations, performance was typically very high (i.e., sounds were easily localizable relative to the visual reference). As for the analysis of the data, again only those trials were considered when the test sound came from the reference location. Third, several blocks of posttest trials were now interspersed between larger, one minute blocks of exposure trials.

7.3.1 Method

Participants

Twenty first year students from Tilburg University participated in six sessions each. They were naïve to the purpose of the study and all had corrected, or corrected to normal vision and normal hearing.

Apparatus and material

The testing was carried out in a completely dark soundproof booth. The participant's head was fixed by means of an adjustable chinrest. The same display units as in Experiment 1 were used but their number was increased to seven, and were occluded by means of a black, acoustically transparent cloth. They were arranged, at eye level, along a semi-circular array at 85 cm from the participant's head, ranging from -27° to +27° at 9° intervals.

Procedure

The combination of three adaptation loci (-18°, 0°, or +18°) and the direction of the visual distracter (to the adapter sound's left or right) produced six conditions. They were tested in six sessions, run in a quasi random order such that in consecutive conditions both the locus and direction of discordance was different from the preceding one. Between sessions there was short break that was spend outside of the laboratory setting.

Each session consisted of a pretest and four adaptation-posttest blocks. On pretests and posttest trials a visual reference was flashed for 750 ms in one the three adaptation loci followed after 750 ms by a test tone in one of the three locations closest to the reference (i.e., 9° to it's left or right, or at the location of the adapter itself). The participants task was to indicate the location of the test tone relative to the visual reference, by pressing one of two dedicated buttons on a response box. In this manner each test location was probed for a total of twelve times. To prevent effects of speaker familiarity, the sound level of the test tone varied within a 2 dB range.

Each of the four blocks of exposure trials consisted of 60 presentations of the condition's particular auditory-visual spatial conflict at a 1 per second rate. The sound was presented from a single location, which, depending on the condition, was straight ahead or 18° to the left or to the right of straight ahead. The visual distracter was 18° to the left or right of the sound. Each block of exposure trials was immediately followed by 27 posttest, three for each of the nine test locations.

7.3.2 Results

Aftereffects were calculated by subtracting the proportion of auditory-visual discordance congruent responses on pretests from those on posttests. Thus, if the visual distracter was to the left of the sound during exposure, the difference between the proportion of left responses in the pre and posttest were calculated, and vice versa. The resulting functions are plotted in figure 7.1. The three panels, from left to right, represent the generalization functions for when the adapter sound was in the left hemispace, the straight ahead location, or in the right hemispace, respectively.

The overall impression of the results is that the patterns of spatial generalization are very similar, even in morphology, to those in chapter 6. First, aftereffects were generally largest at the locus of adaptation and decreased with increasing distance from this locus. This is nicely demonstrated by, for instance, the curve in the leftmost panel corresponding to when the visual distracter was to the right (black squares).

Second, there again appears to be a discordance direction effect. Aftereffects are found to generalize to the side occupied by the visual distracter during exposure, whereas no aftereffects at all were found in the locations opposite. This is most clearly visible in the middle panel. Consider the curve showing the results for when the visual distracter was to the right (i.e., black squares). Aftereffects are found in the location of adaptation and to its *right*. To the left of the location of adaptation aftereffects are non-existent. The exact same pattern of generalization can be seen for when the visual distracter was to the left, and in fact in all other curves.

Third, of potential interest is that the two functions obtained from recalibration in the left hemispace are further apart than those in the right hemispace, as was also the found with the pointing localization task in chapter 6.

For the statistical analysis, aftereffects were considered separately for the central and the two peripheral adaptation locations. For the central location, a 2 (Direction of Visual Distracter: To the left or to the right) x 3 (Test Location) repeated measures ANOVA was run. Neither main effect was significant but their interaction was, $F(2,38) = 3.57, p < .05$, reflecting the opposite trends of the curves. On the aftereffects from exposure in the two peripheral locations a 2 (Direction of Visual Distracter) x 2 (Location of Adaptation: Left or right hemispace) x 3 (Test Location) repeated measures ANOVA was conducted. There was a significant interaction between Direction and Location of Adaptation, $F(1,19) = 14.10, p < .01$, which reflects the

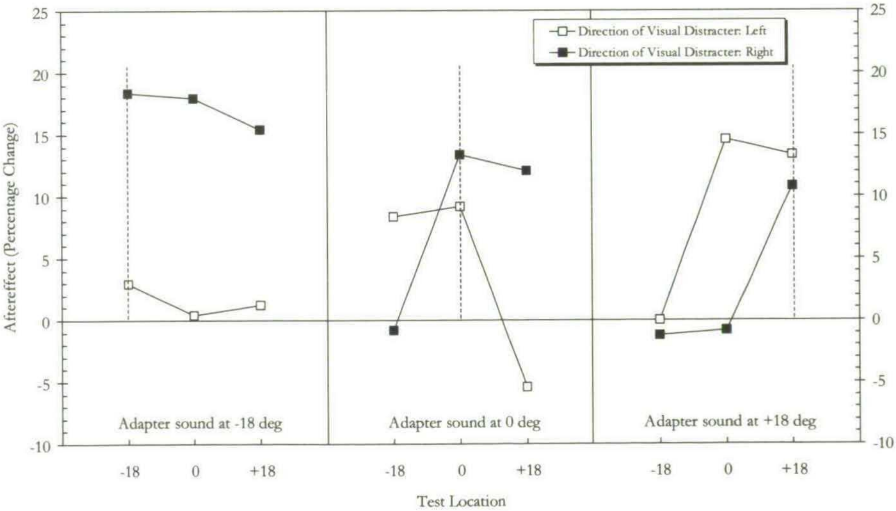


Figure 7.1 Patterns of spatial generalization of aftereffects following local recalibration in the left or right hemispace (left and right panels, respectively), and following recalibration in the median plane (middle panel). The dotted lines represent the location of the adapter sound during exposure.

"discordance direction effect" (Chapter 6), and there was a marginally significant interaction between Location of Adaptation and Test Location, $F(2,38) = 2.51$, $p < .1$, which reflects the opposite directions of the decreasing gradients in the two hemispaces (i.e., going down to the right in the left hemispace and vice versa).

As in chapter 6, the results were further examined for their dependence on the discordance direction. Thus the aftereffects were considered in terms of their being measured on the side on which the visual distracter was presented during exposure ("distracter side"), or on the opposite site ("non-distracter side"). For adaptation in the central location the aftereffects corresponding to the 0° position were not included in the analysis (see §6.4.2). Thus, a 2 (Distracter Side) x 2 (Direction of Visual Distracter) repeated measure ANOVA revealed a significant main effect of Distracter Side, $F(1,19) = 5.50$, $p < .05$. The remaining terms were non-significant (both F's < 1). For the two peripheral adaptation locations a 2 (Distracter Side) x 2 (Direction of Visual Distracter) x 3 (Distance from Location of Adaptation) was run. There was a significant effect of Distracter Side, $F(1,19) = 14.10$, $p < .01$, and a marginally significant effect of Distance, $F(2,38) = 2.62$, $p < .1$.

7.3.3 Discussion

The purpose of the present experiment was to test the generality of the spatial generalization patterns found in Chapter 6, by applying a different localization method that does not require a directed motor response from the participant. The results are in striking agreement with those in the previous Chapter which argues in favor of their generality.

Especially the surprising direction discordance effect was again apparent, with generalization being restricted to those locations that are in the direction of the visual distracter and none in the opposite direction. This excludes the use of hand pointing as a possible cause for the asymmetry.

The morphological resemblance of the generalization functions between chapter 6 and the present ones would suggest that the left hemispace is more susceptible to recalibration than the right one. There is little evidence in the literature that would support such conjecture. In fact, the little available evidence would suggest that auditory localization in the left hemispace is better than in the right (Burke, Letsos, & Butler, 1994), although when controlled for the number of front-back reversals (e.g., Wightman & Kistler, 1999) the left side advantage was no longer significant.

7.4 General discussion

The experiments in this chapter addressed the same questions as those in previous chapters. That is, do aftereffects generalize across sound-frequencies, and what are the generalization patterns from local recalibration. A localization task was used that does not require the participant to point at the apparent sound source but instead to judge it location relative to a visual reference. Both these experiments turned out to be replications of their "pointing" counterparts in the preceding chapters.

The change in localization task *did not* have any effect on the patterns of generalization. Substantial generalization across sound-frequencies was obtained (Experiment 1) with no change in the size of the aftereffects when the adapter and test frequency were two octaves apart as compared to when they were the same. This then constitutes a replication of the findings of chapters 2 through 5, and thereby gives even further support to the conclusions drawn there, that recalibration is not frequency specific, and that its functional locus of is more central than the frequency specific sound localization mechanisms based on interaural time and level differences.

As for the second experiment, finding the same patterns of spatial generalization, even in morphology, as their pendants in chapter 6, and especially their asymmetry around the location of adaptation, irrespective of localization task further strengthens our confidence in their veracity. The results suggest that the relation between the auditory and visual maps of space in the brain are based on a point-to-point correspondence, which is consistent with animal data collected in owls (e.g., Hyde and Knudsen, 2002; see also §1.3.2).

In conclusion, a relative localization method as used here produces qualitatively the same results as when a pointing method was used, and therefore the patterns of generalization found thus far cannot be attributed to the pointing method as such. This seems a validation of the assumption made by Radeau and Bertelson (1974) that the bimodal monitoring task does not affect pointing behavior as such. In addition, the results suggest that the PPC is not the main site of recalibration since aftereffects and their generalization to other stimulus values also occur when a non-directed motor response is required.

Chapter 8

The Time Course of Visual Recalibration of Auditory Localization

8.1 Introduction

The final empirical chapter of this thesis is dedicated to the time course of visual recalibration of auditory localization. The two most basic questions concerning time course are how fast recalibration is acquired and how quickly it dissipates.

Acquisition functions of adaptation to intersensory spatial discordance are routinely reported for the case of visuo-proprioception (e.g., Bock & Burghoff, 1997; Fernández-Ruiz & Díaz, 1999; Fernández-Ruiz, Díaz, Hall-Haro, Vergara, Mischner, Nuñez, et al., 2003; Kitazawa & Yin, 2002; Martin, Keating, Goodkin, Bastian, & Thach, 1996; Roller, Cohen, Kimball, & Bloomberg, 2001; Yin & Kitazawa, 2001). They typically show (very) rapid adaptation to the discordance. Sometimes as little as six exposure trials are required for compensation for the discrepancy, although this number increases, apparently linearly, with the size of the discrepancy (Fernández-Ruiz & Díaz, 1999).

For adaptation to a ventriloquism situation only very little acquisition data is available. 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, 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 to 25 exposure blocks. Bertelson (1993) also found that visual recalibration of auditory localization is fast. After as little as 5 to 8 exposure episodes to an auditory-visual spatial discrepancy (each consisting in six single presentations of spatial conflict) 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.

All aftereffects dissipate, but they do not all dissipate at the same rate. For instance, aftereffects of a peripheral locus of adaptation, such as the color afterimage, tend to decay in a matter seconds whereas more complex aftereffects, such as the contingent color aftereffect can still be effective days after exposure (e.g., McCollough, 1965). Dissipation times then can be informative as to the locus of adaptation. Very fast dissipation betrays a peripheral or sensory locus whereas extremely long retention times must mean that central processes are involved.

Acquisition and dissipation functions, either on their own 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 nice example. They recently showed that exposure to incongruent auditory-visual speech (i.e., a McGurk type situation) 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 speech adaptation (Bertelson, Vroomen, & de Gelder, 2003, see also §1.2.1). This opposition was already an indication that they were dealing with two different perceptual processes. Two subsequent time course studies, one on acquisition and another on dissipation gave further evidence of this. The acquisition study (van Linden, Vroomen, de Gelder, & Bertelson, 2004) showed that, whereas recalibration quickly reached asymptote and after a while even decreased back to baseline, the speech adaptation effect continued to increase slowly as exposure continued. Similarly, the dissipation study showed differential patterns of decay (Vroomen, van Linden, Keetels, de Gelder, & Bertelson, 2004).

