Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Neuroscience Letters xxx (2008) xxx–xxx
Neuroscience Letters
www.elsevier.com/locate/neulet

Time course of auditory masker effects: Tapping the locus of audiovisual integration?

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

[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*

Received 27 November 2007; received in revised form 22 January 2008; accepted 6 February 2008

Abstract

In a focused attention paradigm, saccadic reaction time (SRT) to a visual target tends to be shorter when an auditory accessory stimulus is presented in close temporal and spatial proximity. Observed SRT reductions typically diminish as spatial disparity between the stimuli increases. Here a visual target LED (500 ms duration) was presented above or below the fixation point and a simultaneously presented auditory accessory (2 ms duration) could appear at the same or the opposite vertical position. SRT enhancement was about 35 ms in the coincident and 10 ms in the disparate condition. In order to further probe the audiovisual integration mechanism, in addition to the auditory non-target an auditory masker (200 ms duration) was presented before, simultaneous to, or after the accessory stimulus. In all interstimulus interval (ISI) conditions, SRT enhancement went down both in the coincident and disparate configuration, but this decrement was fairly stable across the ISI values. If multisensory integration solely relied on a feed-forward process, one would expect a monotonic decrease of the masker effect with increasing ISI in the backward masking condition. It is therefore conceivable that the relatively high-energetic masker causes a broad excitatory response of SC neurons. During this state, the spatial audio–visual information from multisensory association areas is fed back and merged with the spatially unspecific excitation pattern induced by the masker. Assuming that a certain threshold of activation has to be achieved in order to generate a saccade in the correct direction, the blurred joint output of noise and spatial audio–visual information needs more time to reach this threshold prolonging SRT to an audio–visual object.

Keywords: Multisensory integration; Auditory temporal masking; Focused attention paradigm; Saccadic reaction time; Warning effect

---

In their natural environment, humans and animals alike are commonly quite effective in integrating information being transmitted from their sensors specialized for particular modalities like vision and audition. Combined information from more than one modality tends to be more reliable than information from a single modality alone. However, in natural environments multisensory target stimuli are often accompanied by external noise presumably affecting the process of multisensory integration in some way. In the laboratory situation, focusing on a single modality can be part of the instructions for a particular task. Specifically, in a bimodal setup with saccadic reaction time (SRT) [9] as the main dependent variable of interest, visual stimuli are often designated as targets that have to be responded to as quickly and as accurately as possible, whereas stimuli from the other, accessory modality (e.g., auditory) should not be paid any attention. This focused attention (FA) paradigm has emerged as an important testing ground to investigate the spatiotemporal rules of multisensory integration underlying the generation and execution of saccadic eye movements [6,26]. The issue addressed here is how and to what degree these rules of crossmodal interaction in SRT may be affected by acoustical background noise.

A key finding in multisensory eye movement research is that average SRT to a visual target (about 150–250 ms) tends to be shorter when an auditory stimulus is presented in close temporal and spatial proximity: Observed SRT reductions typically range between 10 and 50 ms, and the effect decreases as spatial and temporal disparity between the stimuli increases [3,4,10,13,16]. As explanatory mechanisms for this behavioral effect, concepts like warning, statistical facilitation, or genuine crossmodal coactivation have been proposed [6]. Neuronal correlates in the form of multisensory neurons have been found in the midbrain superior colliculus (SC) and in a wide range of (sub-)cortical regions [8,24,25]. Neurons in the

---

[*] Corresponding author. Tel.: +49 441 798 5152; fax: +49 441 798 5172.
E-mail address: rike.steenken@uni-oldenburg.de (R. Steenken).

0304-3940/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.neulet.2008.02.017

intermediate/deep layers of the SC (dSC) that are involved in the initiation and control of saccades have overlapping receptive fields (RF) forming a multisensory map of space such that a given location in the SC corresponds to a given location in space [15,19]. Their firing rates are greater for spatially aligned stimuli and decrease in magnitude as the overlap between the receptive field becomes smaller [1,11,18]. Moreover, their response onset latency can be shortened by multisensory stimuli compared to unimodal stimuli thereby contributing to a reduction of observable reaction time [2,22].

Notwithstanding the wealth of results on the neurophysiology of multisensory neurons, the correspondence between their activity and behavioral indices like SRT or threshold measurements is not as close as one might expect [14]. For example, a spatially disparate acoustic stimulus not falling within the auditory receptive field of a multisensory neuron often causes a depression in the neuron's response to an excitatory visual stimulus [24], whereas SRT to a visual stimulus in an FA paradigm is rarely delayed by presenting an accessory, spatially disparate acoustic stimulus.

