available at www.sciencedirect.com
ScienceDirect
www.elsevier.com/locate/brainres
ELSEVIER
BRAIN RESEARCH

Research Report

# Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level

*Rike Steenken*[a,*], *Hans Colonius*[a], *Adele Diederich*[b], *Stefan Rach*[a,b]

[a]*Department of Psychology, University of Oldenburg, P.O. Box 2503, 26111 Oldenburg, Germany*
[b]*School of Humanities and Social Sciences, Jacobs University Bremen, Germany*

ARTICLE INFO

*Article history:*
Accepted 8 August 2007

*Keywords:*
Multisensory integration
Cross-modal interaction
Saccadic eye movement
Auditory localization
Auditory masking

ABSTRACT

Saccadic reaction time (SRT) to a visual target tends to be shorter when auditory stimuli are presented in close temporal and spatial proximity, even when subjects are instructed to ignore the auditory non-target (focused attention paradigm). Observed SRT reductions typically range between 10 and 50 ms and decrease as spatial disparity between the stimuli increases. Previous studies using pairs of visual and auditory stimuli differing in both azimuth and vertical position suggest that the amount of SRT facilitation decreases not with the physical but with the *perceivable* distance between visual target and auditory accessory. Here we probe this hypothesis by presenting an additional white-noise masker background of 3 s duration. Increasing the masker level had a diametrical effect on SRTs in spatially coincident vs. disparate stimulus configurations: saccadic responses to coincident visual–auditory stimuli are slowed down, whereas saccadic responses to disparate stimuli are speeded up. As verified in a separate auditory localization task, localizability of the auditory accessory decreases with masker level. The SRT results are accounted for by a conceptual model positing that increasing masker level enlarges the area of possible auditory stimulus locations: it implies that perceivable distances decrease for disparate stimulus configurations and increase for coincident stimulus pairs.


## 1. Introduction

A key observation in multisensory research is that saccadic reaction time (SRT) to a visual target tends to be facilitated when auditory stimuli are presented in close temporal and spatial proximity. In fact, crossmodal effects on these fast voluntary movements of the eyes have been the subject of numerous studies on multisensory integration over the last 10–15 years (for recent reviews see, for example, the chapters by van Opstal and Munoz, 2004; Diederich and Colonius, 2004; Gutfreund and Knudsen, 2004; Wallace, 2004). In the *focused attention paradigm* participants are typically instructed to make a saccade as quickly and as accurately as possible toward a visual target stimulus suddenly appearing at a random position off the fixation point and, simultaneously, to ignore any co-occurring stimuli from other modalities (auditory or tactile). Observed SRT reductions usually range between 10 and 50 ms and decrease as the spatial and temporal separation between the stimuli increases (Colonius and Arndt, 2001; Corneil and Munoz, 1996; Frens et al., 1995; Harrington and Peck, 1998; Hughes et al., 1998).

* *Corresponding author*. Fax: +49 441 798 5172.
E-mail address: rike.steenken@uni-oldenburg.de (R. Steenken).
*Abbreviations:* SRT, Saccadic reaction time; SC, superior colliculus; dSC, intermediate/deep layers of the superior colliculus; RF, receptive field; SOA, stimulus onset asynchrony

0006-8993/$ – see front matter 
doi:10.1016/j.brainres.2007.08.034

The neural correlates of multisensory integration and orienting behavior have been found primarily in the midbrain superior colliculus (SC) (see Stein and Meredith, 1993 for review). Neurons in the intermediate/deep layers of the SC (dSC) are involved in the initiation and control of saccades (for review, see Munoz and Fecteau, 2002; Sparks, 1999; Sparks et al., 2001). Information about stimulus location is represented topographically – the horizontal dimension is mapped rostrocaudally, the vertical dimension mediolaterally on the SC (Middlebrooks and Knudsen, 1984) – by an arrangement of neurons according to the location of their receptive fields (RFs). Many of these same neurons exhibit multisensory activity paralleling the spatio-temporal rules found in behavioral studies (King and Palmer, 1985; Meredith and Stein, 1986; Populin and Yin, 2002; Wallace et al., 1996). The spatial register between the auditory and visual sensory maps is formed by multisensory neurons whose different RFs are in register with one another yielding a common frame of reference (Stein and Meredith, 1993). Their firing rates are greatest for spatially aligned stimuli and decrease in magnitude as spatial disparity increases (Bell et al., 2001; Frens and Van Opstal, 1998; Meredith and Stein, 1996). These sensory maps are also in register with the premotor maps found in SC (e.g., McIlwain, 1986), and many SC neurons are involved in both sensory and motor maps.[1] A recent study by Bell and colleagues (Bell et al., 2005) demonstrated a particularly close link between changes in neural activity related to stimulus modality with changes in gazing behavior of alert monkeys (see Discussion).