To the best of our knowledge there are no published data available speaking to the dissipation of the aftereffects of ventriloquism and we turn to other studies on various types of aftereffects as a frame of reference. Many studies on the dissipation of aftereffects as a function of exposure time (e.g., Hershenson, 1989, 1993; Stager & Burton, 1964; Taylor, 1963) can be summarized quite simply. The more exposure, the slower the dissipation rate. In other words, more exposure allows better retention of the aftereffect.

To explore the time course of recalibration, two experiments were run, one each for the acquisition and dissipation function. Although the experiments were run as pilots they are reported nonetheless because they gave some interesting insights.

8.2 Experiment 1: Acquisition functions

From earlier studies conducted in our laboratory we have established that one minute of exposure is sufficient to establish a reliable aftereffect. The present experiment is aimed at determining the acquisition function of recalibration by measuring the aftereffect at set intervals across approx. 72 sec of exposure to an auditory-visual spatial discrepancy. Participants completed two sessions with a total of eight runs, one for each combination of the direction of the visual distracter (left or right) and one of four levels of auditory-visual spatial discrepancy (5°, 10°, 15°, or 20°).

8.2.1 Method

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.

Apparatus and Stimuli

The 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 a acoustically transparent black cloth, each consisted of a loudspeaker (Visaton, FRWS 5, Ø = 5 cm) with an LED (Ø = 1 cm) over its center. Nine of them 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. 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 generated using CoolEdit 2000 (sample rate: 44.1 kHz, resolution: 16 bits) and converted from digital to analogue by means of a 16-bit PC controlled soundcard (Creative Sound Blaster 16). The stimulus itself was a 200 ms long 750 Hz pure tone, with 5 ms linear rise/fall envelopes, presented at 64 dB (A). The (synchronous) LED flashes also lasted 200 ms, and were clearly visible through the occluding cloth when lit.

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. 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 brief delay for saving the intermediate data and initiating the next run.

Each run was made up of three consecutive phases, a pretest, an auditory-visual spatial discrepancy, and an erasure phase. At this point it should be noted that the experiment is modeled after that by Bertelson (1993). This lead to a departure in procedure from the previous studies in this thesis. The auditory stimulus in the pre and posttest is, instead of the usual single short tone burst, a train of tones extending over a period of several seconds. The *pretest phase* then consisted of 18 randomized auditory localization trials, 6 from 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.

The spatial *discrepancy phase* was further divided into 12 adaptation blocks, each consisting in a number of exposure trials and a single localization trial. Exposure trials were 6 presentations, at 1 sec 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. To ensure that the participant attended the stimuli, there were occasional catch trials (four in total) across the twelve adaptation blocks, which consisted in the single omission of a visual distracter. This could occur in any of the adaptation blocks except for the first one. Within an adaptation block the catch trial could occur at either the second, third, fourth, or fifth presentation. It was the participant's task to detect these occurrences and to indicate this by pressing a button on the response box. A single localization trial followed the exposure trials after 1300 ms. In this manner, the three loudspeakers are 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 except that it was shorter and there was no spatial discrepancy between the auditory and the visual inputs. 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 sec intervals. Here there were also occasional catch trials (two in total) with the same task for the participant. The localization trial followed the exposure trials after 1300 ms and the participant was once again allowed 2500 ms to respond. In this manner, the three loudspeakers are tested twice, in a quasi-random order, across the 6 blocks in the erasure phase.

The participant was instructed to use the dominant hand for all pointing and catch trial responses.

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.

8.2.2 Results

The data of four participants were excluded from further analysis. Two because of sub normal performance on catch trials. Two more were excluded for not being able to reliably discriminate between the three test locations. The remaining participants' catch trials scores were high, ranging from 92% to 100%.

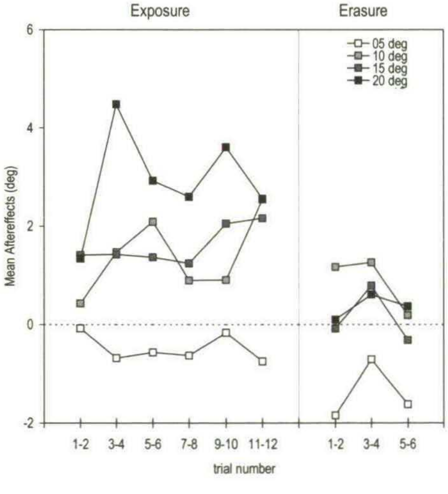


Figure 8.1 Aftereffect acquisition functions per auditory-visual spatial discrepancy.

Aftereffects were calculated by subtracting individual localization responses from the spatial discrepancy and erasure phases from the mean localization response on the corresponding speaker location in the pretest phase. They were counted positive when they went in the direction of the visual distracter (during the discrepancy phase). To reduce to noise in the curves, aftereffects were binned by averaging across two consecutive trials, and pooled across the direction of the visual distracter. The curves corresponding to the exposure and the erasure phase are shown in the left and the right panel of Figure 8.1, respectively.

Three observation are made. First, there are marked differences in the total level of adaptation between the four functions. The 5° auditory-visual discrepancy condition produces no aftereffect at all (mean -.10°, ns.). For the 20° discrepancy, substantial aftereffects are found (mean = 3.06°, $p < .01$), with the 10° (mean = 2.03°, $p < .005$) and 15° (mean = 1.37°, $p < .05$) discrepancy conditions producing intermediate aftereffects.

Second, the curves follow different time courses. The 5° curves stays close to the zero line. The 10° curve at first increases rapidly but shows a drop after 5-6 adaptation blocks, it then increases again, although this particular pattern most likely reflects measurement noise instead of perceptual development. The 15° curve shows a steady, almost monotonously increasing trend. For this and the 10° curve it is most probable that the curves have not reached asymptote yet at end of the exposure phase. The 20° curve on the other hand at first increases very rapidly to its maximum (~4.5°) after only 3-4 adaptation blocks (i.e., 24 single exposures to the auditory-visual discrepancy), it then quickly settles at around 3°.

Third, there still appear to be some aftereffects at the beginning of the erasure phase, and for the 5° curve there is a marked shift *away* from null in the opposite direction of the visual distracter.

8.2.3 Discussion

The main finding of the present experiment is that aftereffects build up rather quickly, although only the 20° function seems to reach asymptote before the completion of the exposure phase. The 10° and 15° functions, on the other hand, still seem to be on the increase, even though in absolute size they have reached the asymptotic level of the 20° function. They are apparently in contradiction with those of Bertelson (1993), since that study suggests that asymptote should have been reached before the end of the exposure. There is, however, no exact time information available from Bertelson's study, only the number of exposure trials. More importantly, each of these exposure trials, in fact, offered six times more exposure since each consisted in six rapid presentations of the discrepancy, whereas in the present experiment there was only a single such presentation. Thus, even if, in objective time, asymptote was reached sooner in Bertelson's study, it was most likely due to the larger *amount* of exposure received in that period than to exposure time per se.

It is unclear why the smallest amount of discrepancy did not produce any aftereffects. Aftereffects have been obtained with discrepancies in this order of magnitude (Bertelson, 1993; Recanzone, 1998). On the other hand, other investigators have failed to find aftereffects even with a 10° discrepancy (Kalil & Freedman, 1967). One interesting account of this is that there exists an inverse relationship between size of discrepancy and the amount of time required to reach asymptote. The present results are consistent with such an interpretation, for the largest discrepancy reached asymptote much sooner than the intermediate ones (i.e., 10° and 15°).

There is also no ready explanation for the negative shift in localization of the 5° curve in the erasure phase.

8.3 Experiment 2: Dissipation functions

Here we aimed at examining the shape and rate of the dissipation functions. Based on the literature on adaptation in general, it is expected that retention increases as a function of the amount of exposure. In other words, the more exposure the slower the dissipation rate. To test this directly an experiment was designed that follows the standard pretest-adaptation-posttest paradigm, where participants are adapted to an auditory-visual spatial discrepancy for one, three, or five minutes.

8.3.1 Method

Participants

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

Apparatus and Stimuli

The setup was the same as in experiment 1. The auditory stimulus was a 200 ms long 750 Hz pure tone, with 5 ms linear rise/fall envelopes, presented at 64 dB (A).

Procedure

We ran a completely within-subject design with two factors, Exposure Duration (1, 3, or 5 minutes) and Direction of the Visual Distracter (to the left or to the right of the sound). All six conditions were run twice, for a total of 12 runs. Runs were 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 posttest. 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 sec, 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 *posttest* was the same as the pretest except for the randomization of the trials. Posttest 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 posttest trial orders were created which were rotated across runs.

The *spatial discrepancy phase* consisted in 1, 3, or 5 minutes of exposure to the condition's particular auditory-visual discrepancy, at a presentation rate was 1 per-sec, so there were 60, 180, or 300 presentations, respectively. The spatial discrepancy, fixed at 15°, was presented across the array of three speakers in sets of five. To ensure that the participant attended to the exposure stimuli, there were occasional catch trials (2, 6, or 9, depending on the exposure duration), which consisted in the omission of one visual distracter. It was the participant's task to detect these occurrences and to indicate this by pressing a button on the response box.

8.3.2 Results

Overall, performance on the catch trials was high (> 80%) and none of the participants were excluded.

As before, aftereffects were calculated by subtracting the individual posttest 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. Aftereffects were binned per three consecutive posttest trials. Since each localization trial lasted 3.33 sec, one bin corresponds to a period of 10 seconds. The results are shown in figure 8.2.

The most obvious observation is that there is no evidence of any dissipation over time. All three functions are as good as level and show no sign of declining. At the same time, there is a clear effect of the duration of exposure. The longer the exposure the larger the overall aftereffect. To further illustrate this the means across the whole posttest have been plotted alongside the "dissipation" functions.

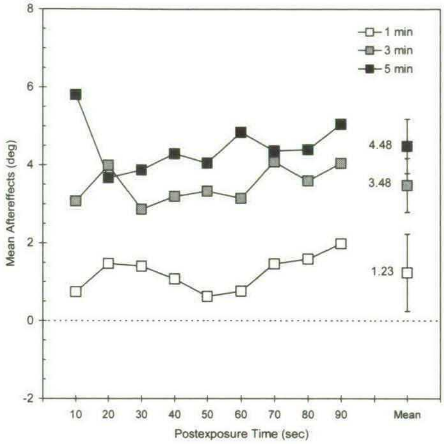


Figure 8.2 Aftereffect dissipation functions per duration of exposure to auditory-visual spatial discrepancy.

These observations were confirmed by statistical analysis. A 3 (Exposure Duration) x 2 (Direction of Visual Distracter) x 9 (Post Exposure time) repeated measures ANOVA revealed three significant effects, the overall effect, $F(1, 7) = 31.15, p < .001$, a main effect of Exposure Duration, $F(2, 14) = 5.32, p < .02$, and a main effect of Direction of Visual Distracter (showing generally larger aftereffects for adaptation to the right), $F(1, 7) = 15.89, p < .01$. All other effects, $p > .33$.

The main effect of Exposure Duration was further explored with paired t-tests, which revealed a marginally significant difference between 1 and 3 minutes of adaptation, $t(7) = 1.96, p < .1$, a significant effect 1 and 5 minutes of adaptation, $t(7) = 2.96, p < .05$, and no difference between 3 and 5 minutes of adaptation, $p > .23$. In addition, the individual correlations between exposure duration and mean aftereffect (i.e., averaged across the whole posttest) were calculated. These ranged from -.2 to 1, with a median of .8. Six out of eight participants showed correlations between .6 and 1, the two others were zero and -.2.