It appears that very few studies have addressed the effect of varying levels of acoustical background noise on multisensory integration in saccadic responses in a visual–auditory stimulus configuration in any systematic way. Corneil et al. [5] presented visual, auditory, and spatially coincident bimodal targets on a complex audiovisual noise background. They observed increasing SRT gain for bimodal stimuli with higher auditory noise levels, i.e., decreasing signal-to-noise (S/N) ratios, relative to the unimodal responses. A recent study from our lab [23] found effects of a white-noise masker (of 3 s duration) on bimodal SRT to depend on spatial configuration: saccadic responses towards a visual target stimulus that was presented on the medial plane (top or bottom position) were slowed down with increasing masker level when an acoustical non-target (100 ms pink-noise) was presented spatially coincident, whereas saccadic responses with the auditory non-target presented spatially disparate were facilitated when masker level increased. Note that this somewhat amazing result becomes reconcilable with the ubiquitous inhibitory effect of spatial disparity between target and non-target on SRT in the FA paradigm once it is assumed that the essential inhibitory factor is not the physical but, rather, the perceived – or *perceivable* (see [23]) – distance between target and non-target: In accordance with previous studies (e.g., [12]), we found a diminishing vertical localizability of the auditory (non-target) stimulus with decreasing S/N ratio. This suggests that a high-level masker blurs the perceivable distance between visual target and auditory non-target mitigating the inhibitory effect of spatial disparity on SRT facilitation.

What remains unclear, as yet, is the level of perceptual processing at which the masker affects the visual–auditory integration of spatially coincident or disparate stimuli. Given that target and non-target were always embedded within the 3 s time interval of the masker in Steenken et al. [23], the latter may have operated at a peripheral level even before any multisensory integration could have occurred. In order to better gauge the time course of the masker effect on audiovisual interaction, here a FA paradigm with an auditory temporal masker presented simultaneous, immediately before, or up to 40 ms after an auditory non-target was employed. Importantly, masker and non-target did not overlap except in the simultaneous condition.

Subjects were asked to make a saccade as quickly and as accurately as possible toward a visual target LED appearing at the top or bottom position in the medial plane. Note that because localization of acoustical stimuli in elevation is already affected at higher signal-to-noise ratios compared to the left–right dimension [5,12] spatial position was only varied in the vertical axis to facilitate masking of the localizability. Participants were instructed to ignore the acoustical non-target presented either coincident or opposite to the visual target.

Five students (one male), aged 21–31, served as paid voluntary participants. All had normal or corrected-to-normal vision, four had right eye dominance and three were right-handed (self-description). They were screened for their ability to follow the experimental instructions (few blinks during trial, and saccades towards visual target). All participants gave their informed consent prior to their inclusion in the study. The experiment was conducted in accordance with the ethical standards described in 1964 Declaration of Helsinki.

The experiment took place in a completely darkened, sound-attenuated chamber. Fixation point was a red LED (13 cd/m^2, distance to the monitor 107 cm) positioned at eye level. Two red-light-emitting diodes (0.34 cd/m^2) fixed on two loudspeakers served as visual targets presented for 500 ms. LEDs were placed 21° above and below the fixation point. Acoustic stimuli were presented via two loudspeakers (Canton Plus XS) vertically aligned to the fixation point. Auditory non-target was a pink noise (1/*f* noise, 45 dB SPL, frequency range 100–22,000 Hz, sampling rate 44,100 Hz) presented for 2 ms, and a running uncorrelated white noise (52 dB SPL, sampling rate 44,100 Hz) served as auditory masker presented for 200 ms. This masker was perceived by the subjects as a spatially diffuse sound distributed somewhere between the two loudspeakers.

Eye movements were controlled and recorded by the infrared light video system Eyelink-II (SR Research) running on two PCs (DELL Dimension-3100); they 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.

The participant sat on a chair with the head on a chin rest preventing head movements. 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). The visual target appeared alone or in combination with the auditory accessory in the coincident or disparate configuration, either with or without masker. All conditions were equally likely to occur.