The focus of this paper is on a particular aspect of multisensory integration rules, i.e., the effect of spatial disparity between visual and auditory stimuli on saccadic eye movements. In studies with animal recordings, basic multisensory spatial effects have been explained in terms of (non-) overlapping receptive field structures: When both stimuli originate from the same location in space, they are likely to fall within the respective excitatory RFs of the same multisensory SC neuron, thereby triggering a superadditive response enhancement of the neuron's activity. For spatially disparate presentations, when the stimulus of one modality is falling outside the borders of the neuron's RF, response depression occurs as a result of the antagonism between an inhibitory input derived from activation of the extra-receptive field region of that stimulus and the excitatory input from the within-field stimulus of the other modality (Stein, 1998; Kadunce et al., 1997). Note that the graded decline of response enhancement of SRT typically observed in behavioral studies is consistent with these RF mechanisms if one assumes that orienting behavior is the collective result of a potentially large number of multisensory neurons with gradually shifted RFs. On the other hand, there are many phenomena of crossmodal spatial effects found in behavioral studies that do not seem to be easily reducible to the receptive field structure in the superior colliculus, the "ventriloquist effect" (Bertelson and Radeau, 1981) perhaps being the most prominent one.

Here we study, within a focused attention task, how the effect of visual–auditory spatial disparity on SRT relates to the localizability of the auditory stimulus. Note that, in principle, participants are not required to localize the auditory stimulus at all in this paradigm, since they are only asked to move the eyes toward the visual target while ignoring any accessory auditory stimulation. Nevertheless, e.g., Frens et al. (1995) found that saccadic latencies increased with about 0.5 ms per degree stimulus separation, up to about 35 ms, using broadband white noise as auditory stimuli. Note that the fact that one can find such a lawful relationship suggests that the magnitude of visual–auditory spatial disparity has a direct and mandatory access to the mechanism of multisensory integration under the focused attention condition. More specifically, the Frens et al. (1995) study revealed that it is not the physical but rather the *perceived* (Frens et al., 1995, pp. 807–808)[2] distance between visual and auditory stimulus position that matters. In fact, with 700-Hz tones as auditory accessory, only the horizontal separation determined the strength of the SRT reduction, whereas the actual vertical position did not play a role. This corresponds to the fact that the elevation of tonal acoustical stimuli is not reflected in the audio-oculomotor system (Frens and Van Opstal, 1995). Whereas the azimuth of a sound source is derived from binaural cues, such as interaural timing and intensity differences, estimating the elevation component is based on spectral filtering by the pinnae/head and, in the case of a tonal stimulus, this monaural cue cannot deliver unambiguous information on the vertical sound source position (Wightman and Kistler, 1989; Blauert, 1997).

Information on the *time course* of the effect of localizability on SRT enhancement comes from the study by Heuermann and Colonius (2001). They presented visual–auditory stimulus pairs varying both in elevation and azimuth with SOAs ranging from −60 to 40 ms, where negative values mean that the auditory accessory was presented prior to the visual target. As no maskers were presented, the auditory white-noise stimuli were easily localizable with maximal bimodal enhancement for the coincident condition at −60 ms. Interestingly, there was no difference in the level of enhancement between pairs differing in azimuth only and pairs differing in both elevation and azimuth when the auditory stimulus was presented simultaneous to or after the visual stimulus, although bimodal enhancement was still observable under both conditions. Presumably, when the auditory stimulus was presented "too late" there was no time for the elevation component to be computed by the pinnae/head system – which is slower than the horizontal system – utilizing binaural cues, so that saccade initiation was already under way. In other words, the perceivable distances between the visual and auditory stimulus at the time of saccade initiation were the same in the two types of configuration.

In this study, we further probe the dependence of cross-modal SRT enhancement on the perceivable distance between visual and auditory stimuli and, thereby, on the localizability of the auditory stimulus position by manipulating the background masker level. To ensure that the participants are able to localize the acoustical stimulus in elevation (Frens and Van Opstal, 1995; Heuermann and Colonius, 2001) when no

[1] While neurons in the rostral pole of SC are involved in fixation, neurons at more caudal sites direct the eyes, ears, and head progressively more contralateral, specifically in the cat. Those in medial SC direct movements upward and those in lateral SC direct movements downward (cf. Stein, 1998).

[2] We prefer the term "perceivable" to the more common "perceived" because it is not clear whether an act of (conscious) perception is involved.