8.3.3 Discussion

The two main findings of the experiment were both somewhat surprising. First, there was no indication of any dissipation as time progressed. This could mean that retention after as little as 1 minute of exposure is already very strong. One could even claim that aftereffects are permanent until new information becomes available, such as, for instance, erasure trials as in experiment 1. Another possibility is that the time range of post exposure testing was simply to short to detect a decrease in aftereffects. In that case, dissipation apparently occurs *after* at least 90 seconds.

The second finding allows a fortuitous insight in the acquisition of recalibration. Aftereffects got larger with increasing exposure time, as shown both by the ANOVA and the individual correlation results. This supports the claim made after the first experiment that, for a 15° discrepancy, asymptote has not yet been reached after 1 minute of exposure (or better, 60 exposure trials). Since there was no reliable difference between the 3 and 5 minutes function, we assume that both were at asymptote and thus argue that asymptote is reached somewhere between 1 and 3 minutes, or between 60 and 180 exposure trials.

The main question of the experiment thus remains largely unanswered.

8.4 General Discussion

The two pilot experiments aimed at exploring the acquisition and dissipation functions of the aftereffects of ventriloquism, and the results can be summarized as follows. Consistent with previous reports (Bertelson, 1993; Radeau & Bertelson, 1976), acquisition was found to be relatively fast with (at least for the 20° function) asymptote reached after only a few exposure trials. The second experiment, on the other hand, seems to indicate that more than 60 exposure trials (i.e., one minute) are necessary in order to reach asymptote (for a 15° discrepancy), and probably more close to 180 (i.e., three minutes). As for dissipation, the aftereffect is retained for longer that 1.5 minutes even after only 1 minute of exposure.

Given their preliminary status the results do not allow any hard conclusions, but they do offer some interesting insights. First, the apparently long retention of aftereffects points at a locus of recalibration that is central (i.e., not peripheral). This is consistent with what was suggested in previous chapters.

Second, one surprising finding was that with a spatial discrepancy of 5° no aftereffects were obtained at the end of the adaptation phase, despite previous demonstrations that such a relatively small discrepancy can produce aftereffects (Bertelson, 1993; Recanzone, 1998). The explanation offered is that there is an inverse relationship between size of discrepancy and the amount of time required for adaptation. Such a relationship could make functional sense. Large errors are less likely to be due to noise in the system but more so because to internal errors and should be quickly corrected for. Small errors, on the other hand, could be the result of noise and more evidence of systematic error is needed before the perceptual system is compelled to correct for it.

It was mentioned that for the acquisition experiment there was a departure from regular procedure. Instead of a single short tone burst, the auditory stimulus that participants were to localize was a relatively long train of sounds. Could this have had an impact on the results? It could, for instance, be argued that the longer duration of the auditory stimulus allowed a more reliable perceptual estimate of its location. On this notion alone one expects less visual bias (Alais & Burr, 2004; Ernst & Banks, 2002). Likewise, a more reliable auditory could be less susceptible to visual recalibration as well. Although there have been no direct tests of this interesting possibility, the results from the study by Lewald (2002) suggest its plausibility. In this study, aftereffects obtained with the optimal type of auditory signal for sound localization (i.e., white noise) were considerably smaller (mean = 2.4°) than those obtained with suboptimal pure tones (mean = 4.5°). In addition, the longer stimulus duration also allows eye positioning on the auditory target, which could be used as an additional localization cue, and is thereby contributive to a more reliable perceptual estimate of the auditory target.

In any case, these experiments were just a first steps toward studying the time-course of adaptation to auditory-visual spatial conflict. Many questions remain, and obviously the truism holds that more experiments are needed to answer these. For instance, what is the amount of exposure needed to reach asymptote? How should this amount be quantified, in terms of time or, much more likely, in terms of the number of exposure events? Is there, as was suggested, an inverse relation between the size of discrepancy and the amount of exposure needed to reach asymptote? How long is the retention of the aftereffects? Is there the presumed inverse relation between amount of exposure and dissipation rate? What is the shape of the dissipation function?

Chapter 9

General Discussion

9.1 Introduction

The research presented in this thesis is concerned with the aftereffects of ventriloquism. After an observer has been exposed to an auditory-visual spatial discrepancy, there is a shift in sound localization in the direction of the visual input, and in some studies the reverse has also been found. The importance of studying these effects lies in the fact that they reduce the registered spatial discrepancy between the two senses. This is taken as a manifestation of a *recalibration* of stimulus-to-percept matches, which may serve to maintain intersensory coherence. However, there has been surprisingly little systematic work on the aftereffects of ventriloquism (see §1.3.4). The present work has taken a closer look at two of the phenomenon's more fundamental properties, its extension and time course. The following sections recapitulate the main findings and discus the functional and possible neurophysiological locus of the recalibration, and their relation to other studies.

9.2 The extension of visual recalibration of auditory localization

The work in the first part of the thesis was based on a particular research strategy which in essence is a stimulus generalization paradigm as we know it, for instance, from work on classical condition (e.g., Hovland, 1937). Two types of generalization have been investigated, that across *sound-frequencies* and across *space.*

9.2.1 Generalization across sound-frequencies

The present work has consistently found substantial generalization across sound-frequencies. Irrespective of the adapter frequency, significant aftereffects were obtained for all test frequencies used. Small generalization decrements, in the order of 0.2°/octave, become apparent only when the distance between the adapter and test frequency was as large as four octaves (Chapter 3). Overall, aftereffects for the test frequencies at a four-octave distance to the adapter were in the order of 70% of those at a zero-octave distance (i.e., adapter and test frequency were the same; Chapters 3 and 4). Two further studies show that generalization is not dependent on the modality attention is allocated to during exposure (Chapter 2, see §9.4), on the octave intervals between adapter and test frequencies used (Chapter 5), or on the particular onset times of the auditory stimuli (Chapter 4).

There are three possible patterns of generalization. The first is complete transfer, aftereffects are found for all other stimulus values. The second is partial transfer, aftereffects are maximal around the adapting frequency and go down with increasing distance from that frequency. The third is no transfer whatsoever. Overall, the second pattern fits the present data best. In all experiments the largest aftereffects were found for the adapter frequency and those at the other frequencies were always smaller. This was even the case in the experiments of Chapter 2, even though the difference here was not statistically reliable.

The main conclusion drawn from this set of experiments is that the locus of visual recalibration of auditory localization is beyond the frequency specific, sound localization mechanisms based on interaural differences. That is, interaural time (ITD) and level differences (ILD). Instead a more central site is suspected. It also calls into question the involvement of the frequency-tuned neurons of the primary auditory cortex (Lewald, 2002; Recanzone, 1998), or any other part of the auditory localization pathway that is strictly frequency tuned. The specific issue of the locus of recalibration is treated in more detail in §9.4.

9.2.1.1 Outstanding issues

We are still faced with the discrepancy between our results and those of Recanzone (1998) and Lewald (2002), who found no generalization whatsoever across a two-octave difference between adapter and test frequency. In this respect, there remains at least one concern and a number of differences between these studies and the present work that need to be addressed. They are all related to the particular acoustics in the experiments. We deal with them in turn in order to be complete, but it should be noted that there is no clear idea, except potentially the first one, as to how these could actually *explain* the contrasting results.

To begin, there might have been differences in the particular quality of speakers. It might be that the speakers produced energy at frequencies other than the intended ones. Furthermore, there were differences in the testing environment, the distance of the sound sources to the observer, and the intensity level of the auditory stimuli.

Speaker quality

One potential problem, and in particular for the frequency generalization studies, is that the loudspeakers may have produced tones at frequencies other than the intended tonal frequency. To check for this, post hoc spectral analyses were done on the auditory stimuli to see whether the speakers (and the experimental set up in total) produced any energy at frequencies other than intended. A microphone (Sennheiser, MD421N) fed the signal, recorded from where the center of the participant's head would be, into a preamp (K+H, sv 4) and then onto DAT (Sony, 55ES). The DAT recording was transferred to hard disk and fast Fourier transforms were computed using Cool Edit 2000 (Syntrillium). Recordings were made of all stimuli used in the frequency generalization experiments, from all relevant speakers (the number of loudspeakers used varied across experiments between three and five).

The frequency response characteristics of the speakers was determined. The signal was a 1000 ms of white noise. All speakers had a characteristic that was flat within 12 dB between .3 and 15 kHz, with 5-6 dB/octave roll-off. This is within the performance range used in other laboratories researching sound localization (e.g., Zwiers, van Opstal, Cruysberg, 2001).

For the tonal stimuli, the finding was that there was no substantial extraneous energy at any frequency other than that of the adapter stimuli. There were some small (but largely inconsequential) anomalies. For instance, for the 750 Hz tone (Chapter 2), only one speaker (at 9° to the left of straight ahead) produced a small peak (still more than 36 dB below the 750 Hz peak) at the first (non-tested) octave (i.e., 1500 Hz), but nothing at the important second octave (i.e., 3000 Hz). The four other speakers showed no such "impurities". The 3000 Hz, on the other hand, did produce an additional, but again very small, peak (for all five speakers at least 40 dB below the 3000 Hz peak). It was however, not at 750 Hz, but at the nearby 800 Hz. The 400 Hz adapter produced no anomalies whatsoever in any of the speakers, nor did the 6400 Hz adapter. The intermediate frequencies (i.e., test frequencies only) of the experiments in chapters 3 and 5, on occasion, produced very small peaks, and at the first octave only, but critically never at one of the adapter frequencies. The stimuli with the different rise/fall times in chapter 4 also did not produce any extraneous energy.

In summary, part of the generalization from a 3000 Hz adapter to the 750 Hz test tone in chapter 2 could, theoretically, be attributed to a small impurity in the loudspeakers. But, at the same time, it seems very unlikely that so little energy at a frequency *nearby* the test frequency could be responsible for the nearly total generalization. In addition, all other tonal stimuli of the remaining experiments did not cause any such impurities. An explanation solely based on impurities in speaker characteristics then does not apply to the experiments.

Reverberations versus anechoic

Lewald (2002) and Recanzone (1998) used (nearly) anechoic environments whereas the present work was carried out in a semi-reverberant environment. Reverberation times were 510 ms at 125 Hz, 400 ms at 250 Hz, 330 ms at 500 Hz, 270 ms at 1000 Hz, and 280 ms at 2000 Hz.

In psychoacoustics reverberations are generally thought to have little effect on directional hearing (Shinn-Cunningham, 2001) because of phenomena such as the precedence effect (e.g. Litovsky, Colburn, Yost, & Guzman, 1999). There are several studies in the (psycho)acoustical literature that used semi-reverberant rooms to test the utilization of interaural difference cues (e.g., Abel, Figueiredo, Consoli, Birt, & Papsin, 2002; Abel, Giguere, Consoli, & Papsin, 2000; Abel & Paik, 2004). All these were conducted in a semi-reverberant test booth (with, incidentally, longer reverberation times than ours, between 600 and 300 ms for frequencies between 125 and 8000 Hz). Auditory stimuli were 500 and 4000 Hz, one-third octave band signals to assess the utilization of interaural time and level differences, respectively.

Only recently studies have been conducted to explore the actual effects of echoes and reverberation on localization cues and performance (see Shinn-Cunningham, 2003, for a brief review). On an acoustical level, reverberations distort all aspects of the signal reaching the ear. ITDs and ILDs are distorted with reverberation due to a de-correlation of the signals at the ears, especially with increasing distance and laterality of the source. Simply put, reverberation adds noise to the ITDs and tends to reduce ILDs. However, although both types of interaural difference cues are affected by reverberation, they do not change in a qualitative manner. That is, ITDs are still ITDs but more noisy, and ILDs are still ILDs but somewhat reduced. The effects of reverberation on localization *performance* is relatively modest, typically an increase in the variability of azimuthal localization judgments, while distance perception actually improves (Shinn-Cunningham, 2003). Functionally then there seems little difference between testing in an anechoic room and one with moderate amounts of reverberation, such as ours.