Please cite this article in press as: R. Steenken, et al., Time course of auditory masker effects: Tapping the locus of audiovisual integration?, Neurosci. Lett. (2008), doi:10.1016/j.neulet.2008.02.017

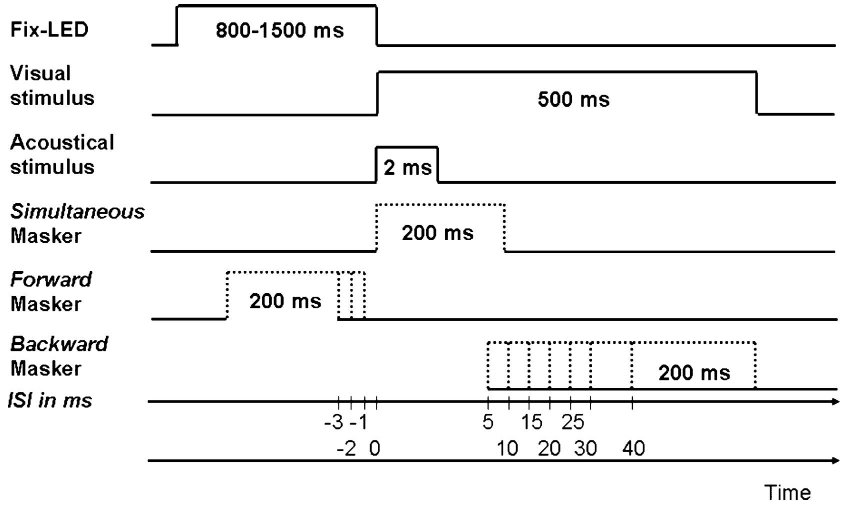

Fig. 1. Time course of a trial. After fix-LED offset, the visual target was either presented alone, or with masker, or with masker and (coincident or disparate) acoustical non-target. The masker was presented before, simultaneously, or after the presentation of the acoustical non-target. There was no temporal overlap between masker and non-target except for the simultaneous condition.

Visual target and auditory non-target always started simultaneously. The interstimulus interval (ISI) between masker and the auditory non-target was varied: In the *forward masking* condition, the masker ended 3, 2 or 1 ms before the presentation of the auditory accessory; in the *simultaneous masking* condition, both auditory stimuli were presented at the same time; and in the *backward masking* condition, the masker followed the end of presentation of the accessory by 5–40 ms (Fig. 1). Trials were separated by a break of 2000 ms. This choice of ISI values was based on testing the effectiveness of the temporal masker in a pilot experiment. In particular, the backward masking was found to be more efficient than the forward masking.

After extensive training (60 min), each participant completed nine blocks of 288 trials, resulting in a total of 2592 trials (72 trials/condition). There were a total of 36 conditions: 11 ISIs times three configurations (coincident/disparate/visual-only) plus three configurations without masker (coincident/disparate/visual-only). Presentation order in each block was completely randomized over conditions. Each block lasted about 20 min, and the entire data collection was spread over two weeks.

Trials with anticipations (SRT <80 ms), misses (SRT >500 ms), hypometric saccades (amplitude <5°), and direction errors were excluded from the analysis (1.8% of all data). Among selected saccades, only about 1.5% were followed by a secondary saccade toward the target. Given this low rate of occurrence, the following analyses were focused on the primary saccades. Inspecting the SRT distributions, we did not find any indication for express saccades. Moreover, no change in accuracy was observed across the various stimulus conditions.

Fig. 2a shows mean SRTs averaged over all subjects. The three lines indicate mean SRTs for the conditions without the masker (visual-only, coincident, and disparate audiovisual condition).

Without the masker, there was a clear audio–visual facilitation of about 35 ms for the coincident and 10 ms for the disparate condition, relative to the visual-only SRT of 240 ms.