Please cite this article as: Steenken, R., et al., Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level, Brain Res. (2007), doi:10.1016/j.brainres.2007.08.034

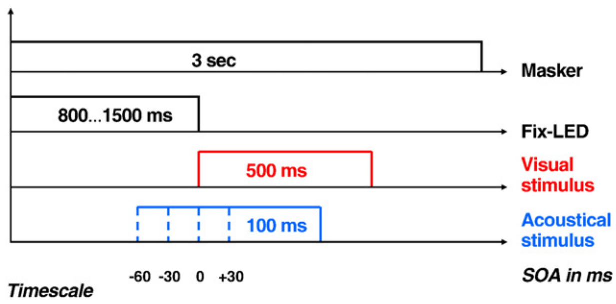


**Fig. 1 – Time course of a trial. Onset of fixation LED and masker defined the start of a trial. Visual (target) stimulus of 500 ms duration was presented simultaneous to the offset of the fixation point. Onset of an acoustical accessory (of 100 ms duration) occurred at 4 different levels of stimulus onset asynchrony (SOA) relative to the target.**

masker is presented, we administered broadband pink noise bursts (1/f noise, range 100 Hz to 22 kHz) of 100 ms length as auditory accessory stimuli together with visual target light-emitting diodes (LEDs). Fig. 1 depicts the time course of a trial.

In order to modulate the perceivable distance between the stimuli, the noise bursts were masked by uncorrelated white noise that was switched on at the beginning of the trial and lasted for 3 s. Therefore, as masker level increases, localizability of the auditory accessory should decrease leading to a modification of perceivable visual–auditory distance and, by hypothesis, to a corresponding change in SRT enhancement. Good and Gilkey (1996), for example, have shown that with decreasing signal-to-noise ratio the percentage of possible responses of the participants in an up/down judgment task is more and more uncorrelated with the target indicating that the perceivable distances broaden with increasing masker level.

As localization of acoustical stimuli in elevation is already affected at higher signal-to-noise ratios compared to the left–right dimension (Good and Gilkey, 1996; Lorenzi et al., 1999), spatial positions were only varied in the vertical axis (one position above and one below the fixation point) facilitating the manipulation of the localizability in a stepwise manner. In the *coincident* condition, visual target and auditory accessory were both presented at the top or at the bottom position, whereas in the *disparate* condition the stimulus of one modality was presented at the top and the stimulus of the other modality was presented at the bottom position. Let us assume now that with the growth of the masker level the size of the point-image of the auditory stimulus (i.e., the area of its possible locations, or its location 'volume') grows as well. This should have opposite effects under the two configuration conditions: in the coincident condition, the diminishing localizability of the auditory stimulus with increasing masker level should allow the occurrence of larger and larger perceivable distances between visual and auditory stimulus even though both stimuli remain at their (nearly) identical physical position. By contrast, in the disparate condition the masker level increase should allow the occurrence of smaller and smaller perceivable distances between the stimuli even though their physical vertical distance remains invariant.

In order to check whether localizability does in fact change in the anticipated manner, a separate auditory localization task was given to the participants using the same visual–auditory spatial configurations but without any speed stress. Localization judgments were recorded in blocks of trials interspersed with those from the saccadic reaction time experiment.

## 2. Results

### 2.1. *Saccadic reaction times*

Most notably, mean SRTs are affected by increasing masker level in two diametric ways: in the coincident condition, they tend to increase across SOAs, whereas in the disparate condition there is a tendency of the SRTs to speed up. This result is in line with the prediction based on the presumed localizability of the auditory accessory derived above. As an additional check, we analyzed all SRTs with respect to the absolute spatial position of the target (top vs. bottom vertical location). Except for a general slow-down for saccades directed upwards (about 20 ms), no effects other than the ones already described were found, however.

Trials with saccades in the wrong direction were excluded from the analysis (less than 2.4% of all data). An analysis of variance on SRTs was performed with subject as random factor. We further defined three ANOVA factors as follows: (1) configuration with levels *coincident* (visual and auditory stimulus presented at the same position), *disparate* (visual

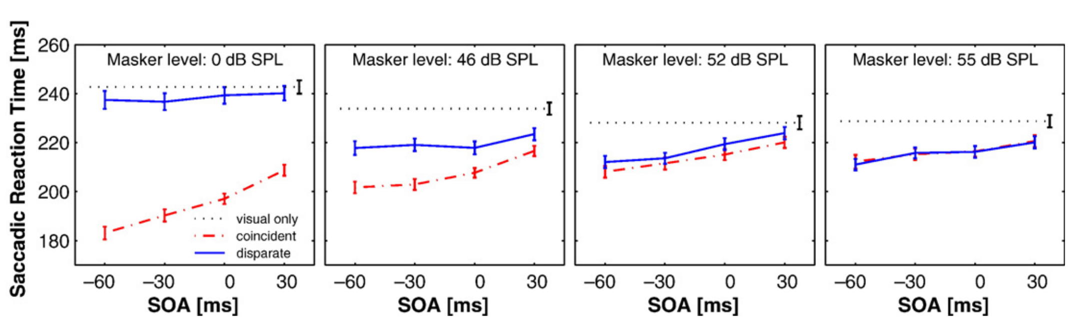


**Fig. 2 – Mean saccadic reaction time. Each panel refers to a specific level of the white-noise masker, from 0 dB to 55 dB (from left to right). Mean SRTs (±standard error) are plotted against stimulus onset asynchrony (SOA), separately for coincident (dashed line) and disparate (continuous line) stimulus configurations. Mean SRTs to unimodal visual target presentation are indicated by the dotted lines in each panel. Data are averaged over 4 subjects (excluding participant HV).**