Thus, making inferences, like we have done in chapters 2 through 5, as to the involvement of the peripheral interaural localization mechanisms still seems warranted even when testing was conducted in a moderately reverberant room. It would of course be prudent, although at present impossible for us due to the lack of the necessary equipment, to determine experimentally to what extent room reverberations might have actually contributed to the generalization across sound-frequencies.

Near versus far field localization

Sounds in the Lewald study and the Recanzone study were presented in the so-called far-field (i.e., the sounds were presented further than 1 meter from the observer) whereas ours were in the near-field (< 1 m; e.g., Brungart, 1999). By far most of the work on sound localization has been in the far-field and only few studies have looked at near-field localization. A notable exception is the work by Brungart and colleagues (e.g., Brungart, 1999; Brungart, Durlach, & Rabinowitz, 1999; Brungart & Rabinowitz, 1996). As for the physics of near-field localization, it turns out that only ILDs are markedly affected by bringing the sound source closer to the head, whereas ITD and spectral cues are more or less unchanged (Brungart & Rabinowitz, 1996). As was to case for the effects of reverberation, localization performance itself is only marginally affected and again auditory distance perception improves (Brungart, Durlach, & Rabinowitz, 1999; Brungart & Rabinowitz, 1999). As far as auditory spatial *perception* is concerned there are no marked differences between near and far-field.

Stimulus intensity

It is known that the frequency-specific response of neurons in the auditory primary cortex are not fixed but are sensitive to the stimulus intensity (diRibaupierre, 1997). Could it be that the intensity level of the auditory stimulus played a critical role, with generalization only occurring for somewhat louder stimuli because of, for instance, a change in neuronal tuning of the substrates involved? The presentation level of auditory stimuli in Recanzone's study was 45 dB(A), whereas in the present work presentation was between 64 and 70 dB(A). Interestingly, a recent study on the generalization of aftereffects across sound-frequencies, with macaque monkeys, found significant generalization to the none adapter frequency in the order of 50% with a level of 65 dB(A), although it was not recognized as such by the authors (Woods and Recanzone, 2004). Unfortunately, this trend no longer holds when considering Lewald's stimuli (producing no generalization), which were at 60 dB(A), taking them close to the range of ours and those of Woods and Recanzone.

9.2.2 Generalization across space

Probably the most surprising set of results were obtained in the study of the spatial generalization. During exposure an auditory-visual spatial discrepancy was presented in a single spatial location and aftereffects were measured at that location and more distant ones (Chapter 6). The results clearly showed generalization gradients, with aftereffects generally being the largest at the location of exposure and decreasing with distance from that location. This then suggest that the visual signal affected only a limited region of auditory space.

The most surprising result was that the generalization gradients were strongly asymmetrical. Aftereffects were only obtained on *one* side of the exposure location, namely that side occupied by the visual distracter during exposure. This can be seen most clearly in Figure 6.1 for when the visual distracter was to the left (open squares). There were only aftereffects in the left part of that function but none on the opposite. This asymmetry was apparent in all other conditions across the whole study.

We have argued that effects of eye position could not explain the asymmetrical functions, although *contributions* of such a factor can of course not be excluded. Participants were instructed to look at the visual distracter during the exposure phase. It is known that eye position has an effect on spatial hearing, and the most frequent finding is that the apparent location of a sound shifts in a direction *away* from that of the eye position (e.g., Bohlander, 1984; Lewald, 1997, 1998). This is clearly opposite the direction necessary to account for visual bias and also for the dependence of its aftereffects on the position of the visual distracter during exposure. Also, if the immediate and aftereffects of ventriloquist effect were simply the consequence of fixation on the visual distracter, both would occur irrespective of the relative timing of the auditory and visual stimulus (i.e., whether they are synchronous or asynchronous). Desynchronization, in fact, eliminates (Bertelson & Aschersleben, 1998) or at least strongly reduces the immediate effects (Bertelson, Vroomen & de Gelder, 1997; Choe, Welch, Gilford & Juola, 1975; Radeau & Bertelson, 1987; Thomas, 1941; Warren, Welch & McCarthy, 1981) as well as aftereffects (Radeau & Bertelson, 1977). Finally, Radeau and Bertelson (1977, Experiment 2) found no aftereffect when participants monitored an LED, located 20° to one side of the sound location. Nevertheless, it would be prudent to conduct further studies on this with more strict control of eye position.

9.2.2.1 Relation to other studies

There is no precedence in the ventriloquism aftereffect literature for spatial generalization. Although it was not the aim of the studies, some of the earlier work (e.g., Radeau & Bertelson, 1977, 1978) could have provided relevant data, since exposure was also restricted to a single location in space, but unfortunately the reported aftereffects were always averaged across the several test locations and direction of visual discordance.

Another potential source was a study by Zwiers, van Opstal, and Paige (2003). Recall from §1.3.4 that these investigators used lenses to alter the visual array. They did so by compressing it by a factor 0.5, and only within a radius of 20°, the rest of the peripheral visual field was masked. As expected, adaptive changes in sound localization were found for those locations directly influenced by visual information. The question of interest here is of course what happened in the part of space that was masked. It turned out that the wearing of the lenses clearly affected localization in this region. In fact, the largest shifts were found at the edge of the adapted field and then continue *unaltered* into the periphery.

There are, however, several major methodological differences that make comparison of these results to ours cumbersome. For instance, the time frame of the studies are very different. Participants wore the lenses for long periods for time (for two to three days, continuously during the day), whereas exposure in our case was very brief (seven times one minute). Since very little is known about the time course of recalibration it is very hard to judge how the achieved end points of adaptation in both cases compare. Also, the lenses affected a relatively large and *continuous* part of visual space. Conflict in our study, on the other hand, was in principle punctiform. Finally, the lens-wearing participants "were encouraged to proceed with active natural behavior" (p. 180). This necessarily also includes instances of auditory-tactile stimulation. It is known that that tactile information is capable of biasing sound localization (Caclin, Soto-Faraco, Kingstone, & Spence, 2002) and that exposure to auditory-tactile spatial conflict can be adapted to (i.e., it is inductive to recalibration; Freedman & Wilson, 1967). Thus, it becomes very hard to separate the tactile influences from the visual ones.

9.2.2.2 A possible model for spatial generalization

In this section a simple model is proposed to account for both the spatial generalization gradients and their asymmetry around the locus of adaptation. The model assumes that recalibration occurs at a site that is topographically organized and has a point-to-point correspondence between its auditory and visual parts. Auditory space is represented by an array of place specific units. These units represent a restricted part of space (i.e., have limited spatial receptive field) and their spatial tuning can be recalibrated by, among other factors, visual information from the corresponding visual map of space.

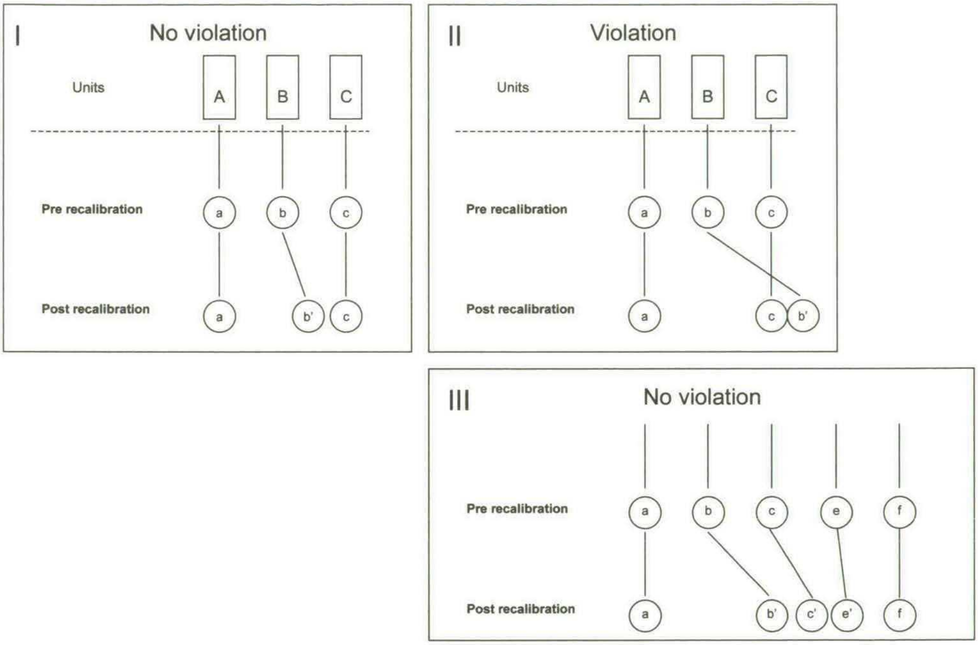


Figure 9.1 The relative location constraint on visual recalibration of auditory localization. Units symbolize neural structures that represent a particular part of space, and the circles with lower case letters symbolize those corresponding locations. Symbols with lower case letters and apostrophes represent the newly learned (through recalibration) locations. Panel I shows a situation where the constraint is not violated. Panels II and III show a violation and its resolution, respectively.

The conditions for recalibration is that units preserve their spatial tuning *relative* to those of their neighboring ones. Consider figure 9.1 for a graphical depiction of this principle. The units (upper case letters) in all panels represent a subset of neural entities that each represent a certain but adjacent part of auditory space (lower case letters in the figure). The lower case letters with an apostrophe represent spatial locations *after* recalibration. Panel I shows a situation where unit B is recalibrated (to the right) to such an extent that *b'* is still to the left of *c*. This clearly is not a violation of the constraint since all relative positions are preserved. Aftereffects are expected only for unit B. Now consider the situation in panel II. In this case unit B is recalibrated to such an extent that *b'* would end up beyond the location of *c* (notice that the lines connecting the pre and post conditions cross). In other words, *b'* is now to the right of *c* instead of to the left, its proper location. This then constitutes a violation of the constraint, which needs to be resolved. The solution (panel III) is to also recalibrate unit C such that *c* is again to the right of *b'* (i.e., *c'*). The gradients found in chapters 6 and 7 suggest that the need for more distant units to recalibrate in order to conform to the relative location rule decreases with distance. In the figure (panel III) this can be seen in the decreasing change in slope of the lines connecting the pre and post conditions.

From this model a number of testable predictions can be made. For instance, local recalibration affects only those units that are in the direct "path" prescribed by the visual distracter. Thus, sound localization in locations that are not in that direction should remain unaffected (e.g., unit A in figure 9.1). This is what was found in the experiments of chapters 6 and 7. Additional experiments could explore locations that are orthogonal to the direction but also those that are in the same direction but on a different elevation. That is, recalibration is for instance, in the straight ahead location with the sound at 0° azimuth and 0° elevation, and the visual distracter at 15° to the right of that location. Testing is done along the azimuth on both 0° and, for instance, 15° elevation. The model predicts there are no, or at least diminished, aftereffects on the higher elevation tests. Another prediction can be derived from what is depicted in panel I of figure 9.1. As long as there is no violation of the relative location constraint there should be no generalization to other locations.

9.3 The time course of visual recalibration of auditory localization

Little is known about the time course of recalibration. Such knowledge, however, is crucial for a complete understanding of the mechanisms involved. It was already suggested in the previous section and by others that different amounts of exposure can lead to (quantitatively) different end points of adaptation. Lewald (2002), for instance, expresses a concern in relating animal studies, typically using very long exposure periods of up to several weeks, to human studies, using relatively brief periods of exposure, as short as one minute. He also distinguishes between three levels of adaptation, short term (obtained in the order of less than one second), intermediate (minutes), and long term (weeks). The intermediate stage (typical for the human studies) could function as a preliminary stage for longer lasting adaptive processes important to development (Held, 1965) and subsequent maintenance of intersensory coordination. There is no a priori reason to assume that these stages necessarily have the same properties.