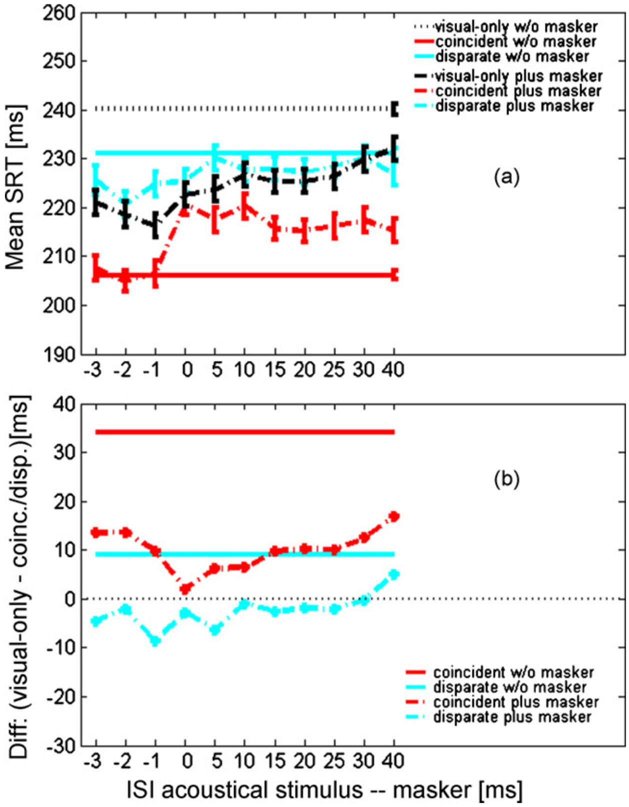

Fig. 2. (a) Mean SRT over five participants (±standard error) across all spatial and masker conditions as a function of ISI. Horizontal lines indicate mean SRT when no masker was present. (b) Net spatial audio–visual interaction effects: Plotting the data of panel (a) after subtracting mean SRT of the visual-only-plus-masker condition (from coincident/disparate-plus-masker curves) and after subtracting mean SRT of the visual-only condition (from coincident/disparate-without-masker lines). Spatial audio–visual interaction takes place when the curves or lines lie above (facilitation) or below (inhibition) the black dotted reference line.

Please cite this article in press as: R. Steenken, et al., Time course of auditory masker effects: Tapping the locus of audiovisual integration?, Neurosci. Lett. (2008), doi:10.1016/j.neulet.2008.02.017

The three curves show, in the presence of the masker, the dependence of SRT on the masker's time of occurrence relative to the simultaneously presented auditory non-target, i.e., the ISI, and the visual target. The bottom curve presents the spatially coincident condition indicating multisensory facilitation relative to both the visual-only-plus-masker condition and the spatially-disparate-plus-masker condition. Comparing the spatially coincident and disparate conditions with and without the masker, it appears as if the masker had an opposite effect under the two spatial configurations: slowing down mean SRT in the coincident condition (at least, for positive ISI values) and somewhat facilitating mean SRT in the disparate condition. However, this interpretation fails to take into account the facilitative effect the masker, as an auditory spatially unspecific accessory stimulus, produces in the visual-only condition: it speeds up responses by 10–20 ms depending on the visual target relative to masker ISI. The actual degree to which the masker influences the spatial interaction between target and non-target is better seen by plotting mean SRTs with reference to the visual-only-plus-masker curve as a baseline (see Fig. 2b: in the coincident condition there is an – albeit reduced – facilitative effect of the auditory non-target across all ISI values), whereas in the disparate condition, there is a tendency for a slight inhibition, except for the ISI = 40 ms value.

No spatial effect is discernible in the simultaneous masking condition (ISI = 0); for all positive ISI values (backward masking), there is a relatively constant difference between coincident and disparate mean SRTs of about 10 ms, i.e., the original spatial effect is reduced from about 25 ms to about 10 ms. In the forward masking conditions (ISI = −3, −2, −1), the spatial effect is slightly larger, about 15–20 ms. The two lines in Fig. 2b indicate the coincident and disparate difference mean SRTs without the masker, relative to the visual-only condition without the masker. The vertical distance of about 25 ms between the two allows gauging the reduction of the spatial effect generated by the masker. Moreover, the figure makes it plain that SRT is more severely impaired in the coincident than in the disparate configuration.

A repeated-measures analysis of variance on SRTs was performed with factors: (1) configuration with levels *coincident* (i.e., visual and auditory stimulus presented at the same position), *disparate* (i.e., visual and auditory stimulus presented at vertically opposite positions), and *visual-only*; (2) masker with levels *absent* and *present*; and (3) interstimulus interval (ISI) with levels −3, −2, −1, 0, 5, 10, 15, 20, 25, 30, and 40 ms (where negative values mean that the masker ended prior to the presentation of the acoustical stimulus), and subject as random factor. The main effect of masker was significant ($F(1,4) = 46$, $p < 0.002$) indicating that participants were on average faster when the masker was present. Factor configuration also reached significance ($F(2,8) = 39$, $p < 0.0001$), but a significant interaction with the factor subject ($F(8,7) = 4$, $p < 0.038$) reveals some subject-specific effects. Whereas mean SRT to a visual target was generally shorter when target and accessory were presented coincident rather than disparate, there was no consistent ordering between the disparate and visual-only conditions across participants.