Please cite this article as: Steenken, R., et al., Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level, Brain Res. (2007), doi:10.1016/j.brainres.2007.08.034

and auditory stimulus presented at vertically opposite positions), and *visual-only*, (2) masker level with levels of 0, 46, 52, and 55 dB (SPL), and (3) stimulus onset asynchrony (SOA) with levels −60, −30, 0, 30 ms. While the subject factor was significant, there was no significant interaction between subjects and the other three factors (configuration and subject ($F(8, 9.8)=2.5$; $p<0.09$); masker and subject ($F(12, 4.9)=2.2$; $p<0.198$); SOA and subject ($F(12, 3.14)=1.36$; $p<0.442$)) suggesting that there were general interindividual differences in the speed but no specific distinct response pattern concerning the experimental factors. Main effects of configuration ($F(2, 8)=85$; $p<0.0001$) and of SOA ($F(3, 12)=50$, $p<0.0001$) were significant, as was the two-way interaction between masker and configuration ($F(6, 24)=67$, $p<0.0001$) indicating that the difference in SRT between the coincident and disparate condition diminishes over increasing masker level. A significant interaction was also observed between SOA and configuration ($F(6,24)=7$, $p<0.005$) (mean SRTs in the *disparate* condition are almost constant over the different SOAs whereas the facilitative influence of the auditory accessory in the *coincident* condition decreases with increasing SOAs). The two-way interaction of SOA and masker level failed to reach significance ($F(9, 36)=2$, $p<0.065$) indicating that no interaction between the time course of the effect of localizability on SRT enhancement and the perceivable distance exists. Post-hoc tests (Tukey) revealed that mean SRT to a visual target was shorter in the presence of the accessory ($p<0.0001$) and shorter as well when target and accessory were presented *coincident* rather than *disparate* ($p<0.0001$). There was a significant difference in SRT between −60 and −30 ms SOA ($p<0.006$) and ($p<0.002$) between the negative SOAs (−60 and −30 ms) and the positive SOAs (0 and 30 ms).

Mean SRTs for the *coincident* and the *disparate* condition, as a function of SOA and Masker level, and averaged over all participants (but see footnote 4), are depicted in Fig. 2.

Finally, a separate analysis revealed that SRTs to unimodal visual stimuli decrease with increasing masker level ($p<0.00$).

Next, we consider the results from the separate localization task.

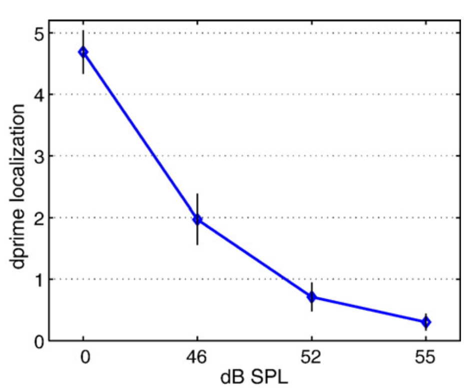


**Fig. 3 – Auditory localizability as a function of masker level. For each level of the auditory white-noise masker, *d*-prime values (excluding participant HV) calculated for each subject separately and standard errors were computed from the frequencies of responses about position of the auditory stimulus ("top" or "bottom").**

### 2.2. Localization accuracy

For each of the 4 possible auditory–visual stimulus configurations AV, av, Av, and aV (where the capital letters refer to the top position, the lower-case letters to the bottom position) the relative frequency of responses "acoustic stimulus at top position" was computed at every level of the auditory masker and for each participant. Given that participants exhibited a very similar pattern of results, we combined their data (participant HV, however, whose performance was practically errorless, was analyzed separately[3]). These frequencies were used to estimate the probabilities *P*(response "acoustic stimulus at top position"|AV or Av) and *P*(response "acoustic stimulus at top position"|aV or av), from which a *d*-prime value was computed as measure of localization accuracy. Fig. 3 depicts the mean *d*-prime value (calculated for each of the four subjects separately) and standard error as a function of masker level showing a decrease of auditory localizability from perfect to marginal. A more detailed analysis, not presented here, shows that this decline cannot be accounted for by a response bias possibly caused by the visual stimulus position (cf. Hairstone et al., 2003; Bertelson and Radeau, 1981).

