Acoustic Noise Improves Visual Perception and Modulates Occipital Oscillatory States

Stephanie Gleiss¹ and Christoph Kayser¹,²

Abstract

Perception is a multisensory process, and previous work has shown that multisensory interactions occur not only for object-related stimuli but also for simplistic and apparently unrelated inputs to the different senses. We here compare the facilitation of visual perception induced by transient (target-synchronized) sounds to the facilitation provided by continuous background noise like sounds. Specifically, we show that continuous acoustic noise improves visual contrast detection by systematically shifting psychometric curves in an amplitude-dependent manner. This multisensory benefit was found to be both qualitatively and quantitatively similar to that induced by a transient and target synchronized sound in the same paradigm. Studying the underlying neural mechanisms using electric neuroimaging (EEG), we found that acoustic noise alters occipital alpha (8–12 Hz) power and decreases beta-band (14–20 Hz) coupling of occipital and temporal sites. Task-irrelevant and continuous sounds thereby have an amplitude-dependent effect on cortical mechanisms implicated in shaping visual cortical excitability. The same oscillatory mechanisms also mediate visual facilitation by transient sounds, and our results suggest that task-related sounds and task-irrelevant background noises could induce perceptually and mechanistically similar enhancement of visual perception. Given the omnipresence of sounds and noises in our environment, such multisensory interactions may affect perception in many everyday scenarios.

INTRODUCTION

Multisensory interactions between visual and auditory stimuli are often considered in the context of complex stimuli such as object or speech recognition or the localization of spatio-temporal attributes. Indeed, much progress has been made in elucidating the perceptual and neural constraints underlying multisensory integration (e.g., Spence, 2011; Angelaki, Gu, & DeAngelis, 2009; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008; Stein & Stanford, 2008; Kayser & Logothetis, 2007). Previous multisensory research has shown that sensory interactions are not bound to object-like features but also occur for very simplistic and apparently unrelated stimuli (Spence, 2011; Noesselt et al., 2010; Burr & Alais, 2006; Seitz, Kim, & Shams, 2006; Doyle & Snowden, 2001). For example, it has been shown that visual contrast detection can be enhanced when the visual target is accompanied by a simultaneous auxiliary sound, although this sound does not carry specific information relating to the visual task but simply flags the relevant moment in time (Caclin et al., 2011; Lippert, Logothetis, & Kayser, 2007; Odgaard, Arich, & Marks, 2004). Transient sounds can also facilitate visual search (Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008), and related effects occur for various combinations of sensory modalities (Thorne & Debener, 2008; Bresciani et al., 2005; Odgaard et al., 2004) and can persist for extended periods (Fiebelkorn et al., 2011; Naue et al., 2011). Research on multisensory perception hence suggests that auditory and visual information can interact even for apparently unrelated stimuli that simply occur within short spatio-temporal windows (Stein & Stanford, 2008).

Another line of research has emphasized the effects of sensory noise on perception and neural activity and has also provided evidence for multisensory interactions arising from apparently unrelated stimuli. Research on stochastic facilitation (McDonnell & Ward, 2011; Ward, MacLean, & Kirschner, 2010) has shown that continuous noise in one sensory modality can facilitate sensory detection (Lugo, Doti, & Faubert, 2012; Manjarrez, Mendez, Martinez, Flores, & Mirasso, 2007; Harper, 1979) or alter neural responses in another modality (Goksoy, Demirtas, & Ates, 2005). Related benefits of apparently continuous and irrelevant stimuli on perception have also been reported in the context of music (Jausovec & Habe, 2004) and may have implications on sensory learning or cognitive performance in general (Shams & Seitz, 2008; Dalton & Behm, 2007; Sikstrom & Soderlund, 2007; Soderlund, Sikstrom, & Smart, 2007).

Although both lines of research provide evidence that functionally well characterized multisensory perceptual benefits can arise from continuous and apparently unrelated auxiliary stimuli, the direct relation between the