The interaction between masker and configuration was significant ($F(2,8) = 37$, $p < 0.0001$), but a three-way interaction with subject ($F(8,80) = 5$, $p < 0.0001$) precludes a consistent interpretation across participants. All participants slowed down in the coincident condition in the presence of a masker, but again there was no consistent ordering between the other configurations across subjects. Random factor subject was significant reflecting interindividual differences in absolute SRTs. The main effect of ISI was significant ($F(10,40) = 5$, $p < 0.0001$), but the interaction of ISI and configuration was not ($p = 0.129$), indicating that the spatial effects were stable over the entire time range. Comparing forward masking (negative ISIs) with backward masking (positive ISIs) in a separate analysis revealed the masking direction as the driving source of the ISI main effect (216 ms vs. 224 ms, respectively, $p < 0.0001$). In addition, all SRTs were analyzed with respect to their absolute spatial position (top or bottom). Except for a general slow-down for upwards directed saccades of about 15 ms, no further effects were found.

In a focused attention task, saccadic reaction time to a visual target has previously been shown to be facilitated in the presence of a spatially coincident auditory accessory stimulus [3,10]. This has been replicated here with an auditory stimulus with a duration as short as 2 ms. Moreover, no improvement in audiovisual saccade accuracy was observed, compared to responses to visual targets alone. Typically, in a more natural environment humans and animals are faced with additional auditory noise stimuli. Probing for possible effects on multisensory interaction in such an enriched situation, several studies have observed that SRT facilitation for spatially coincident stimuli diminishes with decreasing S/N ratio within a simultaneous masking paradigm [5,23]. This paradigm has been extended here utilizing a temporal masking paradigm where a (200 ms white noise) masker was presented before, after, or simultaneous with the non-target. It turned out that the masker reduces multisensory facilitation across the entire range of ISI values (Fig. 2b).

In order to understand the role of the masker in spatiotemporal interaction between visual target and auditory non-target, one must distinguish the effect of the masker as a possible additional auditory cue from its effect on the visual–auditory integration mechanism. Comparing the visual-only SRT with the visual-only-plus-masker SRT discloses a facilitatory effect of about 20 ms that has the hallmark of a temporal warning effect: it is most pronounced when the masker terminates 3, 2, or 1 ms before target onset, and it gradually diminishes when the onset of the masker lags the target onset up to 40 ms (Fig. 2a).

Once this warning effect of the masker has been taken into account, its decelerating impact on the multisensory spatial SRT effect becomes visible in the coincident as well as in the disparate configuration (see Fig. 2b). This finding differs from a recent study in our lab [23] where visual targets and auditory non-targets were presented in identical spatial conditions: we found that a simultaneous masker had a diametrical effect on SRT in the coincident and disparate spatial condition. Whereas SRT in the coincident condition was slowed down, it was speeded up in the disparate condition with increasing noise level. A sepa-

Please cite this article in press as: R. Steenken, et al., Time course of auditory masker effects: Tapping the locus of audiovisual integration?, Neurosci. Lett. (2008), doi:10.1016/j.neulet.2008.02.017

rate auditory localization task revealed a particular sensitivity to masking noise in the vertical dimension, as previously found in Good and Gilkey [12]. The slow-down of SRT in the coincident condition of Steenken et al. [23] may therefore have been a consequence of the impaired localizability of the auditory non-target with increasing background noise. The simultaneous speed-up of SRT in the spatially disparate condition, however, may have had its source in a strategic behavior of the participants: in order to suppress a saccadic response to the auditory non-target (49 dB SPL noise of 100 ms duration) in the disparate condition, participants had to inhibit their oculomotor system somewhat if they were to keep erroneous responses at the required low level; an increasing background noise level, however, reduces the stimulative nature of the non-target and, thereby, the tendency to direct gaze in the wrong direction allowing participants to relax their inhibitory control accordingly. The discrepancy between the two studies may be explained by the physical properties of the auditory non-targets: here a rather weak auditory stimulus (2 ms duration) was employed with a very low chance of triggering a saccade in the wrong direction. Therefore, the presumed inhibition of the oculomotor system in the first study was not necessary for the participants to avoid errors. Support for this hypothesis comes from comparing the SRTs in the visual-only condition of three subjects participating in both experiments. Here, these subjects were, on average, 25 ms faster although the energy of the visual stimulus was the same.