Interviewing our subjects made it obvious that (i) the acoustical stimulus was clearly detectable under all noise levels and that (ii) they had a "spatial impression" in most trials. In those cases where they were uncertain about the location of the auditory target they were asked to guess.

## 3. Discussion

In the focused attention task, saccadic reaction time to a visual target is accelerated in the presence of an auditory accessory stimulus. It has also been established in several studies that increasing spatial disparity between the visual and auditory stimulus has a diminishing effect on this speed-up. However, the results by Frens et al. (1995) have suggested that it is not the physical distance *per se* but rather the *perceivable* distance that determines this effect. Using different levels of broadband white noise masking an auditory accessory (pink noise) the study presented here lends further support to this hypothesis. As we verified in an auditory localization task separate from the SRT measurements, localizability of the auditory accessory in elevation decreases with masker level. At the same time, decreasing localizability of the accessory had a diametrical effect on SRTs in spatially coincident vs. disparate stimulus configurations: saccadic responses to coincident visual–auditory stimuli are slowed down, whereas saccadic responses to disparate stimuli are speeded up. Note that this behavior is elucidated by defining an effective *perceivable distance* in the following way: each visual or auditory stimulus in the medial plane is assigned to a set of possible stimulus locations, *V* or

[3] This participant was highly trained and may have become sensitive to some specific cues emitted by the loudspeakers. Note that his SRT data were also not included in the aggregated data of Fig. 2 but, consistent with our hypothesis and his high localizability score, he showed even higher SRT effects of configuration than the other participants.

ARTICLE IN PRESS



A, respectively. Increasing the masker level enlarges the area of possible locations for the auditory stimulus (see Fig. 4). Next, we define the set-theoretic difference,

$$(V - A) \cup (A - V),$$

i.e., the set of locations corresponding to either a visual or an auditory stimulus, but not to both. The effective *perceivable distance* on the sets is defined as $D(V,A)=m(V-A)+m(A-V)$, where $m$ is some additive numerical function on the sets of possible stimulus locations.[4] Interestingly, $D$ predicts the SRT changes observed: Increasing masker levels correspond to decreasing set differences in the disparate presentation and to increasing set differences in the coincident presentation leading to smaller perceivable distances in the former, to larger perceivable distances in the latter. Note that the concept of defining perceivable distance in terms of sets of possible stimulus locations may be useful in unifying the large number of empirical results on multisensory spatial fusion (e.g., Godfroy et al., 2003; Wallace et al., 2004).

Note that the time course of mean SRT over SOA for the coincident and the disparate configuration converge with increasing masker level reflecting, according to our hypothesis, a growing correspondence of the perceivable distances under both conditions.

Moreover, the speeding up of SRT in the visual-only condition with increasing masker level could be due to a preparation enhancement (Nickerson, 1973) or general warning effect (Diederich and Colonius, 2006) with the masker serving as a spatially unspecific warning cue. An alternative – more cognitive or context-dependent (see Nickerson, 1973) – explanation could be that a general strategy of participants is to try to avoid errors. Because the clearly localizable auditory stimulus presented here has a stimulative nature to respond in its direction, the participants per default delay the response until they are secure about the correctness of their response. If now the masker is added this stimulative nature of the auditory stimulus deteriorates and hence the need to withhold the response to the visual stimulus leading to an overall decrease in SRT in the visual-only condition. Note that the same interpretation could hold for the decreasing of SRT in the disparate condition with increasing masker level, whereas it contradicts the increase of SRT in the coincident condition.

To sum up, there is converging evidence that perceivable distance drives the crossmodal SRT effect in elevation, but it remains unclear how these distance values are processed by the multisensory integration mechanism. In the time-window-of-integration (TWIN) model of multisensory interaction in saccades, recently proposed by two of the authors (Colonius and Diederich, 2004; Diederich and Colonius, 2007a), crossmodal properties like distance are processed in a second stage, after termination of a peripheral stage where the occurrence of crossmodal interaction depends on whether or not the sensory information from stimuli of different modalities is registered within a certain window of time (of the order of 200 ms). Since here the masker was switched on well before the occurrence of the visual–auditory stimulus pair, the masker would not be expected to elicit crossmodal interaction with the target stimulus, but it could have a general alerting or warning effect. (The fact that the SRTs in the visual-only condition decrease with increasing masker level supports this idea.) The spatially specific contribution of the masker level, however, would be effected through a modification of the perceivable distance controlling the magnitude of multisensory enhancement in the second stage of TWIN (see Diederich and Colonius, 2007b, for a modeling of spatial integration rules for visual–tactile stimulus configurations).

