

The Contributions of Sensory Dominance and Attentional Bias to Cross-modal Enhancement of Visual Cortex Excitability

Vincenzo Romei^{1,2}, Micah M. Murray^{3,4}, Céline Cappe^{5,6},
and Gregor Thut¹

Abstract

■ Approaching or looming sounds (L-sounds) have been shown to selectively increase visual cortex excitability [Romei, V., Murray, M. M., Cappe, C., & Thut, G. Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology*, 19, 1799–1805, 2009]. These cross-modal effects start at an early, preperceptual stage of sound processing and persist with increasing sound duration. Here, we identified individual factors contributing to cross-modal effects on visual cortex excitability and studied the persistence of effects after sound offset. To this end, we probed the impact of different L-sound velocities on phosphene perception postsound as a function of individual auditory versus visual preference/dominance using single-pulse TMS over the occipital pole. We found that the boosting of phosphene perception by L-sounds continued for several tens of milliseconds after the end of the L-sound and was temporally sensitive to different L-sound profiles (velocities).

In addition, we found that this depended on an individual's preferred sensory modality (auditory vs. visual) as determined through a divided attention task (attentional preference), but not on their simple threshold detection level per sensory modality. Whereas individuals with “visual preference” showed enhanced phosphene perception irrespective of L-sound velocity, those with “auditory preference” showed differential peaks in phosphene perception whose delays after sound-offset followed the different L-sound velocity profiles. These novel findings suggest that looming signals modulate visual cortex excitability beyond sound duration possibly to support prompt identification and reaction to potentially dangerous approaching objects. The observed interindividual differences favor the idea that unlike early effects this late L-sound impact on visual cortex excitability is influenced by cross-modal attentional mechanisms rather than low-level sensory processes. ■

INTRODUCTION

Approaching or looming signals are indicative of potential dangers. Mechanisms that preferentially respond to such signals would facilitate adaptive behavior and yield an evolutionary advantage. There are several lines of evidence demonstrating an evolved capacity to detect looming signals when conveyed through one or multiple senses. Already human infants (Ball & Tronick, 1971) and animals (Schiff, 1965; Schiff, Caviness, & Gibson, 1962) produce avoidance responses to looming, but not to receding visual cues. Similarly, both humans (Seifritz et al., 2002) and monkeys (Maier & Ghazanfar, 2007; Ghazanfar, Neuhoff, & Logothetis, 2002) show a perceptual bias for auditory looming stimuli. Human listeners reliably overestimate rising (looming) as compared with equivalent falling (receding) sound intensity (Neuhoff, 1998), and it has been proposed that, in natural conditions, this overestimation could provide a se-

lective advantage by increasing the margin of safety for responses to looming objects. Finally, perceptual and neural advantages have been observed also during the processing of multisensory looming (but not receding) signals both in monkeys (Maier, Chandrasekaran, & Ghazanfar, 2008; Maier, Neuhoff, Logothetis, & Ghazanfar, 2004) and humans (Cappe, Thelen, Romei, Thut, & Murray, 2012; Leo, Romei, Freeman, Ladavas, & Driver, 2011; Cappe, Thut, Romei, & Murray, 2009).

Following these observations and in line with growing evidence showing early cross-modal and multisensory modulation of low-level visual cortices (e.g., Cappe, et al., 2012; Murray, Cappe, Romei, Martuzzi, & Thut, 2012; Romei, Gross, & Thut, 2012; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011; Raij et al., 2010; Wang, Celebrini, Trotter, & Barone, 2008; Kayser, Petkov, Augath, & Logothetis, 2007; Martuzzi et al., 2007; Romei, Murray, Merabet, & Thut, 2007; Molholm et al., 2002; Giard & Peronnet, 1999), we have recently shown that looming sounds dramatically and selectively increase phosphene perception, induced by occipital single-pulse TMS at sound offset, relative to stationary or receding sounds (Romei, Murray, Cappe, & Thut, 2009). This effect was observed

¹University of Glasgow, ²University of Essex, ³Center for Biomedical Imaging (CIBM), Lausanne and Geneva, Switzerland, ⁴Vaudois University Hospital Center and University of Lausanne, ⁵Ecole Polytechnique Fédérale de Lausanne, ⁶Université de Toulouse, UPS, CNRS, Centre de Recherche Cerveau et Cognition

across sounds of 250-, 500-, or 1000-msec duration, and a finer-scale examination of its chronometry revealed that looming effects in visual cortex occurred as early as 80 msec after sound onset at a preperceptual stage of sound processing, which provided evidence that these effects were driven by low-level rather than highly processed sound features (Romei et al., 2009). Effects after sound offset were not tested and may prove more effective in deploying higher-order (e.g., attention) processes. More generally, it is unclear if changes in visual cortex excitability persist at latencies beyond sound offset, if such differ according to sound profile, and whether any such effects are driven by low-level versus higher-order mechanisms. For example, Bestmann, Ruff, Blakemore, Driver, and Thilo (2007) demonstrated a role of attentional bias in modulation of phosphene perception; phosphenes were more readily induced at attended locations. Attentional mechanisms have likewise been put forth as the predominant interpretation of the preferential processing of L-sounds, as they serve as a salient warning cue (Bach et al., 2008; Seifritz et al., 2002; Neuhoff, 1998; see also Kayser, Petkov, Lippert, & Logothetis, 2005, for a discussion of auditory saliency maps) that might in turn facilitate detection in other sensory modalities (Bach et al., 2008; Ghazanfar et al., 2002, but see Cappe et al., 2012; Romei et al., 2009, for a different account of early-latency effects). Moreover, recent studies have shown that attending to auditory stimuli in a multisensory context may affect visual processing, especially with respect to its temporal dynamics (cf. Mishra, Martínez, & Hillyard, 2010; Talsma, Senkowski, & Woldorff 2009).

Here, we systematically tested the dynamics of excitability changes in visual cortex beyond sound offset and whether an attentional mechanism (vs. early low-level sensory processing) drives such effects. We aimed to distinguish between low-level sensory processes and attentional mechanisms as mediators of the *late* L-sound impact on phosphene perception by exploring the contribution of looming velocity on phosphene perception as a function of two individual factors. Looming velocity was chosen on the premise that a more rapidly approaching sound source would be more alerting and would receive more attention than one approaching more slowly or not approaching at all. Our interest in interindividual differences was inspired by the seminal ERP findings of Giard and Peronnet (1999), who showed that multisensory effects on neural activity depended on an individual's dominant sensory modality. We therefore determined per participant (1) the individual threshold per sensory modality using a simple detection task (sensory dominance) as well as (2) the individual's preference for one modality over the other in a task in which sounds and visual events competed for attentional resources (attentional preference); two measures which we found to be dissociated in this study. If effects of L-sounds on phosphene perception were to depend on individual sensory dominance per modality, this would speak in favor of low-level process

as mediators of these cross-modal effects. Conversely, if these were to depend on attentional preferences across modalities, this would favor the implication of a higher-order mechanism.

Specifically, we manipulated the perceived velocity of L-sounds by varying the exponential dB increment per 1 sec of L-sound (set at +15, +25, and +35 dB from start to end,