The most interesting finding of this study, besides demonstrating a weakening effect of the masker on multisensory enhancement, is the observation that the amount of this effect remains rather unaffected by delaying the onset of the masker up to 40 ms after presentation of the target and the non-target. This suggests that whatever the masker does to the mechanism of visual–auditory interaction, it will have the same effect whether it occurs 5 ms or 40 ms after the onset of the target/non-target pair. That timing is critical is demonstrated by the ISI = 0 condition, the only condition when masker and auditory non-target overlap temporally: the spatial effect of multisensory SRT breaks down completely and the masker seems to act as a non-specific warning signal [7].

Note that this virtual invariance of the backward masking effect on multisensory enhancement hints at an involvement of feedback processes in visual–auditory integration [8,17]. Specifically, if multisensory integration solely relied on a feed-forward (or bottom-up) process, then one would expect a monotonic decrease of the masker effect with increasing ISIs rather than the constancy as observed here.

On a neuronal level, it is therefore conceivable that the relatively high-energetic backward masker causes a broad excitatory response of SC neurons [5] that lasts for a certain time. During this state, the spatial audio–visual information from multisensory association areas is fed back and merged with the unspecific excitation pattern induced by the masker. Assuming that a certain threshold of activation has to be achieved in order to generate a saccade in the correct direction, the blurred joint output of noise and spatial audio–visual information needs more time to reach that threshold and thereby prolongs SRT to an audio–visual object in a noisy environment.

It may be instructive to relate this hypothesis to some more specific estimates of the timing and presumed feedback interactions of the suggested subprocesses.[1] After about 18 ms, an acoustical stimulus (here, the auditory non-target) is already delivered to the intermediate deep layers of the SC (dSC) [27], and after about 28 ms to the auditory cortex A1 [20]. It is known that top-down influences of the anterior ectosylvian sulcus (AES) and/or the lateral suprasylvian cortex rLS are necessary for multisensory integration to occur on the neuronal (dSC) as well as on the behavioral level. Note that although the association area contains multisensory neurons, the unimodal neurons almost exclusively send their projections to the dSC and are therefore responsible for multisensory integration in the midbrain structure [25]. Concerning the auditory information this suggests that the non-target activity is sent to the auditory subarea of the AES (FAES) and from there back to the dSC. Bell et al. [2] observed that the earliest response onset latencies to spatially aligned audio–visual stimuli that determine a saccade in the correct direction are seen at about 80 ms. However, since here the masker's duration was 200 ms it could innervate the SC over a time span from 5 ms (in the backward masking condition, with respect to the termination of the acoustical stimulus) to 240 ms. When the feedback process from the AES carrying spatial audio–visual information [25,21] arrives during this constant and broad excitation of the SC, it results in a reduced spatial resolution leading to the reduction in multisensory SRT effects observed in our data.

Alternatively, it cannot be excluded that the relatively weak auditory stimulus is processed more slowly than the high energetic masker. Within the limited headstart of up to maximally 40 ms, the masker could catch up with the auditory non-target merging into a less specific representation of spatial auditory information. In that case, a pure feed-forward process would be compatible with the present results. However, given that an auditory stimulus is (on average) delivered to A1 already after about 28 ms [20], it seems questionable whether the masker can really compensate for the headstart.

To sum up, a masker impaired spatial audio–visual interaction in all three temporal masking conditions: in the simultaneous condition, spatial interaction was destroyed, whereas in the forward and backward condition a reduced level of interaction was still preserved. The fact that the impact of the masker remained rather invariant across the backward conditions points to an involvement of feedback processes in multisensory integration. Taking into account the published estimates about neural processing times suggests an interplay between cortical and subcortical areas even for a basic object binding task with audio-visual stimuli as simple as utilized here.

Acknowledgement

This research was supported by Deutsche Forschungsgemeinschaft: International Graduate School for Neurosensory

---

[1] Note, however, that these estimates were derived from different species (humans, cat, monkey) and different measurement techniques.

Please cite this article in press as: R. Steenken, et al., Time course of auditory masker effects: Tapping the locus of audiovisual integration?, Neurosci. Lett. (2008), doi:10.1016/j.neulet.2008.02.017

Science and Systems, Oldenburg, GK 591 (R.S.); grants Di 506/8-1 and Di 506/8-3 (A.D.), and SFB-TR31 (H.C.).