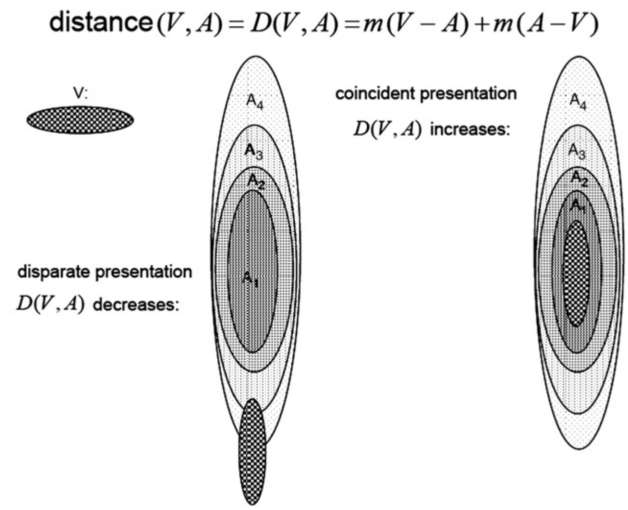


**Fig. 4 – Perceived distance between visual and auditory stimulus in elevation. Hypothetical areas of possible stimulus locations for visual stimulus and, for each level of white-noise masker, of auditory accessory. Set function *m* indicates the number of possible stimulus locations in each subset. Note that perceivable distance *D* decreases, respectively increases, with growing auditory areas for disparate, respectively coincident, presentation.**

Further research on the effect of masker level should also consider stimuli of varying intensity. It is well known that stimulus intensity strongly influences the magnitude of crossmodal interactions: they are often most pronounced for pairings of weaker stimuli (principle of "inverse effectiveness"; Meredith and Stein, 1986; see also Rach and Diederich, 2006, for a behavioral variant). Using auditory and visual stimuli of different intensities, presented either in spatial alignment or to opposite hemifields, Bell et al. (2005) found that spatially aligned audiovisual stimuli evoked the shortest SRTs. In the case of low intensity stimuli, the response to the auditory component of the aligned audiovisual target increased the activity preceding the response to the visual component, accelerating the onset of the visual response and facilitating the generation of shorter-latency saccades. In the case of high-intensity stimuli, the auditory and visual responses occurred much closer together in time and so there was little opportunity for the auditory stimulus to influence pre-visual activity. Instead, the reduction in SRT for high-intensity, aligned audiovisual stimuli was correlated with increased premotor activity.

[4] This can be made precise by postulating appropriate probability distributions.

Please cite this article as: Steenken, R., et al., Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level, Brain Res. (2007), doi:10.1016/j.brainres.2007.08.034

## 4. Experimental procedures

### 4.1. Participants

Five students (3 female), aged 21–31, served as paid voluntary participants. All had normal or corrected-to-normal vision, 3 participants were right-handed (self-description), and all were naïve as to the purpose of the study. They were screened for their ability to follow the experimental instructions (proper fixation, few blinks during trial, saccades towards visual target and the ability to localize an acoustical stimulus in the elevation, respectively). They gave their informed consent prior to inclusion in the study. The experiment was conducted in accordance with the ethical standards described in 1964 Declaration of Helsinki.

### 4.2. Apparatus and stimulus presentation

Two red light-emitting diodes (LED, 0.34 cd/m$^2$) mounted on two loudspeakers served as visual targets. They were placed 38.5 cm (22.6°) of visual angle above or below a fixation point (a red LED, 14.5 cd/m$^2$, distance to monitor 100 cm). Sound stimuli (amplified by a NAD, Stereo Integrated Amplifier 306) were a pink noise (49 dB SPL, sampling rate 44,100) and a running uncorrelated white noise of four intensities (0 dB, 46 dB, 52 dB, 55 dB SPL, sampling rate 44,100) presented via two speakers (Canton Plus XS) vertically aligned to the fixation point. The speakers were driven by a Creative Labs Audigy2 soundcard. The infrared light video system Eyelink II (SR Research) running on two personal computers (DELL Dimension 3100) was used to control the experiment and to record responses; manual responses were registered by the Eyelink ButtonBox.