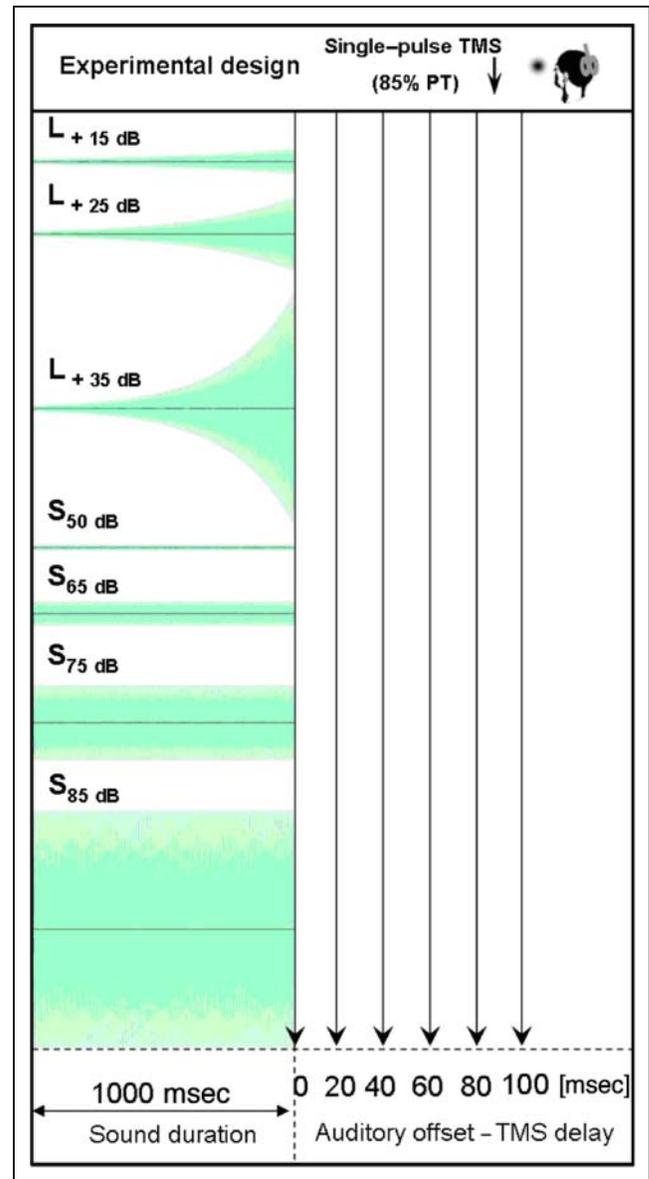


Figure 1. Stimuli and experimental design. For a comparison across different sound categories, looming sounds of three different velocities (from 50 to 65 dB: L + 15 dB; from 50 to 75 dB: L + 25 dB; from 50 to 85 dB: L + 35 dB) were presented together with stationary sounds equated for start (50 dB: S50 dB) and end intensities (65 dB: S65 dB; 75 dB: S75 dB; 85 dB: S85 dB; all 1-sec duration). Visual cortex was stimulated by TMS at sound offset (0 msec) or at several time delays postsound (20–100 msec range in 20-msec steps), to sample its excitability at sound offset and beyond, through quantification of TMS induced illusory visual perceptions (phosphenes). Participants were thus involved in a visual task (phosphene perception) for which the auditory stimulation was irrelevant. TMS was applied below PT (85% PT), which was defined at baseline (i.e., in the absence of any sound).

with start intensities at 50 dB; see Figure 1, top left sound profiles). We controlled for the impact of absolute intensity on phosphene modulation by presenting stationary sounds (S-sounds) of matched start and end amplitude (Figure 1, bottom left sound profiles). We determined the cross-modal impact of differential L-sound velocities by testing phosphene perception at sound offset and at several time delays after sound exposure (0–100 msec in 20-msec steps; Figure 1, right). Depending on whether cross-modal low-level processes or attentional mechanisms are at play, we predicted a differential boosting of phosphene perception as a function of either sensory dominance or attentional preference. To foreshadow our results, we show that phosphene perception is further boosted after L-sound offset depending on sound velocity. Interestingly, no differential effect was found when groups were split according to the simple threshold detection level. Instead and in line with an attentional account, we found modulation of phosphene perception to change with looming velocity in the group with “auditory” but not “visual attentional preferences.”

METHODS

Participants

Fourteen participants with normal hearing by self-report participated in the study (six women, mean age = 27.5 years, range = 20–41 years). The experiment was performed at the Centre for Cognitive Neuroimaging, University of Glasgow. All participants gave written informed consent to the procedures that were approved by the local ethics committee (fIMS).

Occipital TMS, Auditory Stimuli, and Procedure

Illusory visual percepts (phosphenes) were induced by TMS over the occipital pole via a 70-mm figure-of-eight TMS coil connected to a Magstim Rapid2 TMS (Magstim Company, Spring Gardens, UK). This type of protocol has been extensively used to probe visual cortex excitability (e.g., Romei et al., 2007, 2009, 2012; Silvanto, Muggleton, Lavie, & Walsh, 2009; Romei, Brodbeck, et al., 2008; Romei, Rihs, Brodbeck, & Thut, 2008; Bestmann et al., 2007; Bolognini & Maravita, 2007; Ramos-Estebanez et al., 2007; Cowey & Walsh, 2000), based on findings of phosphenes to originate from early visual areas (V1/V2; e.g., Bestmann et al., 2007; Cowey & Walsh, 2000).

All participants underwent one or more training sessions before the experiment. These sessions included careful determination of the site from which the occipital TMS pulse induced a phosphene, at the minimum intensity needed to evoke a phosphene on approximately 50% of trials (phosphene threshold [PT]). This session also served to evaluate the consistency of phosphene perception through repeated stimulation over time. The optimal TMS coil position over the occipital pole as well as the

shape, size, and position of the perceived phosphenes varied somewhat across participants but was constant for each participant across the different TMS blocks. The phosphenes consistently appeared within the lower visual field quadrant opposite to the stimulated occipital cortex, corresponding to the stimulation of the dorsal part of the occipital pole representing the central part of the lower visual field and in accordance with previous reports (e.g., Romei et al., 2007, 2009, 2012; Silvanto et al., 2009; Romei, Brodbeck, et al., 2008; Romei, Rihs, et al., 2008; Ramos-Estebanez et al., 2007; Cowey & Walsh, 2000).

During the experiment, TMS was applied at individual subthreshold intensity of 85% PT always targeting the same optimal stimulation site localized on average ($\pm SEM$) 3.47 ± 0.14 cm above the inion and 0.58 ± 0.33 to the right of the midline. The 85% sub-PT intensity was chosen, because it was previously shown to induce phosphene perception in approximately 50% of trials when single-pulse TMS is paired with brief auditory stimuli (Romei et al., 2007) and thus to be optimized for the prevention of both floor and ceiling effects. With mean PT across participants being 75.71% (± 2.58) of maximum stimulator output, stimulation intensity amounted to 64.65% (± 2.3).

During all TMS blocks, participants were blindfolded. Participants were instructed to make a button press with their right index finger whenever a phosphene was perceived and with the right middle finger whenever no phosphene was perceived. In addition, the blindfold was removed in between TMS blocks to prevent systematic drifts in PT (Pitskel, Merabet, Ramos-Estebanez, Kauffman, & Pascual-Leone, 2007) by adaptation to darkness or drowsiness.

L- and S-sounds consisted of pure tones with 0.4-kHz carrier frequency, composed of triangular waveforms and generated in Cool Edit Pro Software (Syntrillium software Corp, www.syntrillium.com). In our previous study (Romei et al., 2009), we manipulated sound length by using L-sounds of 250-, 500-, and 1000-msec duration while keeping constant their intensity increase (set to 20-dB increase) from start to end. In that case, we did not find any significant modulation of phosphene perception according to the sound duration. Instead, the highest modulation of phosphenes occurred for L-sounds. As modulating sound duration did not lead to any change in phosphene perception, here we kept sound duration constant while instead modulating approaching intensities to levels higher than in our prior study. Specifically, L-sounds of three different ramping intensities and S-sounds of matching start and end intensities of 1-sec duration were presented (Figure 1). L-sounds rose exponentially in intensity: (i) from 50 to 65 dB [L_{+15} dB], (ii) from 50 to 75 dB [L_{+25} dB], and (iii) from 50 to 85 dB [L_{+35} dB] sound pressure level (SPL) at the ear (44.1 kHz sampling rate). Control S-sounds were presented at 50 (S_{50} dB), 65 (S_{65} dB), 75 (S_{75} dB), and 85 dB (S_{85} dB) SPL at ear. Exponential changes were used because they more closely approximate the changes in the intensity that occurs because of constant source velocity in natural environments

than do linear changes (Ghazanfar et al., 2002). The maximum SPL (85 dB) measured at a distance of 75 cm was well below those inducing a startle response (Blumenthal et al., 2005; Blumenthal, 1996; Brown et al., 1991) as previously shown (Romei et al., 2007). Sounds were presented through two loudspeakers at 75 cm distance in front of the participant and were thus perceived centrally.

TMS was applied either at sound offset (i.e., 0 msec) or at five time delays from sound offset (20–100 msec in 20-msec steps; Figure 1). Sound conditions and trials without sound presentation (baseline [BSL], i.e., TMS alone) were randomly intermixed at a rate of 2–4 sec following the participant's manual response.

Each participant completed five blocks of 86 trials, with each block lasting approximately 5 min. A 2-min break was introduced in-between blocks, which led this session to last approximately 30 min. There were a total of 430 trials resulting in 10 repetitions per condition (Sound profile \times TMS delay from sound offset + TMS alone (BSL): $(7 \times 6) + 1$).