References

[1] A.H. Bell, B.D. Corneil, M.A. Meredith, D.P. Munoz, The influence of stimulus properties on multisensory processing in the awake primate superior colliculus, Can. J. Exp. Psychol. 55 (2001) 123–132.
[2] A.H. Bell, M.A. Meredith, A.J. Van Opstal, D.P. Munoz, Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements, J. Neurophysiol. 93 (2005) 3659–3673.
[3] H. Colonius, P. Arndt, A two-stage model for visual-auditory interaction in saccadic latencies, Percept. Psychophys. 63 (2001) 126–147.
[4] B.D. Corneil, D.P. Munoz, The influence of auditory and visual distractors on human orienting gaze shifts, J. Neurosci. 16 (1996) 8193–8207.
[5] B.D. Corneil, M. Van Wanrooij, D.P. Munoz, A.J. Van Opstal, Auditory-visual interactions subserving goal-directed saccades in a complex scene, J. Neurophysiol. 88 (2002) 438–454.
[6] A. Diederich, H. Colonius, Modeling the time course of multisensory interaction in manual and saccadic responses, in: G. Calvert, C. Spence, B.E. Stein (Eds.), Handbook of Multisensory Processes, MIT Press, Cambridge, 2004, pp. 395–408.
[7] A. Diederich, H. Colonius, Crossmodal interaction in saccadic reaction time: separating multisensory from warning effects in the time window of integration model, Exp. Brain Res., in press, doi:10.1007/s00221-007-1197-4.
[8] J.J. Foxe, C.E. Schroeder, The case for feedforward multisensory convergence during early cortical processing, Neuroreport 16 (2005) 419–423.
[9] M.A. Frens, A.J. Van Opstal, A quantitative study of auditory-evoked saccadic eye movements in two dimensions, Exp. Brain Res. 107 (1995) 103–117.
[10] M.A. Frens, A.J. Van Opstal, R.F. Van der Willigen, Spatial and temporal factors determine auditory–visual interactions in human saccadic eye movements, Percept. Psychophys. 57 (1995) 802–816.
[11] M.A. Frens, A.J. Van Opstal, Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus, Brain Res. Bull. 46 (1998) 211–224.
[12] M.D. Good, R.H. Gilkey, Sound localization in noise: the effect of signal-to-noise ratio, J. Acoust. Soc. Am. 99 (1996) 1108–1117.
[13] L.K. Harrington, C.K. Peck, Spatial disparity affects visual-auditory interactions in human sensorimotor processing, Exp. Brain Res. 122 (1998) 247–252.
[14] N.P. Holmes, The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability, Neuropsychologia 45 (2007) 3340–3345.
[15] N.P. Holmes, C. Spence, Multisensory integration: space, time, and superadditivity, Curr. Biol. 15 (2005) 762–764.
[16] H.C. Hughes, M.D. Nelson, D.M. Aronchick, Spatial characteristics of visual-auditory summation in human saccades, Vis. Res. 38 (1998) 3955–3963.
[17] E. Macaluso, Multisensory processing in sensory-specific cortical areas, Neuroscientist 12 (2006) 327–338.
[18] M.A. Meredith, B.E. Stein, Spatial determinants of multisensory integration in cat superior colliculus neurons, J. Neurophysiol. 75 (1996) 1843–1857.
[19] J.C. Middlebrooks, E.I. Knudsen, A neural code for auditory space in the cat's superior colliculus, J. Neurosci. 4 (1984) 2621–2634.
[20] T. Raij, Onset timing of audiovisual crossmodal activations in primary sensory cortices. Paper presentation, IMRF, 2007.
[21] G.H. Recanzone, Spatial processing in the auditory cortex of the macaque monkey, Proc. Natl. Acad. Sci. 97 (2000) 11829–11835.
[22] B.A. Rowland, S. Quessy, T.R. Stanford, B.E. Stein, Multisensory integration shortens physiological response latencies, J. Neurosci. 27 (2007) 5879–5884.
[23] R. Steenken, H. Colonius, A. Diederich, S. Rach, Visual-auditory interaction in saccadic reaction time: effects of auditory masker level, Brain Res., in press, doi:10.1016/j.brainres.2007.08.034.
[24] B.E. Stein, A.M. Meredith, The Merging of the Senses, MIT Press, Cambridge, 1993.
[25] B.E. Stein, M.T. Wallace, T.R. Stanford, W. Jiang, Cortex governs multisensory integration in the midbrain, Neuroscientist 8 (2002) 306–314.
[26] A.J. Van Opstal, D.P. Munoz, Auditory-visual interactions subserving primate gaze orienting, in: G. Calvert, C. Spence, B.E. Stein (Eds.), Handbook of Multisensory Processes, MIT Press, Cambridge, 2004, pp. 373–393.
[27] M.T. Wallace, M.A. Meredith, B.E. Stein, Multisensory integration in the superior colliculus of the alert cat, J. Neurophysiol. 80 (1998) 1006–1010.