In any case, here we have made a first start at studying the time course of visual recalibration of auditory localization. For now, only two studies have been conducted. One tracked the acquisition and another the dissipation functions of recalibration. In general, acquisition was fast, although there was considerable variability between the different levels of auditory-visual spatial discrepancy (see Figure 8.1). For instance, in one condition (20° discrepancy), acquisition was very fast, peaking after only 3-4 exposure trials (each consisting in 6 single exposures to spatial discrepancy) and reaching asymptote after 5-6 trials. When the discrepancy was 5° no aftereffects were apparent at all, not even after 12 exposure trials. The two intermediate conditions (10° and 15°) quickly showed aftereffects although they apparently still had not reached asymptote after 12 exposure trials. The results are in general agreement with an earlier study by Bertelson (1993), although that study did find aftereffects for the 5° discrepancy condition.

It remains unclear why the smallest amount of discrepancy used did not produce aftereffects. It is probably not due to the size of the discrepancy per se because spatial conflict in this order of magnitude has been shown to produce aftereffects (e.g., Woods & Recanzone, 2004), although other investigators have failed to find aftereffects with much larger ones (i.e., 10°; Kalil & Freedman, 1967). The functionally sensible hypothesis was put forward that there is an inverse relationship between magnitude of auditory-visual spatial discrepancy and the amount of time required to adapt to it. Large intersensory errors are unlikely to be caused by noise in the perceptual system but more so by more systematic factors such as drift. Consequently the errors should be quickly corrected for. Small errors, on the other hand, are much more likely to be due to internal or stimulus noise. More evidence of systematic error is needed before the perceptual system is compelled to correct the error. The present results are consistent with such an interpretation.

The dissipation functions did not reveal much as to the actual dissipation (see Figure 8.2), mostly because the post-exposure test interval was too short. It did, however, allow another insight into acquisition. Overall aftereffects increased as a function of exposure duration, with aftereffects for 1 minute of exposure significantly smaller than for 3 and 5 minutes. This in contrast to the earlier observation that asymptote is reached fast. Apparently it had not been reached (for a 15° discrepancy) after 1 minute.

It should be kept in mind that both studies were pilots and any conclusions are therefore tentative. These conclusions are that recalibration is relatively fast and is retained for at least 1.5 minutes, but very probably longer, after exposure stopped. Clearly some parameter settings were not optimal. For instance, in the dissipation study the post-exposure test period was far too short and a much longer one is needed. Also, ideally aftereffects should have reached asymptote before starting post-exposure testing to prevent having to correct for starting level.

9.4 Role of attention

Recall from the introductory chapter the model of Canon (1970, 1971; for a similar case but with visuo-proprioceptive conflict see Kelso, Cook, Olson, & Epstein, 1975). It states, in brief, that the two main determining factors for adaptation are, an intermodal inconsistency of inputs, which is similar to the Wallach's (1968) "informational discrepancy", and the direction of attention. The model makes the strong claim that adaptation occurs *only* in the modality that was *not* attended to during exposure. For instance, when attention is on the auditory stimulus during exposure the model predicts there to be no shift in auditory localization.

The results from chapter 2 are clearly at odds with this. A monitoring task during exposure was used to explicitly focus attention on either the visual (Experiment 2) or auditory (Experiment 3) modality. According to the model the latter condition should not produce any or at least significantly reduced aftereffects. The mean overall aftereffects were 1.89° and 1.84° for the visual and auditory attention condition, respectively. Not surprisingly, these were not statistically different from each other.

Although Canon (1970, 1971) did not distinguish between them, it may be that the type of attention makes a difference. In Canon's studies participants engaged in a spatial tracking task during exposure, whereas ours were to simply monitor for occasional deviants in a train of stimuli. It could be argued that Canon's task required spatial attention, and that ours is not spatial in itself and therefore does not require spatial attention. This offers an interesting idea for further study. Does the direction of a certain type of attention during exposure (e.g., spatial, focused, exogenous, or endogenous) make a difference? In any case, if Canon's model is to remain viable, it needs, at the very least, to be qualified to account for our results. Finding specific effects of attention would also be of interest in comparing aftereffects to the immediate effects of ventriloquism since these have already been shown to be independent of at least endogenous (Bertelson et al., 2000b) and exogenous (Vroomen et al., 2001b) visual attention.

9.5 The locus of recalibration

What about the functional and physiological locus of recalibration? The present work allows some conjectures.

9.5.1 Functional locus

Based mainly on the sound-frequency generalization work we can state that the functional locus of recalibration is most likely in a site more central to the main, sound-frequency specific, peripheral auditory localization mechanisms. Instead, it is proposed that the visual influence operates on auditory spatial perception only *after* all relevant auditory cues have been integrated (i.e., after auditory *intra*sensory integration is complete). A central locus is also evident from the work on the dissipation of aftereffects. After as little as one minute of adaptation, aftereffects were retained for up to at least 90 seconds. Observations in our lab and those of others (e.g., Recanzone, 1998) strongly suggest that after 20 minutes of adaptation aftereffects linger for as long as 10 to 20 minutes. Such long retention times cannot be attributed to simple fatigue in peripheral perceptual mechanisms alone.

A central locus, of course, makes good functional sense. The strength of cross-modal integration lies in the reduction of the consequences of modality-specific variability (de Gelder & Bertelson, 2003; Bertelson & de Gelder, 2004). Such variability reduction also occurs on a modality specific level through the combination of cues that provide (redundant) information about an object's particular feature (e.g., its location). The integration of cues from low-level auditory mechanisms like ITDs and ILDs into a unified auditory directional percept is an example of such intra-modal integration. This organization of intramodal and intermodal integration processes is clearly more economical than one where visual information needs to be combined with all separate lower level localization mechanisms.

9.5.2 Physiological locus

Can we relate the present findings to any of the neural structures known to be involved (see §1.4) in the processing of multisensory information? Given the general conclusions about the functional locus of recalibration we can speculate that that locus should (a) have access to both visual and auditory spatial information, (b) have a topographical representation of auditory space (to account for the spatial generalization functions), and (c) be not sound frequency specific (to account for the generalization across sound-frequencies).

The most obvious candidate seems to be the *superior colliculus* (SC), which meets all these criteria. It is also in agreement with the animal literature where it enjoys center stage as the preeminent site of recalibration. One problem with the SC model, however, is that functionally it has been associated with eye/head orienting behavior (e.g., Stein & Meredith, 1993), and both overt and covert attentional processes (e.g., Ignashchenkova, Dicke, Haarmeier, Their, 2004; Muller, Philiastides, & Newsome, 2004). It is unclear whether changes in the spatial representations in the SC would also affect more controlled localization behavior such as the hand pointing typically used to measure the effects of ventriloquism. In other words, the aftereffect can also be obtained when there is no orienting motor response (Chapter 7; Recanzone, 1998).

Perhaps cortical structures are more likely candidates. It has been argued based on two previous studies (Lewald, 2002; Recanzone, 1998) that the unimodal neurons in the *primary auditory cortex* (AI) have a role to play in the aftereffects of ventriloquism. This is based on a number of observations. First, AI has been shown to be important to sound localization (e.g., Heffner & Heffner, 1990; Jenkins & Merzenich, 1984; Kavanagh & Kelly, 1987). Second, its neurons have very sharp frequency tuning (e.g., diRibaupierre, 1997; Kaas, Hackett, & Tramo, 1999). Third, the aftereffects of ventriloquism did not generalize across sound-frequencies in those studies. Taken together then, the suggestion is indeed strong that the AI is involved. However, in the present work we have not been able to replicate the finding of frequency specificity, and in fact find the opposite result. This at least seems to calls into question the involvement of the AI.

The posterior parietal cortex (PPC), and in particular, its lateral intraparietal area (LIP), is another likely site of the plastic alignment of auditory and visual space (Recanzone, 1998). It plays a role in auditory spatial perception (e.g., Lewald, Foltys, & Töpper, 2002; Rauschecker & Tian, 2000) and contains a population of spatially tuned neurons that are responsive to both auditory and visual stimuli (Mazzoni, Bracewell, Barash, & Anderson, 1996).

Samenvatting

Inleiding

Dit proefschrift gaat over het zogenaamde buikspreker effect, of meer in het bijzonder de na-effecten ervan. Na een periode van blootstelling aan een *audio-visuele spatiele discrepantie* (d.w.z. een geluids- en lichtbron van verschillende locaties), vindt men typisch een verschuiving in geluidslokalisatie in de richting van waar voorheen het licht kwam (het tegenovergestelde is overigens ook gevonden). Het belang in het bestuderen van deze effecten ligt in het feit dat zij een weerspiegeling zijn van pogingen van het brein om de geregistreerde discrepantie tussen de twee sensorische systemen te verminderen. Met andere woorden, de na-effecten worden gezien als een manifestatie van een recalibratie van stimulus-tot-percept relaties, welke dienen tot het instandhouden van intersensorische coherentie.

Er blijkt echter verassend weinig systematisch onderzoek te zijn gedaan naar deze na-effecten (zie §1.3.4) met als gevolg dat we nog relatief weinig weten. In dit proefschrift zijn een tweetal fundamentele eigenschappen van het buikspreker na-effect nader onderzocht. Deze zijn de *extensie* en het *tijdsverloop* van het effect. In wat volgt wordt een korte samenvatting gegeven van de resultaten.

De extensie van het buikspreker na-effect

Met de extensie van recalibratie bedoelen we hoe ver de effecten als het waren rijken. Zulke kennis geeft belangrijke informatie over de aard van de verantwoordelijke perceptuele (en wellicht fysiologische) mechanismen. De extensie van een na-effect wordt onderzocht met behulp van een stimulusgeneralisatie paradigma welke met name bekend is uit de klassieke conditionering literatuur (bijv., Hovland, 1937). Stel een dier heeft geleerd om een specifieke reactie te geven bij een bepaalde stimulus (bijv., een 1000 Hz toon). Wanneer het dier dezelfde reactie geeft bij een (lichtelijk) *andere* stimulus (1200 Hz) dan is er sprake van stimulus generalisatie. Er zijn drie mogelijke patronen van generalisatie. *Complete generalisatie*, het dier reageert op iedere toon ongeacht de frequentie ervan. *Complete specificeit*, het dier reageert alleen maar op de geleerde toon (in dit voorbeeld dus de 1000 Hz toon). *Partiele generalisatie*, het dier reageert het meeste op de geleerde toon maar ook op andere tonen maar met afnemende intensiteit naarmate de toon minder lijkt op de geleerde toon. In het huidige onderzoek is gekeken naar twee vormen van generalisatie, die over *geluidsfrequentie* en die over de *ruimte*.

Generalisatie over geluidsfrequentie

Een van de belangrijkste redenen voor het bestuderen van generalisatie over geluidsfrequenties is het gegeven dat het geluidslokalisatie vermogen van het brein tot op zekere hoogte frequentie specifiek is. Met name voor het lokaliseren van geluidsbronnen in het horizontale vlak is geluidslokalisatie afhankelijk van het hebben van twee oren en de tijd en intensiteit verschillen die tussen beide ontstaan als functie van de richting van de geluidsbron. Deze twee vormen van interaurale verschillen worden verwerkt door verschillende (perifere) perceptuele mechanismen welke sterk frequentie specifiek zijn, waarbij tijdsverschillen effectief zijn voor relatief lage (lager dan 1500 Hz) en intensiteitverschillen voor relatief hoge (hoger dan 1500 Hz) frequenties. Het vinden van na-effecten voor tonen in het ene bereik na recalibratie met een toon in het andere zou ons vertellen dat de visuele invloed niet op het niveau van deze twee perifere perceptuele mechanisme werkt maar op een later moment in de perceptuele verwerkingsketen. Alle studies naar generalisatie over geluidsfrequentie in deze thesis wijzen erop dat dit inderdaad het geval is. In alle gevallen werd er substantiële generalisatie gevonden tussen de twee frequentie bereiken.