Indexing Individual Sensory Dominance and Attentional Preference: Visual and Auditory Stimuli and Procedure

To explore differential effects of L-sound velocities as a function of individual sensory or attentional preferences (auditory vs. visual modalities), we assessed for each participant, through separate psychophysical testing, their perceptual threshold in three tasks. Task 1 was designed to determine auditory detection threshold on a purely auditory task (auditory stimuli only). Task 2 was employed to determine visual detection threshold on a purely visual task (visual stimuli only). Task 3 was designed to identify detection threshold for auditory and visual stimuli when these were presented within the same block, such that participants had to divide their attention between the sensory modalities. In Task 3, no instruction was given as to which modality to attend, to allow for inferences on preferred sensory modality.

Procedure

Brief auditory or visual stimuli, parametrically manipulated in intensity (i.e., auditory sound pressure or visual luminance contrast), were presented either separated by modality (Task 1 or 2) or mixed modality (Task 3) for assessment of individual titration curves per task, after piloting a broader range of visual contrast/auditory intensity on three participants to optimally set the intensity range between ceiling and floor performance. Stimuli were presented sequentially every 2 or 4 sec (step size) for \sim 12-msec duration. Tasks were presented in a block design, and stimulus presentation was randomized over trials with regard to intensity (Tasks 1–3) and stimulus type (sounds vs. visual dots, Task 3). Participants were instructed to make a button press with their right index

finger stressing accuracy (rather than speed) whenever a stimulus of any type was detected. Each participant completed two sensory-specific blocks of 210 trials each and two divided attention blocks of 210 trials each (10 repetitions per intensity level), with each block lasting approximately 10 min. This led to a total of 840 trials. A 5-min break was introduced in between blocks, which led this session to last approximately 1 hr.

Visual Stimuli

Monochromatic small gray dots (created in Paint, Version 6.0, Microsoft, Windows) of 21 different red, green, blue (RGB) gradients (RGB: $x,y,z = 128$ [catch trials], 120, 118, 116..., 80) were presented on a monochromatic gray background (RGB: $x,y,z = 128$). The dots subtended 1.2° in diameter and were positioned at 3° vertical eccentricity below a central fixation cross on a CRT computer monitor (85-Hz refresh rate and 75-cm viewing distance).

Auditory Stimuli

Twenty-one auditory stimuli (of 0.9 kHz carrier frequency) were presented at SPL of 26–36 dB (in 0.5-dB step increments, plus catch trials: 0 dB). Stimuli were composed of triangular waveforms (44.1 kHz sampling rate) generated in Cool Edit Pro Software (Syntrillium software Corp, www.syntrillium.com) and presented through two loud speakers on either side of the computer screen at 75-cm distance from the participant.

Grouping: General Principle

We grouped our participants according to sensory dominance per modality and attentional preference across modalities, to then look for effects of this grouping criterion on a completely independent variable (phosphene report).

To assess the influence of level of individual sensory dominance (across participants) on cross-modal L-sound effects, we assigned participants to either of two groups (median split) according to their threshold level per sensory modality (for threshold evaluation see Grouping Formulas below).

To assess the influence of attentional preference on cross-modal L-sound effects, we assigned participants to either of two other groups according to how much their performance was changing during divided attention (as compared with unimodal blocks) in one modality relative to the other (relative sensory modality bias/threshold shift). To infer individual attentional bias, we first compared per participant their sensory detection threshold during the purely unimodal task (i.e., Task 1 or 2) with their threshold for the same modality during the divided attention task (i.e., Task 3). We found that 6 of the 14 participants worsened in both visual and auditory thresholds during divided attention so that a prompt grouping selection based on absolute threshold shifts in only one or the other

modality was not possible. We therefore compared the changes in the divided attention relative to the respective unimodal condition between modalities (whether threshold shifts in the divided attention tasks were stronger for the auditory or visual modality, see Grouping Formulas below). By calculating this index of relative biases, it is possible to assign each participant to a group with relative “auditory” or “visual attentional preferences.”

Grouping Formulas

The logistic curves obtained from psychophysical evaluation in the unimodal and divided attention conditions were fitted with a sigmoid function $y = 1/(1 + \exp(-1*(x - a)/b))$ using nonlinear regression analysis and least-square estimation implemented in STATISTICA software (StatSoft, Inc., 8.0, www.statsoft.com). In the formula, y represents the average percentage increment of sensory perception as a function of x , the increasing sensory intensity (i.e., auditory sound pressure or visual contrast in luminance). a is the center and b is the width of the curve.

We derived individual indices of sensory or attentional preferences from Tasks 1 to 3 and curve fitting as follows:

To index low-level visual and auditory dominance (across participants), we determined the visual or auditory threshold values per participant (curve centers, in [dB] or [RGB], respectively) from performance in unimodal Tasks 1 and 2 and used the median splits for grouping participants with low versus high visual or auditory dominance. For calculating attentional bias, we entered per modality the curve center values obtained in the unimodal and divided attention task in the following formula: $[(\text{unimodal} - \text{divided attention})_{\text{per modality (aud OR vis)}} / (\text{unimodal} + \text{divided attention})_{\text{per modality (aud OR vis)}}] \times 100$, which returns a percentage value, where negative values would represent a disadvantage in the divided attention relative to the unimodal task, and positive values the opposite trend (advantage), whereas 0% would represent no change. This was calculated for each participant and for each sensory modality (aud or vis). We then compared the changes during divided attention relative to unimodal performance (as calculated above) between modalities by subtracting the %-values obtained for the auditory and the visual modalities (aud minus vis). This measure could take any number (negative and positive) with positive values indicating attentional preference for the auditory modality (auditory modality less affected than visual during divided attention) and negative values indicating attentional preference for the visual modality (visual modality less affected than auditory during divided attention). It is important to note that, although the distribution could have taken any form (all values positive or all negative), positive and negative values (attentional preferences for visual or auditory modality) were equally distributed across the 14 participants (7:7, see below). We therefore assigned to each attentional group an equal number of people that during divided attention showed an advantage in one or in the other modality.

Grouping According to Sensory Dominance versus Attentional Preference: Description of Group Characteristics

With regards to attentional preferences, seven participants showed a stronger threshold shift in the auditory than visual modality in divided attention relative to unimodal tasks (negative values in final attentional preference index) and were therefore assigned to the visual attention group. The remaining seven participants showed a stronger threshold shift in the visual than auditory modality in divided attention relative to unimodal tasks (positive values in final attentional preference index) and were therefore assigned to the auditory attention group.

The outcome of our attentional preference grouping is further illustrated in Figure 2, depicting logistic curves after reaveraging according to attentional preference indices. The figure shows curves for detection of auditory and visual stimuli in unimodal (black lines) and divided attention tasks (red lines), respectively, and illustrates the directionality of performance differences in the divided attention task per modality (relative to unimodal performance). Importantly, in the visual attentional group, visual performance remained unchanged during divided attention relative to the unimodal task (Figure 2, top right, curve centers nearby at 28.53 RGB contrast vs. 29.19 RGB contrast), but auditory performance worsened relative to the unimodal task (Figure 2, top left, clear rightward shift of curve centers from 31.69 dB to 32.75 dB). Conversely, participants of the auditory attention group did not change as to auditory performance (Figure 2, bottom left, curve centers nearby at 31.91 dB vs. 31.86 dB) but got worse in visual performance in divided attention relative to the respective unimodal task (Figure 2, bottom right, clear rightward shift of curve centers from 32.48 RGB contrast to 38.09 RGB contrast). Note that this is not a trivial outcome of our attentional grouping criterion, as the sigmoid curves in divided attention could be located anywhere relative to the unimodal condition. This outcome of grouping suggests that in the divided attention task, participants focus on one modality (presumably their preferred modality) leading to (i) normal (preserved) performance in this modality (i.e., no threshold shifts = sigmoid curves overlying with unimodal performance in the same modality), but (ii) a (rightward) shift toward worse performance in the other modality (sigmoid curves deviate during divided attention relative to unimodal condition). That is, despite the need for dividing attention, performance in half of the participants was on average preserved in one modality (the preferred modality), at the cost of performance in the other modality (reflected in an unimodal threshold shift). By extension, this also suggests that our grouping criterion allowed us to define two groups of equal size with relative attentional preference in one rather than the other sensory modality.