¹Max Planck Institute for Biological Cybernetics, ²University of Glasgow

© 2014 Massachusetts Institute of Technology

Journal of Cognitive Neuroscience 26:4, pp. 699–711
doi:10.1162/jocn_a_00524

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/jocn_a_00524 by guest on 09 July 2021
multisensory impact of transient and continuous sounds remains unclear. In particular, the use of different experimental tests and hypothesis makes it difficult to judge in how far the perceptual and neural mechanisms underlying the integration of object-based stimuli and those underlying the facilitation by continuous or noise-like stimuli are related. This study directly addresses this issue and studies the enhancement of visual perception by continuous acoustic noise with two objectives. First, we tested whether the auditory enhancement of visual detection by continuous noise is qualitatively comparable and occurs in the very same paradigm that previously revealed multisensory enhancement by transient (target-synchronized) sounds. And second, we studied the underlying patterns of neural activity that mediate such an enhancement of visual perception by continuous sounds. Using a two-alternative forced-choice (2AFC) task to measure visual contrast detection, we found that an auxiliary continuous acoustic noise shifts psychometric curves in a manner dependent on the noise’s amplitude. Using neuroimaging (EEG), we then found that acoustic noise modulates the oscillatory power of parieto-occipital alpha oscillations and the beta coherence of occipital and temporal sites. Thereby, auditory noise systematically alters the excitability of visual areas by tapping into oscillatory mechanisms—the same mechanisms that have been shown to also contribute to the enhancement of visual perception by transient sounds and that supposedly reflect attention-related excitability changes in visual cortices (Romei, Gross, & Thut, 2012; Thut, Miniussi, & Gross, 2012; Romei, Murray, Cappe, & Thut, 2009).

METHODS

General Experimental Procedures

Adult volunteer participants (aged 18–35 years, both sexes) were paid to participate in the experiments. All reported normal hearing and normal or corrected-to-normal vision and gave written informed consent before participation. The experiments were approved by the joint ethics committee of the University Clinic and the Max Planck Institute for Biological Cybernetics Tübingen and performed according to the Declaration of Helsinki. Experiments were conducted in a sound-attenuated and dark room. Visual stimuli were presented on a gamma-corrected monitor (24-in., 60 Hz) positioned 57 cm from the participant’s head, and acoustic stimuli were presented binaurally using a Sennheiser In-Ear headphone (Model PMX 80, Germany). Stimulus presentation was controlled from Matlab (Mathworks, Inc., Natick, MA) using routines from the Psychophysics toolbox (Brainard, 1997; Pelli, 1997). Sound levels were calibrated using a condenser microphone (Bruel&Kjær 4188, Copenhagen) and a sound level meter (2238 Mediator, Bruel&Kjær). The participant’s head was stabilized using a chin rest, and they used a computer keyboard to indicate their responses. We obtained data for two experiments, one only involving psychometric tasks and one involving the same task during EEG measurements.

Sensory Stimuli and Tasks

Participants performed a visual contrast detection task in a 2AFC procedure (Figure 1). Specifically, they were required to detect a dim and vertically oriented Gabor patch (1.2° diameter, two cycles/degree) randomly presented to the left or right (6° eccentricity) of a central fixation dot on a neutral gray screen (3 cd/m² background luminance). This target was shown for one frame (~16 msec) in a random period (600–1200 msec, uniform) after the onset of a central fixation dot. The contrast of the Gabor patch varied in eight steps from 0% to 25% root-mean-square contrast. After a delay period (400–700 msec), a question mark appeared centrally on the screen cueing the participant to respond by indicating on which side the target had appeared. Participants were instructed to respond as accurately as possible. Contrast values and side of presentation were varied pseudorandomly. Trials were grouped in blocks of 256, and intertrial intervals were random to avoid specific regularities in the paradigm (1500–2000 msec). We performed two experiments, one involving only psychophysical tests and one involving one of these tasks combined with EEG measurements.

In Experiment 1, we compared two multisensory versions of this basic visual contrast detection task. In one paradigm (“click sound”), the visual target was accompanied by a “click” sound (65-dB SPL; 20 msec including 5 msec on/off cosine ramp) synchronized to the target randomly on 50% of the trials. This paradigm is very similar to one used in previous studies on the multisensory enhancement of visual perception by synchronized sounds (Chen, Huang, Yeh, & Spence, 2011; Lippert et al., 2007). Participants (n = 13) performed three blocks of 256 trials each, resulting in 48 trials per contrast and sound level. In the second paradigm (“continuous noise”), we presented continuous acoustic white noise (200–16 kHz, 44.1-kHz sample rate). The level of the acoustic noise was constant for a period of 64 consecutive trials and varied in four levels (65, 68, 71, and 74 dB(A) SPL) across groups of 64 trials in a pseudorandom fashion. Importantly, this noise was continuously present during each trial and intertrial epoch and formed a task-irrelevant background sound that changed about every 4 min (corresponding to 64 trials). Participants (n = 13) performed six blocks of 256 trials each, resulting in 48 trials per contrast and noise level.