Generalisatie over ruimte

De meest verassende resultaten werden behaald uit de studie naar de spatiele generalisatie van recalibratie. Blootstelling aan de audio-visuele spatiele discrepantie was maar op een enkele locatie in de ruimte en na-effecten werden bepaald voor deze en meer verder gelegen locaties. Er waren twee opmerkelijke generalisatie karakteristieken. Ten eerste, de na-effecten waren het grootst in de locatie waar de oorspronkelijke blootstelling was maar namen af naarmate men verder van deze locatie af ging. Dit suggereert sterk dat recalibratie slechts een beperkte reikwijdte heeft. Ten tweede, en meest opmerkelijk, was de vondst van een asymmetrisch patroon van generalisatie. Na-effecten werden alleen maar gevonden aan die zijde van de originele locatie waar de visuele bron zich bevond, terwijl aan de andere zijde er geen na-effecten kenbaar waren. De oorzaak voor dit patroon is nog onduidelijk maar we menen dat het in ieder geval niet terug te voeren is naar de reeds bekende invloed van oogbewegingen op geluidslokalisatie (bijv Lewald, 1997, 1998). In hoodfstuk 9 wordt een mogelijk model van spatiele generalisatie gegeven.

Het tijdsverloop van recalibratie

Een geheel ander aspect van recalibratie dat in deze thesis werd onderzocht is het tijdsverloop van recalibratie. Met andere woorden, hoe snel bouwt het zich op en hoe snel verdwijnt het weer? Zulke informatie is kritisch in het volledig beschrijven en begrijpen van de betrokken perceptuele (en net zoals voor de extensie studies, wellicht de fysiologische) mechanismen.

Het moet voorop gesteld worden dat de studies die gerapporteerd worden in deze thesis slechts een begin zijn en nog lang geen volledig beeld geven. Hiervoor zijn nog meer studies nodig. Het beeld dat echter nu geschetst wordt door de resultaten is makkelijk als volgt samen te vatten. *Recalibratie is snel.* Ne-effecten zijn al meetbaar naar een klein aantal blootstellingen aan een audio-visuele spatiele discrepantie. Bovendien blijven de na-effecten relatief lang aanwezig.

De locus van recalibratie

Waar dan in het brein (in functionele termen) vindt recalibratie plaats? De huidige resultaten veroorloven enkele voorzichtige conclusies. Met name de resultaten van de generalisatie studies zijn hier bepalend. Het lijkt erop dat (1) de gevolgen van recalibratie sterk generaliseren over geluidsfrequentie maar in beperkte mate over de ruimte, en (2) dat recalibratie snel is en dat het relatief lang blijft.

De geluidsfrequentie generalisatie studies laten ons toe te concluderen dat recalibratie geschiedt ergens *nadat* de perifere geluidslokalisatie mechanismen hun werk hebben gedaan en dus dat het meer centraal is. Dit gegeven past binnen een meer algemeen beeld van multisensorische integratie waar gesteld wordt dat eerst de intrasensorische integratie mechanismen hun werk doen en dat crossmodale processen werken op de producten van deze integratie. Zulk een centrale locus lijkt tevens zinnig vanuit een "economisch" standpunt. De spatiele generalisatie studies suggereren dat de locus van recalibratie een wellicht "punt-tot-punt" organisatie die de representaties van de auditieve en visuele ruimte met elkaar verbinden. Verder lijkt de ruimtelijke tuning van de "punten" binnen representatie van de auditieve ruimte tot zekere hoogte van elkaar afhankelijk maar dat deze afhankelijkheid snel afneemt met de afstand tussen de punten.

De resultaten van de tijdsverloop studies zijn ook consistent met een centrale locus in zoverre dat de relatief lange retentie van de recalibratie (soms tot meer dan 20 minuten) niet kan worden toegeschreven tot, bijvoorbeeld, simpele vermoeidheid van perifere mechanismen.

References

Abel, S. M., Figueiredo, J. C., Consoli, A., Birt, C. M., & Papsin, B. C. (2002). The effect of blindness on horizontal plane source identification. *International Journal of Audiology, 41,* 285-292.

Abel, S. M., Giguère, C., Consoli, A., & Papsin, B. C. (2000). The effect of aging on horizontal plane sound localization. *Journal of the Acoustical Society of America, 108,* 743-752.

Abel, S. M., & Paik, J. E. S. (2004). The benefit of practice for sound localization without sight. *Applied Acoustics, 65,* 229-241.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14,* 257-262.

Andersen, R. A., Bracewell, R. M,. Barash, S., Gnadt, J. W., & Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a in macaque. *Journal of Neuroscience, 10,* 1176-1196.

Aschersleben, G. & Bertelson, P. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension 2. Evidence from sensory motor synchronization. *International Journal of Psychophysiology, 50,* 157-163.

Baily, J. S. (1972). Adaptation to prisms: Do proprioceptive changes mediate adapted behavior with ballistic arm movements? *The Quarterly Journal of Experimental Psychology, 24,* 8-20.

Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science, 285,* 257-260.

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A-Optics Image Science and Vision, 20,* 1391-1397.

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

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

Bertelson, P. (1993). *The time-course of adaptation to auditory-visual spatial discrepancy.* Paper presented at the 6th Conference of the European Society for Cognitive Psychology, Copenhagen.

Bertelson, P. (1994). The cognitive architecture behind auditory-visual interaction in scene analysis and speech identification. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 13,* 69-75.

Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In M. Sabourin, F. I. M. Craik, & M. Robert (Eds.), *Advances in psychological science, Vol. 2: Biological and cognitive aspects* (pp. 419-439). Hove: Psychology Press.

Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Aschersleben, T. Bachman, & J. Musseler (Eds.), *Cognitive Contributions to the Perception of Spatial and Temporal Events* (pp. 347-362). Amsterdam: Elsevier Science Publishers.

Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin and Review, 5,* 482-489.

Bertelson, P., & Aschersleben, G. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension 1. Evidence from auditory-visual temporal order judgment. *International Journal of Psychophysiology, 50,* 147-155.

Bertelson, P., & de Gelder, B. (2004). The psychology of multisensory perception. In C. Spence & J. Driver (Eds.), *Crossmodal Space and Crossmodal Attention.* Oxford: Oxford University Press.

Bertelson, P., Frissen, I., Vroomen, J., & de Gelder, B. (2004). *The aftereffects of ventriloquism: Patterns of spatial generalization from local recalibration.* Manuscript submitted for publication.

Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & de Gelder, B. (2000a). Ventriloquism is patients with unilateral visual neglect. *Neuropsychologia, 38,* 1634-1642.

Bertelson, P., & Radeau, M. (1976). Ventriloquism, sensory interaction, and response bias: Remarks on the paper by Choe, Welch, Gilford, and Juola. *Perception & Psychophysics, 19,* 531-535.

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

Bertelson, P., Vroomen, J., & De Gelder, B. (1997). Auditory-visual interaction in voice localization and in bimodal speech recognition: The effects of desynchronization. *Proceedings of the Workshop on Audio-Visual Speech Processing*, Rhodes, Greece, 97-100.

Bertelson, P., Vroomen, J., & de Gelder, B. (2003). Visual recalibration of auditory speech identification: A McGurk aftereffect. *Psychological Science, 14,* 592-597.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000b). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics, 62,* 321-332.

Bertelson, P., Vroomen, J., Wiegeraad, G., & De Gelder, B. (1994). Exploring the relation between McGurk interference and ventriloquis. *International Congress on Spoken Language Processing*, Yokohama, Japan, 97-100.

Bisiach, E., & Vallar, G. (1988). Hemineglect in humans. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology: Vol. 1.* (pp. 195–222). Amsterdam: Elsevier.

Blackwell, H. R., & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology, 33,* 407-419.

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

Bock, O., & Burghoff, M. (1997). Visuo-motor adaptation: evidence for a distributed amplitude control system. *Behavioral Brain Research, 89,* 267-273.

Bohlander, R. W. (1984). Eye position and visual attention influence perceived auditory direction. *Perceptual & Motor Skills, 59,* 483-510.

Bolognini, N, Frassinetti, F., Serino, A., & Ladavas, E. (2005). "Acoustical vision" of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Experimental Brain Research, 160,* 273-282.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bronkhorst, A. W., & Houtgast, T (1999). Auditory distance perception in rooms. *Nature, 397,* 517-520.

Brugge, J. F., Dubrovsky, N. A., Aitkin, L. M., & Anderson, D. J. (1969). Sensitivity of single neurons in auditory cortex of the cat to binaural tonal stimulation: effects of varying interaural time and intensity. *Journal of Neurophysiology, 32,* 1005-1024.

Brungart. D. S. (1999). Auditory localization of nearby sources. III. Stimulus Effects. *Journal of the Acoustical Society of America, 106,* 3589-3602.

Brungart. D. S., Durlach, N. I., & Rabinowitz, W. (1999). Auditory localization of nearby sources. II. Localization of a broadband source. *Journal of the Acoustical Society of America, 106,* 1959-1968.

Brungart. D. S., & Rabinowitz, W. (1996). *Auditory localization in the near-field.* Proceedings of the 3rd International Conference on Auditory Display, Palo Alto, CA

Burke, K. A., Letsos, A., & Butler, R. A. (1994). Asymmetric performances in binaural localization of sound in space. *Neuropsychologia, 32,* 1409-1417.

Burns, E. M., & Ward, W. D. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (pp. 241-269). New York: Academic Press.

Caclin, A., Soto-Faraco, S., Kingstone, A., & Spence, C. (2002). Tactile "capture" of audition. *Perception & Psychophysics, 64,* 616-630.

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional imaging studies. *Cerebral Cortex, 11,* 1110-1123.

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes.* Cambridge, MA: MIT Press.

Campbell, R., Dodd, B., & Burnham, D. (Eds.). (1998). *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech.* Hove, UK: Psychology 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.

Canon, L. K. (1971). Directed attention and maladaptive "adaptation" to displacement of the visual field. *Journal of Experimental Psychology, 88,* 403–408.

Choe, C. S., Welch, R. B., Gilford, R. M., & Juola, J. F. (1975). The "ventriloquist effect": Visual discrimination or response bias? *Perception & Psychophysics, 18,* 55-60.

Cohen, H. B. (1966). Some critical factors in prism adaptation. *American Journal of Psychology, 79,* 285-290.

Cohen, Y. E., & Andersen, R. A. (2004). Multimodal spatial representations in the primate parietal lobe. In C. Spence & J. Driver (Eds.), *Crossmodal Space and Crossmodal Attention* (pp. 99-121). Oxford: Oxford University Press.

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.

Colin, C., Radeau, M., Soquet, A., Dachy, B., & Deltenre, P. (2002). Electrophysiology of spatial scene analysis: The mismatch negativity (MMN) is sensitive to the ventriloquism illusion. *Clinical Neurophysiology, 113,* 507-518.

Coren, S., Ward, L. M., & Enns, J. T. (1994). *Sensation and Perception (4th ed.).* Orlando, FL: Harcourt Brace.

Cornsweet, T. N. (1962) The staircase-method in psychophysics. *American Journal of Psychology, 75,* 485-491.

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

diRibaupierre, F. (1997). Acoustical information processing in the auditory thalamus and cerebral cortex. In G. Ehret & R. Romand (Eds.), *The central auditory system* (pp. 317-397). Oxford: Oxford University Press.