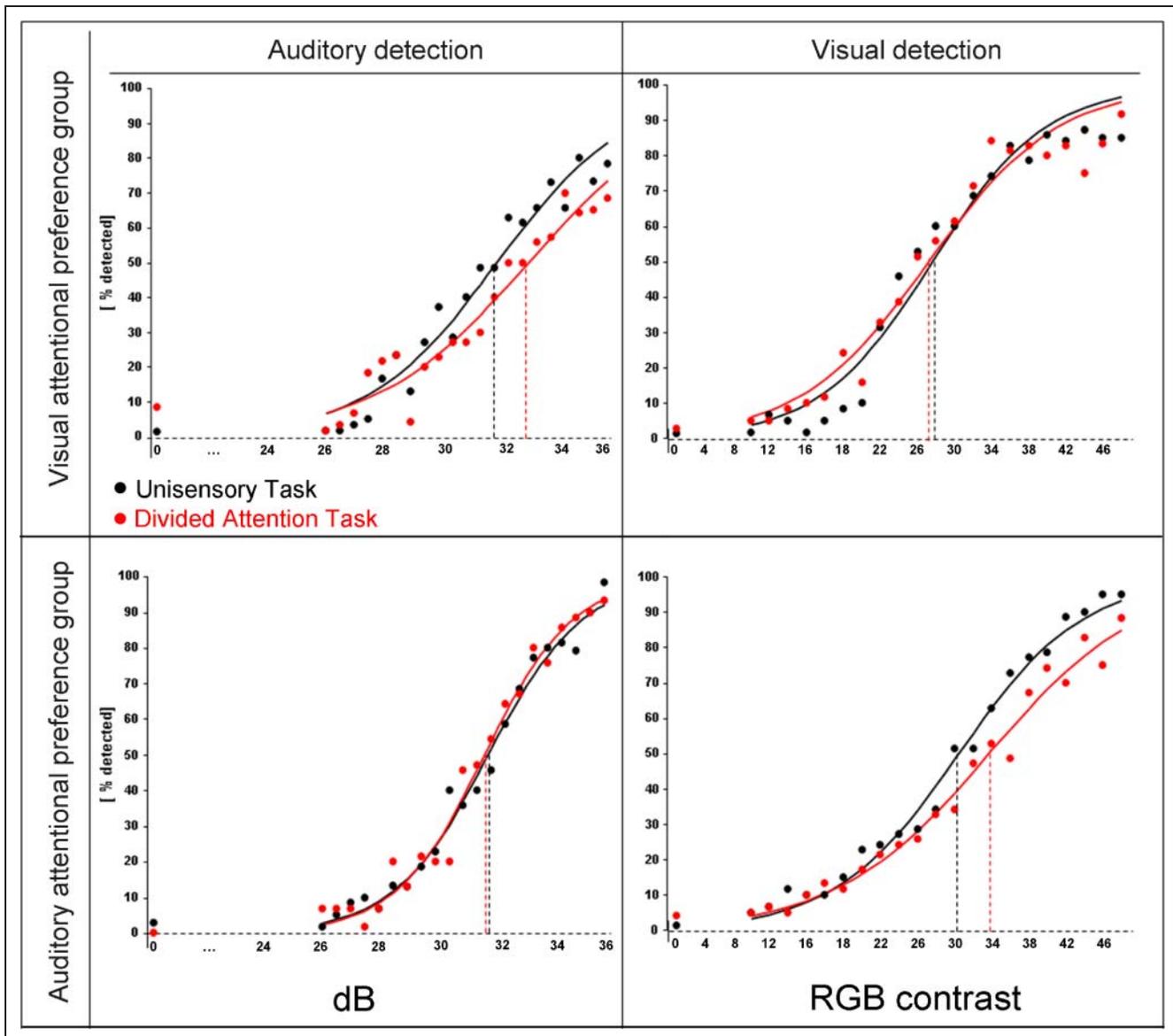


Figure 2. Detection rate (dots) and fitted sigmoid function (curves) for auditory (left) and visual (right) perithreshold stimuli in a unimodal (black curves and black dots) and divided attention task (red curves and red dots) for the group of participants showing visual attentional preference (top) and auditory attentional preference (bottom). Note the rightward shift in the performance curve of the group with visual preference

Grouping According to Sensory Dominance versus Attentional Preference: Independent Measures?

An important question that arises when using the sensory modality thresholds and the attentional preference index to test for sensory versus attentional influences on the L-Sound effects is the extent to which these two measures are dependent (whether participants showing auditory preference also show lower auditory thresholds) or whether they instead constitute distinct measures. In other words, it is important to ascertain the degree to which these measures are decoupled from each other.

To address this question, we first broke down each divided attention group in percentage of participants who showed a visual or auditory bias in unimodal (sensory)

detection tasks. We found that only 57% of participants who showed an auditory preference (worse visual performance) under divided attention were also grouped as auditory (low auditory threshold) in the unimodal auditory task; similarly only 57% of participants who showed a visual preference (lower auditory performance) under divided attention were also grouped as visual (low visual threshold) in the unimodal visual tasks. This suggests that indices of attentional preferences and detection threshold are dissociated.

Second, we contrasted unimodal threshold values (curve centers) between visual and auditory attentional preference groups and found unimodal performance not to depend on divided attention grouping (Auditory vs. Visual preference group: unimodal auditory detection;

31.86 dB vs. 31.69 dB, $p = .71$; unimodal visual detection 32.48 RGB vs. 29.19 RGB, $p = .47$), that is, participants with auditory preference did not show significantly lower unimodal auditory thresholds than participants with visual preferences, and vice versa for unimodal visual thresholds. This provides further evidence that attentional preference and sensory dominance are indeed independent here.

Finally, we correlated the divided attention index with the threshold performance in the simple sensory threshold conditions (auditory and visual) across participants. If the sensory dominance and attention preference measures were coupled, we should find significant correlations between our attention index and the simple threshold measures obtained in our sensory tasks. However, none of these measures significantly correlated (all $r < 0.27$), again showing that sensory dominance and attentional preferences as measured in our study do effectively constitute two independent measures.

In other words, this suggests that any effect of attentional preference on sound-induced visual cortex changes (see below) cannot simply be explained by detection threshold, or vice versa.

Data Analyses: Effects of Grouping on Phosphene Report

To assess cross-modal L-sound effects as a function of TMS delay and sound profile, we first subjected the percentages of phosphene perception (across trials) to repeated-measure (rm) ANOVAs. As within-subject factors, we used temporal sound profile (L_{+15} dB, L_{+25} dB, L_{+35} dB, S_{50} dB, S_{65} dB, S_{75} dB and S_{85} dB) and TMS delay (0–100 msec; 7×6 rmANOVA); or sound category (Looming vs. Stationary), Temporal Sound Profile (L_{+15} dB vs. S_{65} dB, L_{+25} dB vs. S_{75} dB, and L_{+35} dB vs. S_{85} dB) and TMS Delay (0–100 msec; $2 \times 3 \times 6$ rmANOVA). Planned paired t tests were performed to examine differences between conditions. The same rmANOVA was performed on RTs of perceived phosphenes to compare across sound ramping intensity and sound profile.

To test our main hypothesis (low-level processing vs. cross-modal attention account) data were re-analyzed in 3 different rmANOVAs including the factor group (either visual median split, auditory median split or divided attention scores) to test for differential influences of sounds on visual cortex excitability according to individual differences. Planned paired t tests were performed to examine differences between and within groups across conditions.

RESULTS

Effects of Temporal Sound Profiles on Phosphene Perception

Results of the 7×6 ANOVA with the within-subject factors Temporal Sound Profile of different L-sound velocities/matched S-sounds amplitudes (L_{+15} dB, L_{+25} dB, L_{+35} dB,

S_{50} dB, S_{65} dB, S_{75} dB, and S_{85} dB) \times TMS delay (0, 20, 40, 60, 80, 100 msec from sound offset) revealed differential modulations of phosphene perception by different sound profiles, $F(6, 78) = 11.75$, $p < .00001$, at different time delays, $F(5, 65) = 6.04$, $p = .0001$, and a significant interaction between these factors, $F(30, 390) = 1.53$, $p < .04$.