In a third paradigm (“control”), we employed synchronized click sounds but, this time, using clicks of varying intensity (65, 68, 71, and 74 dB SPL; random intensity on each trial). Participants (n = 14) performed six blocks of 256 trials each, resulting in 48 trials per contrast and click level.
In Experiment 2, we repeated the “continuous noise” paradigm while acquiring EEG data. Participants (n = 11) performed six blocks of 256 trials each, resulting in 48 trials per contrast and noise level. For these participants, we also obtained EEG data in response to the presentation of flashed sequences of Gabor patches. This paradigm was used to define those frequency bands of oscillatory activity that are specifically associated with the kind of visual stimuli used in the main study (parafoveal, small Gabor patches), as described below.

EEG Recording Procedures

EEG signals were continuously recorded using a 64-channel actiCAP system (Brain Products, Gilching, Germany) with Ag–AgCl electrodes placed according to the standard 10–10 system. Recordings were referenced to an electrode placed on the nose tip, and the ground electrode was placed in position AFz. A third electrode was placed over the lower left orbit to register eye movements. Electrode impedance was kept under 10 kΩ. Signals were amplified using BrainAmp amplifiers (Brain Products, Gilching, Germany), and data were acquired at a sampling rate of 500 Hz using a band-pass filter of 0.318–250 Hz.

Data Analysis

Data were analyzed in Matlab. Behavioral data were analyzed by calculating the psychometric curves displaying correct performance as a function of target contrast. To this end, we lumped visual targets presented on the left
and right sides of the screen together, as performance did not differ between hemifields (e.g., for Experiment 2, there was no statistical difference between these across participants; paired \( t \) test, \( p = .75 \)). Psychometric curves were fit using a Weibull function for each participant and condition using the psignifit toolbox version 2.5.6 for Matlab, which implements the maximum-likelihood method described by Wichmann and Hill (2001). Error bars shown in Figure 1 were obtained from the bootstrap procedure by Wichmann and Hill. From these fits, we extracted the critical shape parameters of the Weibull function, characterizing contrast threshold and slope. For comparison across participants, we calculated normalized slope and threshold values, obtained for each participant by dividing the actual value by the mean across auditory noise levels.

EEG data were analyzed partly using functions from the EEGLAB toolbox (Delorme & Makeig, 2004). Individual trials were rejected as containing artifacts if the amplitude on any of the channels exceeded \( \pm 100 \) \( \mu \)V anywhere or exceeded 8 SDs of the signal for a period longer than 50 msec. Data from two participants had to be discarded because of excessive noise and movement artifacts. On average, 39 \( \pm 6 \) trials per contrast and noise level remained after artifact rejection. The remaining trials were used to quantify the response evoked by the visual target as well as the oscillatory activity in the pretarget period. Before analysis, the data were rereferenced to a common average reference. Trials were split into those in which participants indicated the correct side of target presentation (“seen” trials, on average 73%, suggesting that they likely perceived the target) and trials wherein they indicated the wrong side (“not seen” trials, 27%). EEG results are presented as topographic maps (where applicable) and were summarized using an ROI analysis based on four ROIs of interest (frontal, central, temporal, and occipital electrodes; cf. Figure 2A). As the visual stimuli used in this experiment were small and, by nature of the task, were of very low contrast, we used a differential procedure to increase the signal-to-noise ratio of the visual evoked response. Specifically, for each electrode, we considered only trials in which the target was presented on the contralateral hemisphere, and we computed the difference of the response evoked on this electrode minus the response on the opposing electrode on the other hemisphere (for which the target was ipsilateral). This procedure effectively reduced noise by considering only stimuli on the dominantly stimulated hemisphere and by subtracting any concomitant trial-by-trial variations in general cortical state also present on the other hemisphere (Donner et al., 2007; Linkenkaer-Hansen, Nikulin, Palva, Ilmoniemi, & Palva, 2004). For the topographic map of visual evoked responses (Figure 2B), we interpolated the values for midline electrodes between hemispheres using the neighboring electrodes. The amplitude of the evoked response (negativity) was quantified in a time window of 60 msec centered on the peak negativity over occipital cortex (210–270 msec). Using somewhat shorter or longer windows did not affect the results.