Dong, C., Swindale, N.V., & Cynader, M.S. (1999). A contingent aftereffect in the auditory system. *Nature Neuroscience, 2,* 863–865.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* New York: Academic Press.

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip reading. *Nature, 381,* 66-68.

Ehrenstein, W. H., & Reinhardt-Rutland, A. H. (1996). A cross-modal aftereffect: auditory displacement following adaptation to visual motion. *Perceptual & Motor Skills, 82,* 23-26.

Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology, 4,* 99-109.

Epstein, W. (1975). Recalibration by pairing: a process of perceptual learning. *Perception, 4,* 59-72.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429-433.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8,* 162-169.

Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics, 63,* 719-725.

Fernandez-Ruiz, J., & Diaz, R. (1999). Prism adaptation and aftereffect: specifying the properties of a procedural memory system. *Learning & Memory, 6,* 47-53.

Fernandez-Ruiz, J., Diaz, R., Hall-Haro, C., Vergara, P., Mischner, J., Nunez, L., Drucker-Colin, R., Ochoa, A., & Alonso, M. E. (2003). Normal prism adaptation but reduced after-effect in basal ganglia disorders using a throwing task. *European Journal of Neuroscience, 18,* 689-694.

Field, D. P., Shipley, T. F., & Cunningham, D. W. (1999). Prism adaptation to dynamic events. *Perception & Psychophysics, 61,* 161-176.

Frassinetti, F., Bolognini, N., & Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147,* 332-343.

Freedman, S. J., & Wilson, L. (1967). Compensation for auditory re-arrangement following exposure to auditory-tactile discordance. *Perceptual & Motor Skills, 25,* 861-866.

Freedman, S. J., Wilson, L., & Rekosh, J. H. (1967). Compensation for auditory re-arrangement in ear-hand coordination. *Perceptual & Motor Skills, 24,* 1207-1210.

Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2003). The aftereffects of ventriloquism: are they sound-frequency specific? *Acta Psychologica, 13,* 315-327.

Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2005). The aftereffects of ventriloquism: Generalization across sound-frequency. *Acta Psychologica, 118,* 93-100.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7,* 773-778.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1996). Generalization to local remappings of the visuomotor coordinate system. *Journal of Neuroscience, 16,* 7085-7096.

Godfroy, M., Roumes, C., & Dauchy, P. (2003). Spatial variations of visual-auditory fusion areas. *Perception, 32,* 1233-1245.

Goldstein, B. (1996). *Sensation and perception* (4th ed.). Pacific Grove, CA : Brooks/Cole.

Gutfreund, Y., & Knudsen, E. I. (2004). Visual instruction of the auditory space map in the midbrain. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 613-624). Cambridge, MA: MIT Press.

Hafter, E. R. (1984). Spatial hearing and the duplex theory: How viable is the model? In G. M. Edelman, W. M. Cowan, & W. E. Gall (Eds.), *Dynamic aspects of neocortical function* (pp 425-448). New York: Wiley.

Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., & Wallace, M. T. (2004). Multisensory enhancement of localization under conditions of induced myopia. *Experimental Brain Research, 152,* 404-408.

Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L., & Schirillo, J. A. (2003). Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience, 15,* 20-29.

Harris, C. S. (1963). Adaptation to displaced vision: Visual, motor, or proprioceptive vhange? *Science, 140,* 812-813.

Hartmann, W. M., & Rakerd, B. (1989). Localization of sound in rooms IV: The Franssen effect. *Journal of the Acoustical Society of America, 86,* 1366-73.

Hay, J. C., Langdon, B., & Pick, H. L. (1971). Spatial parameters of eye-hand adaptation. *Journal of Experimental Psychology, 91,* 11-17.

Heffner, H. E., & Heffner, R. S. (1990). Effects of bilateral auditory cortex lesions on sound localization in Japanese macaques. *Journal of Neurophysiology, 64,* 915-931.

Held, R. (1955). Shifts in binaural localization after prolonged exposures to atypical combinations of stimuli. *American Journal of Psychology, 68,* 526-548.

Held, R. (1961). Exposure-history as a factor in maintaining stability of perception and co-ordination. *Journal of Nervous and Mental Disease, 132,* 26-32

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

Held, R., & Hein, A. V. (1958). Adaptation of disarranged hand-eye coordination contingent upon re-afferent stimulation. *Perceptual & Motor Skills, 8,* 87-90.

Heron, J., Whitaker, D., & McGraw, P. V. (2004). Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research, 44,* 2875-2884.

Hershenson, M. (1989). Duration, time constant, and decay of the linear motion aftereffect as a function of inspection duration. *Perception & Psychophysics, 45,* 251-257.

Hershenson, M. (1993). Linear and rotation motion aftereffects as a function of inspection duration. *Vision Research, 33,* 1913-1919.

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science, 298,* 1627-1630.

Hovland, C. I. (1937). The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *Journal of General Psychology, 17,* 125-148.

Howard, I. P. (1982). *Human visual orientation.* New York: Wiley.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation.* London: Wiley.

Humphreys, L. G. (1939). Generalization as function of the method of reinforcement. *Journal of Experimental Psychology, 25,* 361-372.

Hyde, P. S., & Knudsen, E. I. (2002). The optic tectum controls visually guided adaptive plasticity in the owl's auditory space map. *Nature, 415,* 73-76.

Ignashchenkova, A., Dicke, P. W., Haarmeier, T., & Their, P. (2004). Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nature Neuroscience, 7,* 56-64.

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual & Motor Skills, 37,* 967-979.

Jackson, C. V. (1953). Visual factors in auditory localization. *The Quarterly Journal of Experimental Psychology, 5,* 52-65.

Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature, 309,* 345-347.

Jay, M. F., & Sparks, D. L. (1987). Sensory-motor integration in the primate superior colliculus. II. Coordinates of auditory signals. *Journal of Neurophysiology, 57,* 35-55.

Jenkins, W. M., & Merzenich, M. M. (1984). Role of cat primary auditory cortex for sound localization. *Journal of Neurophsyiology, 52,* 819-847.

Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., & Stein, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *Journal of Neurophysiology, 85,* 506-522.

Jousmaki, V., & Hari, R. (1998). Parchment-skin illusion: sound-biased touch. *Current Biology, 8,* R190.

Kaas, J. H., Hackett, T. A., & Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Current Opinion in Neurobiology, 9,* 174-170.

Kalil, R., & Freedman, S. J. (1967). Compensation for auditory re-arrangement in the absence of observer movement. *Perceptual & Motor Skills, 24,* 475-478.

Kavanagh, G. L., & Kelly, J. B. (1987). Contribution of auditory cortex to sound localization by the ferret (*mustela putorius*). *Journal of Neurophysiology, 57,* 1746-1766.

Keetels, M., & Vroomen, J. (2004, June). The role of space in audiovisual temporal order judgments. Poster presented at the 5th meeting of the International Multisensory Research Forum. Barcelona, Spain.

Kelso, J. A. S., Cook, E., Olson, M. E., & Epstein, W. (1975) Allocation of attention and the locus of adaptation to displaced vision. *Journal of Experimental Psychology: Human Pereption and Performance, 3,* 237-245.

King, A. J. (2002). Neural plasticity: how the eye tells the brain about sound location. *Current Biology, 12,* R393-395.

King A. J. (2004). The superior colliculus. *Current Biology, 9,* R335-R338.

King, A. J., & Doubell, T. P., & Skaliora, I. (2004). Epigenetic factors that align visual and auditory maps in the ferret midbrain. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 599-612). Cambridge, MA: MIT Press.

King, A. J., Schnupp, J. W. H., & Thompson, I. D. (1998). Signals from the superficial layers of the superior colliculus enable the development of the auditory space map in the deeper layers. *Journal of Neuroscience, 18,* 9394-9408.

Kitazawa, S., & Yin, P. B. (2002). Prism adaptation with delayed visual error signals in the monkey. *Expimental Brain Research, 144,* 258-261.

Klemm, O. (1909). Localisation von Sinnineidrucken bei disparaten Nebenreizen (localization of sensory impressions with disparate distractors). *Psychologische Studien (Wundt), 5,* 73-161.

Knudsen, E. I. (2002). Instructed learning in the auditory pathway of the barn owl. *Nature, 417,* 322-328.

Konishi, M. (2000). Study of sound localization by owls and its relevance to humans. *Comparative Biochemistry & Physiology Part A, 126,* 459-469.

Krahe, R., Larsen, O. N., & Ronacher, B. (2000). Directional hearing is only weakly dependent on the rise time of acoustic stimuli. *Journal of the Acoustical Society of America, 107,* 1067-1070.

Ladavas, E., & Pavani, F. (1998). Neuropsychological evidence of the functional integration of visual, auditory and proprioceptive spatial maps. *Neuroreport, 9,* 1195-1200.

Lakatos, S. (1995). The influence of visual cues on the localisation of circular auditory motion. *Perception, 24,* 457-465.

Lewald, J. (1997). Eye-position effect on directional hearing. *Behavioural Brain Research, 87,* 35-48.

Lewald, J. (1998). The effect of gaze eccentricity on perceived sound direction and its relation to visual localization. *Hearing Research, 115,* 206-216.

Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learning & Memory, 9,* 268-278.

Lewald, J., Ehrenstein, W. H., & Guski, R. (2001). Spatio-temporal constraints for auditory--visual integration. *Behavioural Brain Research, 121,* 69-79.

Lewald, J., Foltys, H., & Töpper, R. (2002). Role of the posterior parietal cortex in spatial hearing. *Journal of Neuroscience, 22,* RC207.

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research, 16,* 468-478.

Litovksy, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America, 106,* 1633-1654.

Luksch, H., Gauger, B., & Wagner, H. (2000). A candidate pathway for a visual instructional signal to the barn owl's auditory system. *Journal of Neuroscience, 20,* RC70.

Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J., & Thach, W. T. (1996). Throwing while looking through prisms. II. Specificity and storage of multiple gaze-throw calibrations. *Brain,119,* 1199-1211.

Mateeff, S., Hohnsbein, J., & Noack, T. (1985). Dynamic visual capture: Apparent auditory motion induced by a moving visual target. *Perception, 14,* 721-727.

Matin, L. (1972). Eye movements and perceived visual direction. In D. Jameson & L. M. Hurvich (Eds.), *Handbook of sensory physiology: Vol. VII/4. Visual Psychophysics* (pp. 331-380). Berlin: Springer.

Mazzoni, P., Bracewell, R. M., Barash, S., & Anderson, R. A. (1996). Spatially tuned auditory responses in are LIP of macaques performing delayed memory saccades to acoustic targets. *Journal of Neurophysiology, 75,* 1233-1241.

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

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

Meyer, G. F., & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *NeuroReport, 12,* 2557-2560.

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

Mikaelian, H. H. (1969). Adaptation to rearranged ear-hand coordination. *Perceptual & Motor Skills, 29,* 147-150.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research, 17,* 154-163.

Müller, J. R., Philiastides, M. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences of the USA, 102,* 524-529.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale, NJ: Lawrence Erlbaum.

Nijhawan. R. (1994). Motion extrapolation in catching. *Nature, 370,* 256-257.

Paavilainen, P., Karlsson, M. L., Reinikainen, K., Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography & Clinical Neurophysiology, 73,* 129-141.

Phillips, D. P., & Brugge, J. F. (1985). Progress in neurophysiology of sound localization. *Annual Review of Psychology, 36,* 245-274.

Pick, H. L. Jr., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics, 6,* 203-205.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience, 3,* 741-747.

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

Radeau, M. (1974). Adaptation au deplacement prismatique sur le base d'une discordance entre la vision et l'audition (Adaptation to prismatic displacement on the basis of a discordance between vision and audition). *L'Annee Psychologique, 74,* 23-34.

Radeau, M. (1985). Signal intensity, task context, and auditory-visual interactions. *Perception, 14,* 571-577.