A $2 \times 3 \times 6$ ANOVA was then performed excluding the condition S_{50} dB to allow for direct comparisons between L- and S-sounds of matched end-intensities, with the within-subject factors Sound Category (looming vs. stationary) \times End Intensity (65, 75, 85 dB) \times TMS Delay (0 to 100 msec). We found that modulation of phosphene perception depended on Sound Category, $F(1, 13) = 26.06$, $p = .0002$, Intensity, $F(2, 26) = 11.38$, $p = .0003$, Delay, $F(5, 65) = 4.87$, $p = .0008$, and showed two-way interactions between Sound Category And Intensity, $F(2, 26) = 3.81$, $p = .035$, and between Sound Category and Delay, $F(5, 65) = 2.57$, $p = .034$.

To break down the interactions with the factor Sound category, we looked at the effects of intensity and delay for L- and S-sounds separately. Whereas L-sounds modulated phosphene perception depending on Intensity, $F(2, 26) = 12.25$, $p = .0002$, and Delay, $F(5, 65) = 7.79$, $p < .00001$, with a significant interaction between these two factors, $F(10, 130) = 2.26$, $p = .018$, S-sounds did not produce any significant modulation of phosphene perception either for intensity, $F(2, 26) = 2.41$, $p = .11$, or delay, $F(5, 65) = 1.21$, $p = .31$, nor an interaction between these two factors, $F(10, 130) = 0.71$, $p = .71$. This reveals that phosphene perception is selectively modulated by L-sounds, and that this modulation depends on the delay from sound-offset and on the velocity of the sound presented (see Figure 3).

Each TMS delay of L-sounds was then contrasted against the 0-msec condition (i.e., at sound offset) to test for modulations of phosphene perception across time after sound offset (see Figure 3). This planned comparison showed that L + 25 dB further enhanced phosphene perception at 40 msec after sound-offset ($t(13) = -2.77$, $p = .016$) [with a return toward baseline levels at 100 msec (47.86%, $t(13) = 3.81$, $p = .002$; see Figure 3, middle panel)] and that L + 35 has an effect at 20 msec after sound-offset ($t(13) = -4.61$, $p = .0005$) [again returning toward baseline at 100 msec (52.5%, $t(13) = 2.44$, $p = .03$; see Figure 3, right panel)]. No further enhancement was observed for L + 15 dB (see Figure 3, left panel).

Differential Effects of Temporal Sound Profile as a Function of Attentional Preference but Not Simple Detection Threshold

Analyses were repeated, first independently for the attentional preference grouping and then independently for the two sensory dominance groupings (for description of these groups, see Methods and Figure 2). The split by attentional preference showed a significant three-way interaction between the factors Attentional Preference,

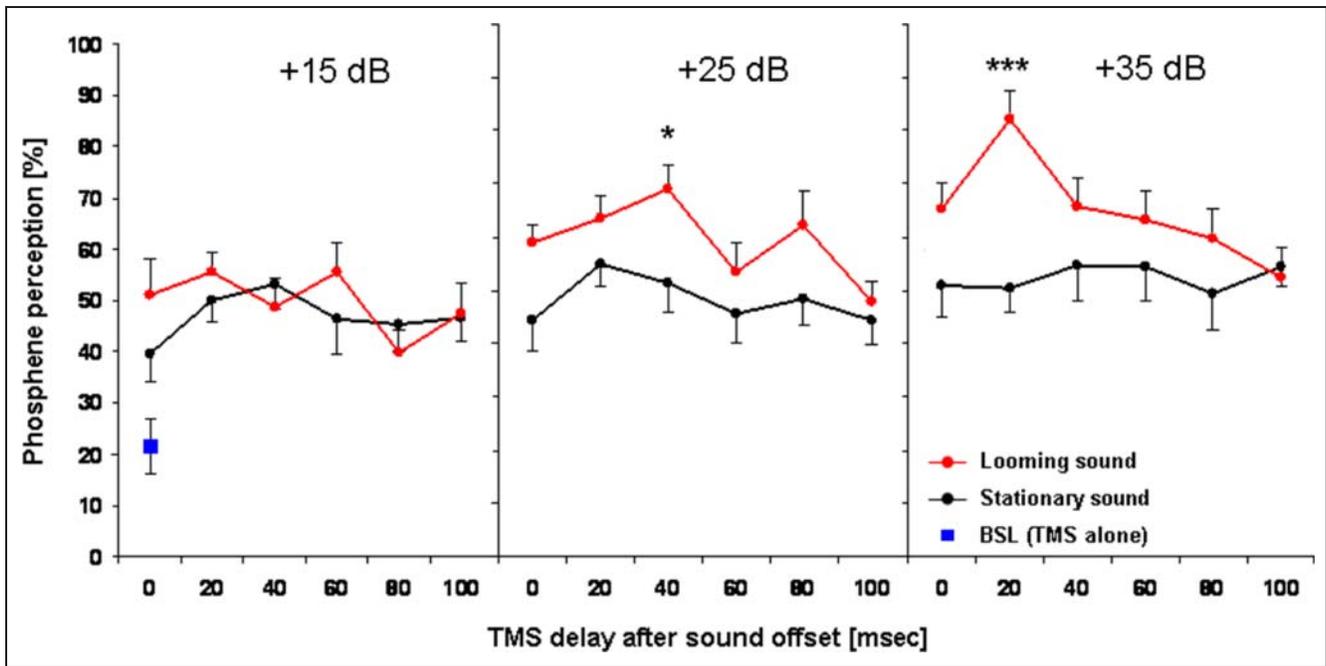


Figure 3. Modulation of phosphene perception by looming sounds of different rising intensities and stationary sounds of matched end intensities. Asterisks point to significant differences in comparison with performance at sound offset of the same sound condition. * $p < .05$; *** $p < .001$.

Sound Category, and Delay, $F(5, 60) = 4.49, p < .001$, a significant three-way interaction between Attentional Preference, Intensity, and Delay, $F(10, 120) = 2.24, p < .02$, and a four-way interaction between the factors Attentional Preference, Sound Category, Intensity, and Delay, $F(10, 120) = 3.23, p < .001$.

To break down the interactions with the factor sound category, we again looked at effects for S- and L-sounds separately. The ANOVA performed on the S-sounds yielded no significant main effect, nor interactions between factors (all $ps > .12$). The same ANOVA performed for L-sounds instead showed main effects of Intensity, $F(2, 24) = 11.95, p < .001$, and Delay, $F(5, 60) = 8.65, p < .00001$, a two-way interaction between Delay and Intensity, $F(10, 120) = 2.8, p < .01$, a two-way interaction between Attentional Preference and Delay ($F(5, 60) = 2.44, p < .05$), and crucially, a three-way interaction Attentional Preference \times Delay \times Intensity, $F(10, 120) = 4.14, p < .0001$.

Post hoc t tests directly contrasting phosphene reports to L-sounds between attentional preference groups (per delay and intensity) showed that in comparison with the visual attention group, the auditory attention group had a tendency for higher phosphene modulations for L-sounds of +35 dB at 20 msec ($p = .083$), for L-sounds of +25 dB at 80 msec ($p < .005$), and for L-sounds of +15 dB at 100 msec ($p < .03$). This temporal profile of phosphene report by L-sound velocity was further corroborated by within-condition comparisons (L-sound) of phosphene report against sound offset (0 msec). In the auditory preference group (Figure 4, top left), phosphene perception was differentially modulated by L_{+35} dB and L_{+25} dB, whereas L_{+15} dB had no effect. L_{+35} dB enhanced phos-