Please cite this article in press as: R. Steenken, et al., Time course of auditory masker effects: Tapping the locus of audiovisual integration?, Neurosci. Lett. (2008), doi:10.1016/j.neulet.2008.02.017

Neuroscience Letters xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Erratum

Erratum to “Time course of auditory masker effects: Tapping the locus of audiovisual integration?” [Neurosci. Lett. 435 (2008) 78–83]

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

[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:
Received 26 March 2008
Accepted 10 April 2008

The authors regret to communicate that, due to a programming error, the duration of the auditory accessory non-target stimulus in the experiment was 20 ms rather than 2 ms, as reported. A corrected depiction of the time course of a trial is presented in Fig. 1 below. The change in the relative timing between the auditory accessory and the masker generates a fourth condition of *partial masking* where the masker follows the onset of the non-target by a certain time delay (stimulus onset asynchrony, SOA) (see Fig. 1). In the original text, the following substitutions should be made: (1) “20 ms” instead of “2 ms” (auditory non-target); (2) SOA “7,12,17” instead of interstimulus intervals (ISI) “5, 10, 15”, and ISIs “2, 7, 12, 22” instead of ISIs “20, 25, 30, 40”; (3) The second paragraph on the last page of text “The most interesting finding . . . as a non-specific warning signal [7].” should be replaced by “The most Interesting finding of this study, besides demonstrating a weakening effect of the masker on multisensory enhancement, is the observation that the amount of this effect remains rather unaffected by delaying the onset of the masker up to 22 ms after presentation of the target and the non-target. This suggests that whatever the masker does to the mechanism of visual–auditory interaction, it will have the same effect whether it occurs with an SOA of 7 ms or 22 ms after the onset of the target/non-target pair. That timing is critical is demonstrated by the ISI = 0 condition, the only condition when masker and auditory non-target completely overlap: the spatial effect of multisensory SRT breaks down entirely and the masker seems to act as a non-specific warning signal [7]. Presumably, the reduced spatial audio–visual effect is determined by that part of the auditory stimulus which is present without the masker. Otherwise, the SRT in the ISI = 0 condition should have been the same as for the other temporal masking conditions. Moreover, it seems that 7 ms of the unmasked non-target are sufficient to generate the reduced spatial audio–visual effect seen here, because no further difference in SRT is discernable, even if the masker follows the end of presentation of the accessory by 2–22 ms.”

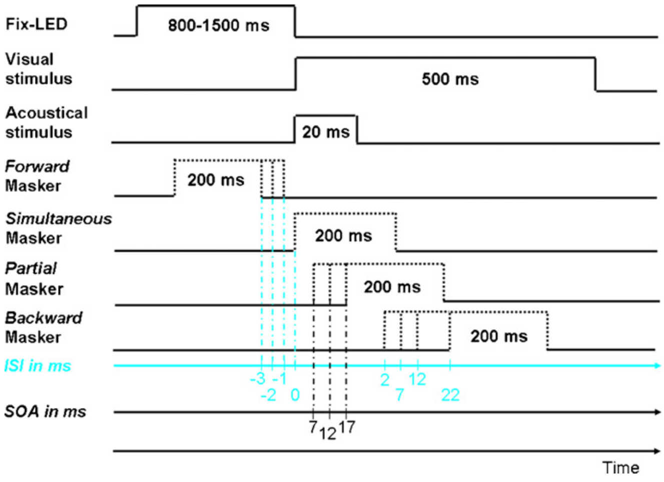

Fig. 1. . Time course of a trial. After fix-LED offset, the visual target was either presented alone, or with masker, or with masker and (coincident or disparate) acoustical non-target. The masker was presented before, simultaneously, or after the presentation of the acoustical non-target. There was no temporal overlap between masker and non-target except for the simultaneous condition.

DOI of original article: 10.1016/j.neulet.2008.02.017.

* Corresponding author. Tel.: +49 441 798 5152; fax: +49 441 798 5172.
E-mail address: rike.steenken@uni-oldenburg.de (R. Steenken).

doi:10.1016/j.neulet.2008.04.028