### 4.3. Procedure: SRT measurements

The experiment was carried out in a completely darkened, sound-attenuated chamber. The participant sat on a chair with the head on a chin rest. Every session began with 10 min of dark adaptation during which the eye movement measurement system was adjusted and calibrated. Each trial began with the appearance of the fixation point. After a variable fixation time (800–1500 ms), the fixation LED disappeared and, simultaneously, the visual stimulus was turned on above or below the fixation point (no gap). Participants were instructed to gaze at the visual target as quickly and as accurately as possible ignoring the auditory stimuli (focused attention paradigm). Depending on the particular condition, the visual target and the masker appeared alone or in combination with the auditory accessory (coincident or disparate configuration). The onset of the accessory was shifted relative to the visual target by a stimulus onset asynchrony of −60, −30, 0, or 30 ms (negative values mean that the accessory was presented prior to the target). The visual targets were presented for 500 ms, the accessory for 100 ms. The masker was turned on at the beginning of each trial lasting for 3 s. Trials were separated by a break of 2 s. Saccadic eye movements were recorded with a temporal resolution of 500 Hz and horizontal and vertical spatial resolution of 0.01°. Criteria for saccade detection on a trial by trial basis were velocity (35°/s) and acceleration (>9500°/s$^2$). The recordings from each trial were checked for proper fixation at the beginning of the trial, eye blinks, and correct detection of start and endpoint of the saccade. Saccades were screened for anticipation errors (SRT<80 ms), misses (SRT>500 ms), and direction/localization errors.

### 4.4. Procedure: auditory localization task

For this task, stimulus configurations identical to the ones used for the SRT measurements were presented, except for leaving out the visual-only conditions with various masker levels. Two responses were solicited: First, participants had to move the eyes towards the perceived (top or bottom) position of the acoustical stimulus; second, they had to indicate via button press whether the acoustical stimulus was presented from top or bottom. There was no emphasis on speed other than to finish both tasks within the 3 s duration of the masker. Moreover, participants were allowed to give two conflicting responses. (They made slightly more errors when they had to respond via saccadic eye movement, the characteristic of the response pattern was not affected. Because the saccadic response was only required to keep the conditions constant between the SRT task and localization task it was not further analyzed.)

### 4.5. Presentation schedule

After extensive training of both tasks (80 min), each participant completed 10 blocks of 216 trials in the SRT task, resulting in a total of 2160 trials (60 trials per condition) and 7 blocks of 192 trials in the localization task, resulting in a total of 1344 trials (42 trials per condition). Presentation order in each block was completely randomized over conditions. All blocks were presented in a pseudo-randomized order. Each block lasted about 20 min, and the entire data collection was spread over 2 weeks.

## REFERENCES

Bell, A.H., Corneil, B.D., Meredith, M.A., Munoz, D.P., 2001. The influence of stimulus properties on multisensory processing in the awake primate superior colliculus. Can. J. Exp. Psychol. 55, 123–132.

Bell, A.H., Meredith, M.A., Van Opstal, A.J., Munoz, D.P., 2005. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. J. Neurophysiol. 93, 3659–3673.

Bertelson, P., Radeau, M., 1981. Crossmodal bias and perceptual fusion with auditory–visual spatial discordance. Percept. Psychophys. 29, 578–584.

Blauert, J., 1997. The Psychophysics of Human Sound Localization, 2nd Edition. MIT Press, Cambridge.

Colonius, H., Arndt, P., 2001. A two-stage model for visual–auditory interaction in saccadic latencies. Percept. Psychophys. 63, 126–147.

Colonius, H., Diederich, A., 2004. Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J. Cogn. Neurosci. 16, 1000–1009.

Corneil, B.D., Munoz, D.P., 1996. The influence of auditory and visual distractors on human orienting gaze shifts. J. Neurosci. 16, 8193–8207.

Diederich, A., Colonius, H., 2004. Modeling the time course of multisensory interaction in manual and saccadic responses. In: Calvert, G., Spence, C., Stein, B.E. (Eds.), Handbook of Multisensory Processes. MIT Press, Cambridge, pp. 395–408.

Please cite this article as: Steenken, R., et al., Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level, Brain Res. (2007), doi:10.1016/j.brainres.2007.08.034