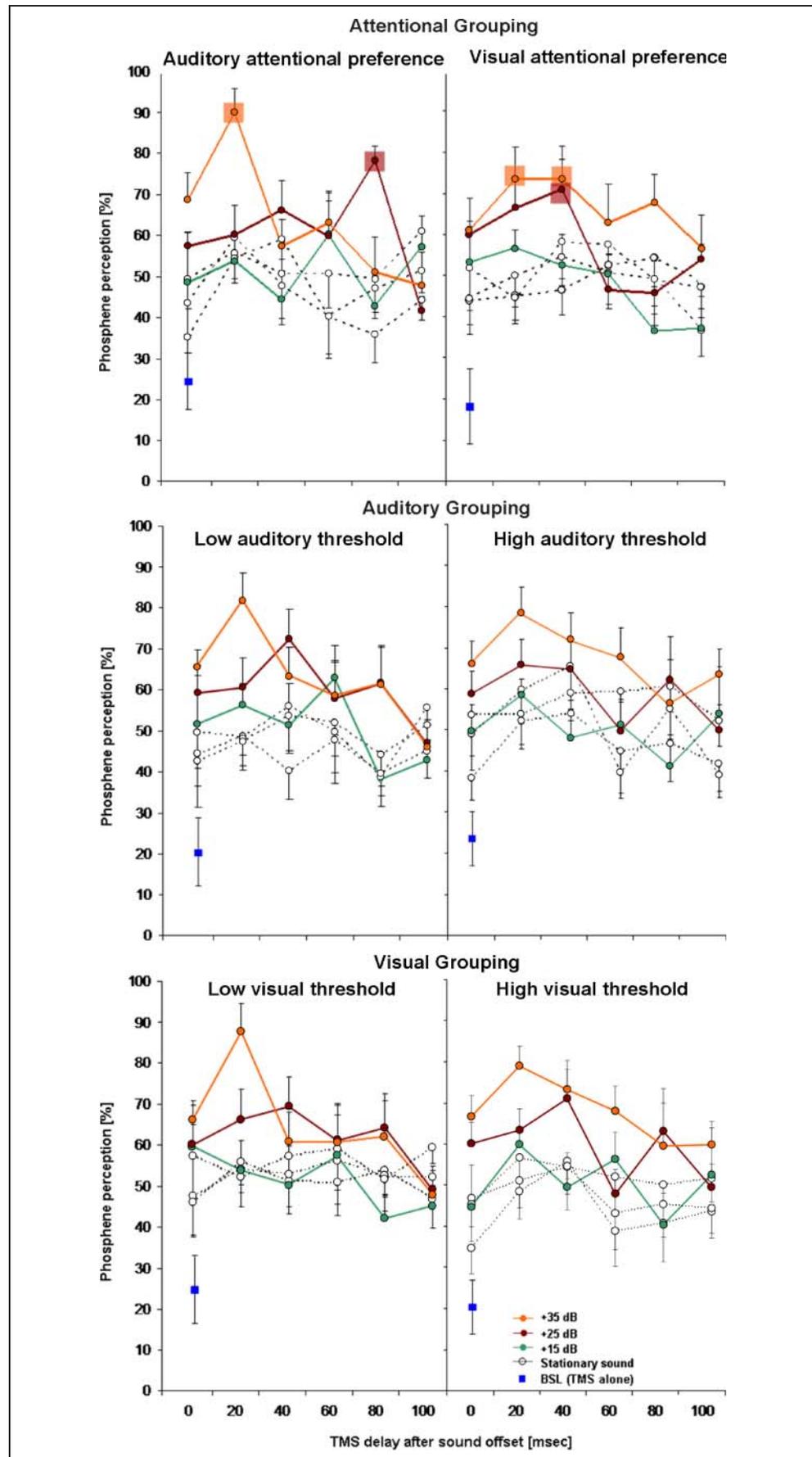
phene perception at 20 msec (90.71%, $t(6) = -3.15, p = .019$) relative to 0 msec (69.29%), whereas L_{+25} dB enhanced perception at 80 msec (78.93%, $t(6) = -7.6, p = .0003$) relative to 0 msec (57.86%). That is, examination of phosphene report to L_{+15} dB, L_{+25} dB, and L_{+35} dB sounds in the auditory preference group per time delay suggests that in this group, phosphene perception is differently modulated by L-sounds as a function of looming velocity, such that faster L-sounds induced maximal phosphene enhancement earlier than slower L-sounds, more following the extrapolated physical property of the sound.

This was not the case in the visual preference group. In this group (Figure 4, top right) phosphene perception was similarly modulated by L_{+35} dB and L_{+25} dB with no significant effects of L_{+15} dB. L_{+35} dB enhanced phosphene perception at 20 msec (73.57%, $t(6) = -2.89, p = .028$) and 40 msec (73.57%, $t(6) = -4.11, p = .006$) relative to 0 msec (61.07%), as did L_{+25} dB (40 msec vs. 0 msec delay: 71.07% vs. 60.0%, $t(6) = -2.12, p = .078$). That is, in the visual preference group, increases of phosphene perception seemed to last longer for L-sounds of higher velocity.

In contrast to attentional preference, the split by sensory dominance according to the visual or auditory thresholds in the unimodal tasks did not reveal any significant difference in the phosphene perception curve profile across groups. Neither the split by visual threshold nor the split by auditory thresholds revealed any significant interaction with the factor group (all $ps > .13$).

Overall, this suggests that high-level attentional mechanisms rather than low-level processes are responsible for our findings. The results indicate that participants with an auditory attentional preference show a temporal modulation

Figure 4. Modulation of phosphene perception by looming sounds of different rising intensities (and stationary sounds of matched end intensities) as a function of attentional preference grouping (top left: auditory attentional preference; top right: visual attentional preference), auditory dominance grouping (middle left: low threshold; middle right: high threshold), and visual dominance grouping (bottom left: low threshold; bottom right: high threshold). Colored squares point to significant enhancements of phosphene perception relative to 0 msec within the same L-sound velocity. All of the dotted lines refer to stationary sounds of different intensities.



of phosphene perception that follows the L-sound velocity profile. Instead, the L-sound velocity profile does not differentially affect the temporal profile of visual cortex excitability in participants with an attentional preference for visual stimuli.

Modulation of Phosphene Perception by L-sounds of Different Velocities: Perceptual or Response Bias?

It could be argued that the present data can be alternatively explained by a response bias, given that the faster L-sounds can be easily perceived as more salient, thus inducing some intrinsic criterion shift. However, several aspects of our data argue against our results being explained by a response bias, rather than a visual perceptual bias of the participants to L-sounds of different velocities. Firstly, we found that RTs were not becoming significantly faster (biased) neither across L-sounds of increasing velocities nor between L- as compared with S-sounds (with this latter evidence replicating what was already noted in Romei et al., 2009): The rmANOVA on RT showed no effect for Sound Category or Delay nor interaction between these factors (all p s > .85). Likewise, when including the factor Intensity in the ANOVA (by excluding the S-sound of matching start intensity), there was no significant effects of Sound Category, that is, looming versus stationary, $F(1, 13) = 0.88, p = .36$, intensity of stimulation, $F(2, 26) = 0.13, p = .88$, and delay, $F(5, 65) = 0.34, p = .96$, nor interactions between these factors (all p s > .24). Second, RTs to L-sounds of different velocities were too slow ($L_{+15 \text{ dB}} = 572 \pm 75$ msec vs. $S_{65 \text{ dB}} = 575 \pm 71$ msec, $L_{+25 \text{ dB}} = 578 \pm 77$ msec vs. $S_{75 \text{ dB}} = 577 \pm 78$ msec, $L_{+35 \text{ dB}} = 590 \pm 79$ msec vs. $S_{85 \text{ dB}} = 571 \pm 73$ msec) to be attributable to a response pattern that might have been triggered by an automatic avoidance mechanism to L-sounds, an enhancement in arousal or other mechanisms that might have differentially affected L-sounds as a function of increasing velocities. Third, if a response bias would be responsible for our differential modulation of phosphene perception depending on the differentially rising sound intensities, then a differential effect between L-sounds of different velocities and stationary sounds of different amplitudes should be observed already at sound offset (0 msec). Contrary to this account, an ANOVA performed on phosphene perception at this time delay revealed a main effect of Sound Category, $F(1, 13) = 39.86, p = .00003$, and an effect of Sound Intensity, $F(2, 26) = 4.27, p = .025$, but no interaction between these two factors, $F(2, 26) = 0.13, p = .87$. Hence, phosphene report did not yet show specific differences between the three L-sound velocity profiles at sound offset (only general differences between intensities, that is, independently of sound category). This pattern thus clearly excludes that response biases can explain our results. As a final remark, previous reports of our and other groups (e.g., Bolognini, Senna, Maravita, Pascual-Leone & Merabet, 2010; Romei et al., 2007, 2009)

independently revealed that the sound-induced increase in likelihood of perceiving a phosphene is not an unspecific effect but shows visual characteristics (not easily reconcilable with a response bias). More specifically, in these previous reports using brief sounds (instead of L- or S-sounds), effects of sound on phosphene report (i) were confined to specific time windows postsound corresponding to response latencies of the visual cortex (Bolognini et al., 2010; Romei et al., 2007) and (ii) have been linked to electrophysiologic activity generated in visual areas (Romei et al., 2012). Overall, this clearly speaks against the argument that response bias might be responsible for variability in phosphene (“yes”) report.