Radeau, M. (1992). Cognitive impenetrability in auditory-visual interaction. In J. Alegria, D. Holender, J. Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 41-55). Amsterdam: Elsevier.

Radeau, M. (1994a). Auditory-visual spatial interaction and modularity. *Cahier de Psychologie Cognitive/Current Psychology of Cognition, 13,* 3-51.

Radeau, M. (1994b). Ventriloquism against audio-visual speech: or, where Japanese-speaking barn owls might help. *Current Psychology of Cognition/Cahier de Psychologie Cognitive,* 13, 124-140.

Radeau, M. & Bertelson, P. (1969). Adaptation à un déplacement prismatique sur la base de stimulations exafférentes en conflit. *Psychologica Belgica, 9,* 133-140.

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *The 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 & Psychophysics, 20,* 227-235.

Radeau, M., & Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. *Perception & Psychophysics, 22,* 137-146.

Radeau, M., & Bertelson, P. (1978). Cognitive factors and adaptation to auditory-visual discordance. *Perception & Psychophysics, 23,* 341-343.

Radeau, M., & Bertelson, P. (1987). Auditory-visual interaction and the timing of inputs: Thomas (1941) revisited. *Psychological Research, 49,* 17-22.

Rakerd, B., & Hartmann, W. M. (1986). Localization of sound in rooms III: Onset and duration effects. *Journal of the Acoustical Society of America, 80,* 1695-1706.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences of the USA, 97,* 11800-11806.

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

Recanzone, G. H. (2003). Auditory influences on visual rate perception. *Journal of Neurophysiology, 89,* 1078-1093.

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

Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science, 143,* 594-596.

Roller, C. A., Cohen, H. S., Kimball, K. T., & Bloomberg, J. J. (2001). Variable practice with lenses improves visuo-motor plasticity. Cognitive Brain Research, 12, 341-352.

Schröger, E. (1996). Interaural time and level differences: integrated or separate processing. *Hearing Research, 96,* 191-198.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317-1323.

Shinn-Cunningham, B. (2001). *Localizing sounds in rooms.* Proceedings of the ACM SIGGRAPH and EUROGRAPHICS, Snowbird, Utah, USA.

Shinn-Cunningham, B. (2003). *Acoustics and perception of sound in everyday environments.* Proceedings of the 3rd international workshop on spatial media, Aisu-Wakamatsu, Japan.

Singer, G., & Day, R. H. (1965). Temporal determinants of a kinesthetic aftereffect. *Journal of Experimental Psychology, 69,* 343-348.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport, 12,* 7-10.

Soto-Faraco, S., & Kingstone, A. (2004). Multisensory integration of dynamic information. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of Multisensory Processes* (pp. 49-67). Cambridge, MA: MIT Press.

Soto-Faraco, S., Kingstone, A., & Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia, 41,* 1847-1862.

Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., & Kingstone, A. (2002). The ventriloquist in motion: illusory capture of dynamic information across sensory modalities. *Cognitive Brain Research, 14,* 139-146.

Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Cross-modal dynamic capture: congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology: Human Perception & Performance, 30,* 330-345.

Spence, C., & Driver, J. (2000). Attracting attention to the illusory location of a sound: reflexive crossmodal orienting and ventriloquism. *Neuroreport, 11,* 2057-2061.

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention.* Oxford: Oxford University Press.

Stager, P., & Burton, A. (1964). Graphic recording of the spiral after-effect: A study of its magnitude and rate of decay. *Canadian Journal of Psychology, 18,* 118-125.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

Stein, B. E., Jiang, W., & Stanford, T. R. (2004). Multisensory integration in single neurons of the midbrain. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 243-264). Cambridge, MA: MIT Press.

Stein, B. E., Meredith, M. A., & Huneycutt, W. S., & McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience, 1,* 12-24.

Stein, B. E., Stanford, T. R., Wallace, M. T., Vaughan, J. W., & Jiang, W. (2004). Crossmodal spatial interactions in subcortical and cortical circuits. In C. Spence & J. Driver (Eds.), *Crossmodal Space and Crossmodal Attention* (pp. 25-50). Oxford: Oxford University Press.

Stekelenburg, J. J., Vroomen, J., & de Gelder, B. (2004). Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neuroscience Letters, 357,* 163-166.

Stratton, G. M. (1897). Vision without the inversion of the retinal image. *Psychological Review, 4,* 341-360.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature, 421,* 911.

Taylor, M. M. (1963). Tracking the decay of the after-effect of seen rotary movement. *Perceptual & Motor Skills, 16,* 119-129.

Thomas, G. J. (1941). Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology, 28,* 163-177.

Thurlow, W. R., & Jack, C. E. (1973). Certain determinants of the "ventriloquism effect." *Perceptual & Motor Skills, 36,* 1171-1184.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Ungan, P., Yagcioglu, S., & Goksoy, C. (2001). Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. *Clinical Neurophysiology, 112,* 485-498.

van Beers, R. J., Wolpert, D. M., & Haggard, P. (2002). When feeling is more important than seeing in sensorimotor adaptation. *Current Biology, 12,* 834-837.

van Linden, S., Vroomen, J., de Gelder, B., & Bertelson, P. (2004, October). *Visual recalibration of auditory speech versus selective speech adaptation: Different build-up courses.* Paper presented at the 8th International Conference on Spoken Language Processing (ICSLP), Jeju, Jeju Island, Korea.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001a). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychologica, 108,* 21-33.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001b). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics, 63,* 651-659.

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

Vroomen, J., & de Gelder, B. (2003). Visual motion influences the contingent auditory motion aftereffect. *Psychological Science, 14,* 257-361.

Vroomen, J., & de Gelder, B. (2004a). Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of Multisensory Processes* (pp. 141-150). Cambridge, MA: MIT Press.

Vroomen. J., & de Gelder, B. (2004b). Temporal ventriloquism: sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception & Performance, 30,* 513-518.

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research, 22,* 32-35.

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

Wallace, M. T. (2004). The development of multisensory integration. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 625-642). Cambridge, MA: MIT Press.

Wallace, M. T., & Stein, B. E. (1994). Cross-modal synthesis in the midbrain depends on input from cortex. *Journal of Neurophysiology, 71,* 429-432.

Wallach, H. (1968). Informational discrepancy as the basis of perceptual adaptation. In S. J. Freedman (Ed.), *The neuropsychology of spatially oriented behavior* (pp. 209-229). Homewood, IL: Dorsey Press.

Warren, D. H., Welch, R. B., & McCarthy, T. J. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics, 30,* 557-564.

Weerts, T. C., & Thurlow, W. R. (1971). The effects of eye position and expectation on sound localization. *Perception & Psychophysics, 9,* 35-39.

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

Welch, R. B. (1986). Adaptation to space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 24.1-24.45). New York: Wiley & Sons.

Welch, R. B. (1994). The dissection of intersensory bias: Weighting for Radeau. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 13,* 117-123.

Welch, R. B. (1999). Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Musseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 371-387). Amsterdam: Elsevier.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88,* 638-667.

Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listeners and source movement. *Journal of the Acoustical Society of America, 105,* 2841-2853.

Wilkinson, L. K., Meredith, M. A., & Stein, B. E. (1996). The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Experimental Brain Research, 112,* 1-10.

Willey, C. F., Inglis, E., & Pearce, C. H. (1937). Reversal of auditory localization. *Journal of Experimental Psychology, 20,* 114-130.

Witkin, H. A., Wapner, S., & Leventhal, T. (1952). Sound localization with conflicting visual and auditory cues. *Journal of Experimental Psychology, 43,* 58-67.

Woods, T. M., & Recanzone, G. H. (2004). Visually induced plasticity of auditory spatial perception in macaques. *Current Biology, 14,* 1559-1564.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology.* New York: Holt.

Wright, B. A., & Fitzgerald, M. B. (2001). Different patterns of human discrimination learing for two interaural cues to sound-source location. *Proceedings of the National Academy of Science of the USA, 98,* 12307-12312.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General, 129,* 291-307.

Yin, P. B., & Kitazawa, S. (2001). Long-lasting aftereffects of prism adaptation in the monkey. *Expimental Brain Research, 141,* 250-253.

Yost, W. A. (2000). *Fundamentals of hearing: An introduction* (4th ed.). San Diego, CA: Academic Press.

Yost, W. A., & Sheft, S. (1993). Auditory perception. In R. R. Fay & A. N. Popper (Series Eds.) & W. A. Yost, A. N. Popper., & R. R. Fay (Vol. Eds.), *Springer handbook of auditory research: Vol 3. Human Psychophysics* (pp. 193-236). New York: Springer.

Young, P. T. (1928). Auditory localization with acoustical transposition of the ears. *Journal of Experimental Psychology, 11,* 399-429.

Zampini, M., Shore, D. I., & Spence, C. (2003). Multisensory temporal order judgments: The role of hemispheric redundancy. *International Journal of Psychophysiology, 50,* 165-180.

Zurek, P. M. (1993). A note on onset effects in binaural hearing. *Journal of the Acoustical Society of America, 93,* 1200-1201.

Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. (2001). Two-dimensional sound-localization behavior of early-blind humans. *Experimental Brain Research,* 140, 206-222.

Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nature Neuroscience, 6,* 175-181.

Acknowledgments

After the technicalities of science let's do the fun stuff.

There is, of course, no way, I could present this work to the reader pretending to have done it on my own. Many people have, either directly or indirectly, contributed and this is the time and place to mention them (and trying not to forget anyone, in which case I apologize and ask the victim to please send me a note so I can make up for it by buying him/her a nice cold beer to ease the pains of disappointment).

Let me start with people that have been most central to this endeavor, Bea de Gelder, Jean Vroomen and Paul Bertelson. There is no way to sum up all the things I learned about doing scientific research from you, so I am not going to. Each of you has a specific flair for doing things and I hope that some of it rubbed off on me and hopefully you'll recognize it in my future works. But let me give the more impatient reader a quick (and by no means complete) peek into what I am hoping for. First, Bea's apparently never-ending appetite for new avenues of exploration. Then, Jean's uncanny ability to see through the merits and problems of an experimental procedure at a glance. Also, Paul's eloquent way of saying things the way they should be said. Not unimportant when your main instrument is the pen (or more precisely the word-processor).

Experimental research cannot be done without the proper equipment and here I have to thank the technical support people, but one person deserves particular praise. Charles Rambelje, you made my life a lot easier, and on the way thought me one or two things about electronics and miscellaneous electrical stuff.

And now the part everyone checks out first, the co-workers! I am not even going to try to make a list because I am sure to forget someone. So, if we had lunch together or went to the "Esplanade" at one point, consider yourself a contributor to my good time at Tilburg. But there are of course some people I want to mention by name. Jeroen, I had a great time being your roommate for the past couple of years. Viona, for being a friend and all the drinks downtown. The psychonomics people (in geographical order), Ruthger, Jan, Sabine, Geert, Ton, Karin, Franc, Mirjam, and Wim. But also, Seger, Marijke, Hellen, Wendy, and Chris. Let me repeat the already mentioned liquid compensation in case I forgot you.

Curriculum Vitae

Ilja Frissen was born on 25th of July in the year 1973 in Sittard, but was raised in Maastricht. After finishing high-school in 1991 he continued his education at a laboratory school. But the lab did not appeal enough and he therefore started his studies of psychology at Maastricht University in 1996, and finished in 2000 with a master's degree in Cognitive Psychology. During this period he spent the fall semester of 1999 at the University of Illinois at Urbana-Champaign. In 2000 he spent half a year at TNO-Human Factors in Soesterberg, from which resulted the master's thesis on human sound localization. In December 2000 he worked as a PhD student at Tilburg University, with as a result the present thesis. As of March 2005 he is a postdoctoral fellow at the Max Planck Institute for Biological Cybernetics in Tübingen.

SBN: 90-9019587-4