Diederich, A., Colonius, H., 2006. Separating multisensory integration from unspecific warning effects in saccadic reaction time. Paper presentation, IMRF.
Diederich, A., Colonius, H., 2007a. Why two "Distractors" are better than one: modeling the effect of non-target auditory and tactile stimuli on visual saccadic reaction time. Exp. Brain Res. 179, 43–54.
Diederich, A., Colonius, H., 2007b. Modeling spatial effects in visual–tactile saccadic reaction time. Percept. Psychophys. 69, 56–67.
Frens, M.A., Van Opstal, A.J., 1995. A quantitative study of auditory-evoked saccadic eye movements in two dimensions. Exp. Brain Res. 107, 103–117.
Frens, M.A., Van Opstal, A.J., 1998. Visual–auditory interactions modulate saccade-related activity in monkey superior colliculus. Brain Res. Bull. 46, 211–224.
Frens, M.A., Van Opstal, A.J., Van der Willigen, R.F., 1995. Spatial and temporal factors determine auditory–visual interactions in human saccadic eye movements. Percept. Psychophys. 57, 802–816.
Godfroy, M., Roumes, C., Dauchy, P., 2003. Spatial variations of visual–auditory fusion areas. Perception 32, 1233–1245.
Good, M.D., Gilkey, R.H., 1996. Sound localization in noise: the effect of signal-to-noise ratio. J. Acoust. Soc. Am. 99, 1108–1117.
Gutfreund, Y., Knudsen, E.I., 2004. Visual instruction of the auditory space map in the midbrain. In: Calvert, G., Spence, C., Stein, B.E. (Eds.), Handbook of Multisensory Processes. MIT Press, Cambridge, pp. 613–624.
Hairstone, 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. J. Cogn. Neurosci. 15, 20–29.
Harrington, L.K., Peck, C.K., 1998. Spatial disparity affects visual–auditory interactions in human sensorimotor processing. Exp. Brain Res. 122, 247–252.
Heuermann, H., Colonius, H., 2001. Spatial and temporal factors in visual–auditory interaction. In: Sommerfeld, E., Kompass, R., Lachmann, T. (Eds.), Proceedings of the 17th Meeting of the International Society for Psychophysics. Pabst Science, Lengerich, pp. 118–123.
Hughes, H.C., Nelson, M.D., Aronchick, D.M., 1998. Spatial characteristics of visual–auditory summation in human saccades. Vis. Res. 38, 3955–3963.
Kadunce, D.C., Vaughan, J.W., Wallace, M.T., Benedek, G., Stein, B.E., 1997. Mechanisms of within-modality and cross-modality suppression in the superior colliculus. J. Neurophysiol. 78, 2834–2847.
King, A.J., Palmer, A.R., 1985. Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. Exp. Brain Res. 60, 492–500.
Lorenzi, C., Gatehouse, S., Lever, C., 1999. Sound localization in normal-hearing listeners. J. Acoust. Soc. Am. 99, 1810–1820.
McIlwain, J.T., 1986. Effect of eye position on saccades evoked electrically from superior colliculus of alert cats. J. Neurophysiol. 55, 97–112.
Meredith, M.A., Stein, B.E., 1986. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J. Neurophysiol. 56, 640–662.
Meredith, M.A., Stein, B.E., 1996. Spatial determinants of multisensory integration in cat superior colliculus neurons. J. Neurophysiol. 75, 1843–1857.
Middlebrooks, J.C., Knudsen, E.I., 1984. A neural code for auditory space in the cat's superior colliculus. J. Neurosci. 4, 2621–2634.
Munoz, D.P., Fecteau, J.H., 2002. Vying for dominance: dynamic interactions control visual fixation and saccadic initiation in the superior colliculus. Prog. Brain Res. 140, 3–19.
Nickerson, R.S., 1973. Intersensory facilitation of reaction time: energy summation or preparation enhancement? Psychol. Rev. 80, 489–509.
Populin, L.C., Yin, T.C., 2002. Bimodal interactions in the superior colliculus of the behaving cat. J. Neurosci. 22, 2826–2834.
Rach, S., Diederich, A., 2006. Visual–tactile integration: does stimulus duration influence the relative amount of response enhancement? Exp. Brain Res. 173, 514–520.
Sparks, D.L., 1999. Conceptual issues related to the role of the superior colliculus in the control of gaze. Curr. Opin. Neurobiol. 9, 698–707.
Sparks, D.L, Freedman, E.G., Chen, L.L., Gandhi, N.J., 2001. Cortical and subcortical contributions to coordinated eye and head movements. Vis. Res. 41, 3295–3305.
Stein, B.E., 1998. Neural mechanisms for synthesizing sensory information and producing adaptive behavior. Exp. Brain Res. 123, 124–135.
Stein, B.E., Meredith, A.M., 1993. The Merging of the Senses. MIT Press, Cambridge.
Van Opstal, A.J., Munoz, D.P., 2004. Auditory–visual interactions subserving primate gaze orienting. In: Calvert, G., Spence, C., Stein, B.E. (Eds.), Handbook of Multisensory Processes. MIT Press, Cambridge, pp. 373–394.
Wallace, M.T., 2004. The development of multisensory integration. In: Calvert, G., Spence, C., Stein, B.E. (Eds.), Handbook of Multisensory Processes. MIT Press, Cambridge, pp. 625–642.
Wallace, M.T., Wilkinson, L.K., Stein, B.E., 1996. Representation and integration of multiple sensory inputs in primate superior colliculus. J. Neurophysiol. 76, 1246–1266.
Wallace, M.T., Roberson, G.E., Hairston, W.D., Stein, B.E., Vaughan, J.W., Schirillo, J.A., 2004. Unifying multisensory signals across time and space. Exp. Brain Res. 158, 252–258.
Wightman, F.L., Kistler, D.J., 1989. Headphone simulation of free field listening I: stimulus synthesis. J. Acoust. Soc. Am. 85, 858–867.

Please cite this article as: Steenken, R., et al., Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level, Brain Res. (2007), doi:10.1016/j.brainres.2007.08.034