DISCUSSION

In the present work, we systematically tested the time course of the cross-modal impact of looming sounds on phosphene perception as a function of their approaching velocity. We characterized the effect induced by looming and stationary sounds at sound offset (as in Romei et al., 2009) and at later stages of sound processing. We previously showed (Romei et al., 2009) that looming sounds impact phosphene perception very early in time after sound onset (~80 msec into sound presentation), relative to other sound types, with this effect lasting throughout the auditory stimulus duration. Although we could previously link the early effect to low-level processing of sensory information—because it empirically proved to be preperceptual—we did not study the origin of any late effect. Here, we tested these late effects of looming sounds for three different velocities. We found a differential time course of the late cross-modal impact on phosphene perception depending on type of sound, sound velocity, and grouping of participants according to attentional preference.

Late Cross-modal Impact of L-sounds on Visual Cortex Excitability Shows Different Time Courses Depending on Sound Velocity

Looming sounds produced the highest modulation of phosphene perception as compared with stationary sounds of matched end intensity (reproducing Romei et al., 2009). In addition, looming sounds approaching with different velocity had different impact on phosphene perception. Specifically, the fastest approaching sounds ($L_{+35 \text{ dB}}$ sounds) produced faster and higher visual cortex excitability enhancements, whereas slower approaching sounds produced slower and lower ($L_{+25 \text{ dB}}$ sounds) or negligible ($L_{+15 \text{ dB}}$ sound) visual cortex excitability enhancements. This differential delay in peak of phosphene perception across sounds, despite simultaneous sound offset, suggests that the late cross-modal effects of L-sounds on phosphene perception did not closely follow the low-level properties of the sounds. This

is in keeping with the hypothesis that attentional processing of L-sound input (instead of a low-level audio–visual interaction at a sensory level) is enhancing visual cortex excitability at these late delays. Note that we could here discount response bias as an explanation of these results (see Results section above). One could argue that sounds moving at different velocities do differ with regard to their low-level properties (e.g., different mean intensities and ramps) and that these low-level differences have sustained effects on visual cortex excitability that persist beyond sound offset without the need for attentional processes. However, to prevent this confound, we matched different L-sound intensities with those of stationary (control) sounds (S-sounds), that is, we control for the effect of low-level properties. Unlike L-sounds, S-sounds did not yield to significant differences on phosphene reports across conditions, delays, etc. Moreover, not only is the intensity difference between looming sounds unlikely to explain our results but also ramps do not seem to have an important influence. In Romei et al. (2009), we tested for different mean ramps (see Experiment 1) with L-sounds of different durations (250, 500, and 1000 msec) but did not find any significant difference across conditions at sound offset that might account for the attentional effect we have instead demonstrated. In keeping with our previous report, we here did not find significant differences with L-sounds of different velocities at sound offset (0 msec). We found such a difference only after sound offset, presumably dissociating low-level versus attentional effects on phosphene perception.

The attentional account was confirmed by directly testing the two accounts (low-level sensory vs. attentional) through splitting our participants in two groups according to either their sensory threshold in simple auditory or visual detection tasks or their attentional preference for auditory versus visual processing under a divided attention condition. We then looked for effects of these grouping criteria on the independent variable of phosphene report and only found effects of attentional but not sensory groupings discussed below.

Late Cross-modal Impact of L-sounds on Visual Cortex Excitability Depends on Individual Attentional Preference but Not Sensory Thresholds

When dividing our participants in two groups according either to their visual or auditory sensitivity, we found no significant effect of sensory thresholds on cross-modal impact. This further substantiates that low-level sensory processing cannot account for the present results. In contrast, attention preference clearly modulated cross-modal impact. The group showing attentional visual preference under divided attention conditions demonstrated early peaks of phosphene modulation independently of looming sound velocity. In contrast, in the auditory attentional preference group, fastest looming sounds lead to earliest

peaks. That is, visual cortex excitability followed a sound velocity-independent time course for participants showing visual preference over audition or instead followed a more fine-grained, sound velocity-dependent time course for participants showing an auditory preference over vision.

Our findings support the view that not only bottom–up cross-modal interactions can capture attention but also top–down attention can facilitate the integration of multisensory inputs (e.g., Santangelo, Olivetti Belardinelli, Spence, & Macaluso, 2009; Talsma, Doty, & Woldorff, 2007; Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Talsma & Woldorff, 2005), which can lead in turn to spreading of attention across modalities (for a review, see Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010). In this respect, our findings further support the notion that preferential attention to auditory stimuli in a multisensory context may affect visual processing, especially with respect to its temporal dynamics (cf. Mishra et al., 2010; Talsma et al., 2009).

Moreover, our findings extend the seminal results of Giard and Peronnet (1999), showing that multisensory effects on neural activity may depend on individual sensory preferences, and demonstrate that taking into account such factors may be useful in the study of the origin of cross-modal effects. We here confirm that low-level sensory processes cannot account for late cross-modal impact of L-sounds, whereas attention preference modulates these effects. Future studies could check if the reverse pattern is true for early effects induced by L-sounds, that is, low-level sensory processing instead of attentional preference accounting for differential effects in phosphene perception, as suggested by Romei et al. (2009). In this previous study, we found a correlation ($r = .93$) between the latency of early enhancement of phosphene perception and an auditory threshold measure (stationary-looming sound discrimination) but did not test for attentional mechanisms. Another question that remains open is what structures can account for the differential cross-modal effects according to preferred sensory modality or sensory dominance. This could be studied by systematically relating individual differences in multisensory processing to individual differences in the underlying anatomy revealed by structural MR scans (see, e.g., de Haas, Kanai, Jalkanen, & Rees, 2012; Kanai & Rees, 2011).

Late Cross-modal Impact of L-sounds on Visual Cortex Excitability: An Instance of Coding for Imminent Visual Events Based on (Nonvisual) Priors?

Our findings favor the idea that cross-modal attentional mechanisms enhance visual cortex excitability (in the absence of visual input), possibly to extrapolate object trajectory and to allocate sensory processing resources to potentially approaching object. This is reminiscent of

a recent fMRI study showing that apparent visual motion leads to extrapolation effects in retinotopic visual cortex, which cannot be explained by sensory responses (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010), in line with models positing that the brain is not merely reactive but also “proactive” or “predictive” (e.g., Enns & Lleras, 2008; Bar, 2007; Rao & Ballard, 1999). Alink et al. (2010) showed that the predictability of apparent motion stimuli affects fMRI responses in primary visual cortex: Visual targets embedded in the illusory motion path evoked smaller visual responses when their onset or motion direction could be predicted from the illusory motion dynamics (Alink et al., 2010). In addition, these smaller responses co-occurred with higher detection rates, thus providing evidence that the human brain extrapolates imminent sensory input from predictable motion paths allowing predictable visual stimuli to be processed with less neural activation (lower thresholds) at early stages of cortical processing in visual cortex (Alink et al., 2010). Such extrapolation effects even seem to generalize to a nonmotion context, as nonstimulated early visual areas carry information about spatial surrounds in complex visual scenes (Smith & Muckli, 2010). In further analogy to our results, temporal expectation modulates activity in the occipital cortex, not only when the expected stimulus is presented in the same visual modality (Bueti, Bahrami, Walsh, & Rees, 2010) but also cross-modally, during the expectation of auditory stimuli (Bueti & Macaluso, 2010). Our findings therefore seem to corroborate the view that coding of expected imminent visual events by priors involves changes in early visual areas. Here, the priors are nonvisual in nature and consist of motion information conveyed through the auditory modality that induces coherent cross-modal effects in early sensory areas of another (visual) modality. This likely helps ensuing extrapolation/interpretation of the visual world from partial information to predict timing and position of possible future events (such as the timing of an approaching colliding object). Note that our finding of a top-down rather than a bottom-up source of the looming effects (driven by attentional preferences rather than sensory processes) further corroborates this predictive coding account. We conclude that, in analogy to Alink et al. (2010), our results suggest that information about motion (presented here in the auditory domain) is extrapolated to the visual system even in the absence of visual signals to generate estimations of expected sensory input based on information from the recent past.

Conclusions

This study provides novel evidence on the processes through which looming sounds selectively affect visual cortex excitability. Our study builds upon our previous findings of early preperceptual cross-modal effects of looming sounds on visual cortex excitability (Romei et al., 2009) but goes beyond them by addressing an

open question: The extent to which low-level sensory versus higher-order attentional processes play a role in the cross-modal impact of these sounds at later stages of sound processing. Here we provide an answer by adopting an interindividual differences approach. We show that grouping participants according to theoretical predictors (attention preferences, auditory and visual thresholds) helps to successfully disentangle the sensory versus attentional contribution to cross-modal effects on visual cortex excitability. We found that, unlike early modulations (Romei et al., 2009), late cross-modal effects are likely brought about by attentional mechanisms. We conclude that looming sounds engage visual cortex excitability through distinct mechanisms at different stages of sound processing.

Acknowledgments

M. M. M. receives support from the Swiss National Science Foundation (Grant 310030B-133136).

Reprint requests should be sent to Vincenzo Romei, Centre for Brain Science, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom, or via e-mail: vromei@essex.ac.uk.

REFERENCES

- Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, *30*, 2960–2966.
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*, 839–843.
- Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., et al. (2008). Rising sound intensity: An intrinsic warning cue activating the amygdale. *Cerebral Cortex*, *18*, 145–150.
- Ball, W., & Tronick, E. (1971). Infant responses to impending collision: Optical and real. *Science*, *171*, 818–820.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Science*, *11*, 280–289.
- Bestmann, S., Ruff, C. C., Blakemore, C., Driver, J., & Thilo, K. V. (2007). Spatial attention changes excitability of human visual cortex to direct stimulation. *Current Biology*, *17*, 134–139.
- Blumenthal, T. D. (1996). Inhibition of the human startle response is affected by both prepulse intensity and eliciting stimulus intensity. *Biological Psychology*, *44*, 85–104.
- Blumenthal, T. D., Cuthbert, B. N., Fillion, D. L., Hackley, S., Lipp, O. V., & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*, 1–15.
- Bolognini, N., & Maravita, A. (2007). Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Current Biology*, *17*, 1890–1895.
- Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., & Merabet, L. B. (2010). Auditory enhancement of visual phosphene perception: The effect of temporal and spatial factors and of stimulus intensity. *Neuroscience Letters*, *477*, 109–114.
- Brown, P., Rothwell, J. C., Thompson, P. D., Britton, T. C., Day, B. L., & Marsden, C. D. (1991). New observations on the normal auditory startle reflex in man. *Brain*, *114*, 1891–1902.

- Buetti, D., Bahrami, B., Walsh, V., & Rees, G. (2010). Encoding of temporal probabilities in the human brain. *Journal of Neuroscience*, *30*, 4343–4352.
- Buetti, D., & Macaluso, E. (2010). Auditory temporal expectations modulate activity in visual cortex. *Neuroimage*, *51*, 1168–1183.
- Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences, U.S.A.*, *102*, 18751–18756.
- Cappe, C., Thelen, A., Romei, V., Thut, G., & Murray, M. M. (2012). Looming signals reveal synergistic principles of multisensory integration. *Journal of Neuroscience*, *32*, 1171–1182.
- Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2009). Selective integration of auditory-visual looming cues by humans. *Neuropsychologia*, *47*, 1045–1052.
- Cowey, A., & Walsh, V. (2000). Magnetically induced phosphenes in sighted, blind and blindsighted observers. *NeuroReport*, *11*, 3269–3273.
- de Haas, B., Kanai, R., Jalkanen, L., & Rees, G. (2012). Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proceedings of the Royal Society: Biological Sciences*, *279*, 4955–4961.
- Enns, J. T., & Lleras, A. (2008). What's next? New evidence for prediction in human vision. *Trends in Cognitive Science*, *12*, 327–333.
- Ghazanfar, A. A., Neuhoff, J. G., & Logothetis, N. K. (2002). Auditory looming perception in rhesus monkeys. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 15755–15757.
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473–490.
- Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Review Neuroscience*, *12*, 231–242.
- Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *Journal of Neuroscience*, *27*, 1824–1835.
- Kayser, C., Petkov, C. I., Lippert, M., & Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: An auditory saliency map. *Current Biology*, *15*, 1943–1947.
- Leo, F., Romei, V., Freeman, E., Ladavas, E., & Driver, J. (2011). Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Experimental Brain Research*, *213*, 193–201.
- Maier, J. X., Chandrasekaran, C., & Ghazanfar, A. A. (2008). Integration of bimodal looming signals through neuronal coherence in the temporal lobe. *Current Biology*, *18*, 963–968.
- Maier, J. X., & Ghazanfar, A. A. (2007). Looming biases in monkey auditory cortex. *Journal of Neuroscience*, *27*, 4093–4100.
- Maier, J. X., Neuhoff, J. G., Logothetis, N. K., & Ghazanfar, A. A. (2004). Multisensory integration of looming signals by rhesus monkeys. *Neuron*, *43*, 177–1781.
- Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., et al. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral Cortex*, *17*, 1672–1679.
- Mishra, J., Martínez, A., & Hillyard, S. A. (2010). Effect of attention on early cortical processes associated with the sound-induced extra flash illusion. *Journal of Cognitive Neuroscience*, *22*, 1714–1729.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Brain Research, Cognitive Brain Research*, *14*, 115–128.
- Murray, M. M., Cappe, C., Romei, V., Martuzzi, R., & Thut, G. (2012). Auditory-visual multisensory interactions in humans: A synthesis of findings from behavior, ERPs, fMRI, and TMS. In B. E. Stein (Ed.), *The new handbook of multisensory processes*. Cambridge, MA: MIT Press.
- Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*, 123–124.
- Pitskel, N. B., Merabet, L. B., Ramos-Estebanez, C., Kauffman, T., & Pascual-Leone, A. (2007). Time-dependent changes in cortical excitability after prolonged visual deprivation. *NeuroReport*, *18*, 1703–1707.
- Raij, T., Ahveninen, J., Lin, F. H., Witzel, T., Jääskeläinen, I. P., Letham, B., et al. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, *31*, 1772–1782.
- Ramos-Estebanez, C., Merabet, L. B., Machii, K., Fregni, F., Thut, G., Wagner, T. A., et al. (2007). Visual phosphene perception modulated by subthreshold cross-modal sensory stimulation. *Journal of Neuroscience*, *27*, 4178–4181.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
- Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior {alpha}-band EEG activity reflect variability in excitability of human visual areas. *Cerebral Cortex*, *18*, 2010–2018.
- Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, *22*, 807–813.
- Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology*, *19*, 1799–1805.
- Romei, V., Murray, M. M., Merabet, L. B., & Thut, G. (2007). Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: Implications for multisensory interactions. *Journal of Neuroscience*, *27*, 11465–11472.
- Romei, V., Rihs, T., Brodbeck, V., & Thut, G. (2008). Resting electroencephalogram alpha-power over posterior sites indexes baseline visual cortex excitability. *NeuroReport*, *19*, 203–208.
- Santangelo, V., Olivetti Belardinelli, M., Spence, C., & Macaluso, E. (2009). Interactions between voluntary and stimulus-driven spatial attention mechanisms across sensory modalities. *Journal of Cognitive Neuroscience*, *21*, 2384–2397.
- Schiff, W. (1965). Perception of impending collision: A study of visually directed avoidant behavior. *Psychological Monographs*, *79*, 1–26.
- Schiff, W., Caviness, J. A., & Gibson, J. J. (1962). Persistent fear responses in rhesus monkeys to the optical stimulus of "looming." *Science*, *136*, 982–983.
- Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., et al. (2002). Neural processing of auditory looming in the human brain. *Current Biology*, *12*, 2147–2151.
- Silvanto, J., Muggleton, N., Lavie, N., & Walsh, V. (2009). The perceptual and functional consequences of parietal top-down modulation on the visual cortex. *Cerebral Cortex*, *19*, 327–330.

- Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 20099–20103.
- Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? *Cerebral Cortex*, *17*, 679–690.
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cogn Sciences*, *14*, 400–410.
- Talsma, D., Senkowski, D., & Woldorff, M. G. (2009). Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. *Experimental Brain Research*, *198*, 313–328.
- Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, *17*, 1098–1114.
- Van der Burg, E., Talsma, D., Olivers, C. N., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage*, *55*, 1208–1218.
- Wang, Y., Celebrini, S., Trotter, Y., & Barone, P. (2008). Visuo-auditory interactions in the primary visual cortex of the behaving monkey: Electrophysiological evidence. *BMC Neuroscience*, *9*, 79.