Oscillatory activity was analyzed in the prestimulus period using either spectral Fourier analysis or narrowband filters and considering only “seen” trials. For this
analysis, we considered all trials and used the original (nondifferentiated) signal for each electrode. To estimate the oscillatory power in the pretarget period (−400 to −100 msec before target), we computed the spectral density across trials using Welch’s periodogram method by concatenating all prestimulus periods and obtaining the overall spectral density estimate. Importantly, for this analysis, we considered only this pre-stimulus window to avoid potential carryover effects from the poststimulus period (VanRullen, 2011). Using slightly different window durations did not affect results. To estimate the coherence between frequency bands in the pre-target period, we computed the spectral cross-density across trials for pairs of electrodes using Welch’s periodogram method. Both power and coherence were quantified at the sensor level and averaged across ROIs within each participant. To estimate the time-resolved power in individual frequency bands of interest, we filtered single trial data using wavelets (Gabor wavelets, centered on the band of interest, two side lobes). Filtering was applied to the interval of −400 to +400 msec around the target, to be able to assess the impact of prestimulus alpha on the subsequent perceptual outcome (Figure 3). The power was obtained from the filtered signal as the squared magnitude of the Hilbert transform.

We performed the following analysis to determine whether the amplitude of the evoked response or the power and coherence of oscillatory activity varied systematically with the level of the auditory noise in a similar manner as the perceptual threshold. For each participant and ROI (or electrode), we correlated the quantity of interest with the perceptual thresholds across acoustic noise levels. The resulting correlation values were averaged across participants using the z transform (Corey, Dunlap, & Burke, 1998). For statistical assessment (both at the single participant and group level), we computed a bootstrap baseline by shuffling the assignment of noise levels across our sample of participants, we performed the following bootstrapping procedure. Random subsets of nine participants (\( n = 1000 \), drawn with replacement) were selected from the population, and for each, we calculated the correlation of coherence (“coherence effect”) and the correlation of power (“power effect”) with noise level within each ROI. We then asked whether a coherence effect between two areas occurs independently of a power effect within each area. For the temporal–occipital coherence, there was an overall effect of coherence within the bootstrap distribution (mean: \( r = -0.88, 99\% \text{ CI } [-1, -0.45] \)), whereas there was no effect of power in either of the two ROIs (temporal: mean, \( r = -0.08 \) with CI \([-0.52, 0.7]\); occipital: mean, \( r = -0.13 \) with CI \([-0.55, 0.50]\)). Importantly, the coherence effect remained when analyzed for a stratified subset of random pools that exhibited no effect of power in either ROI (requiring a power effect in each ROI of \(-2 < r < 0.2\)): For this subset, the coherence effect persisted as well and was of similar magnitude as for the full sample (mean: \( r = -0.85, 99\% \text{ CI } [-0.96, -0.58] \)).

**Selection of Frequency Bands**

Oscillatory bands of interest were determined using a paradigm involving the passive presentation of Gabor patches during fixation. From these, we calculated the stimulus-evoked oscillatory activity as the ratio of the
RESULTS

Behavioral Results: Experiment 1

In the first experiment, we compared the perceptual enhancement of visual detection provided by a brief, target-synchronized sound and a continuous acoustic noise. We measured visual contrast detection curves using a 2AFC task involving spatial and temporal uncertainty about target presentation (Figure 1). Previous work has shown that, under such conditions, a transient sound can enhance detection performance when the sound is reliably synchronized to the visual target (Chen et al., 2011; Lippert et al., 2007). We here asked whether continuous acoustic noise can induce a similar perceptual benefit dependent on its amplitude, as suggested by work on stochastic facilitation. To this end, we directly compared psychometric performance in two multisensory variants of the same visual task, one involving a target-synchronized click sound (either present or absent; Figure 1A and B) and one involving continuous white noise of four levels (Figure 1C and D). Importantly, the noise was continuously present for periods of 64 trials, and noise levels varied only across these periods (lasting about 4 min).

The data for the “click sound” paradigm (n = 13 participants) revealed a systematic leftward shift of the psychometric curves toward lower contrast values when the sound was present. This was visible in single participant data (Figure 1A) and was quantified using thresholds and slopes derived from psychometric fits for each participant. Across participants, contrast thresholds were significantly lower in the sound compared with the no-sound condition (paired t test, p = .012; Figure 1B), whereas the slope of psychometric curves did not change (p = .9). This confirms previous reports, obtained in a very similar paradigm, that an auxiliary transient sound reduces visual contrast thresholds (Lippert et al., 2007).

For the “continuous noise” paradigm, we pooled the behavioral data from both Experiment 1 (only psychophysics, n = 13) and Experiment 2 (EEG, n = 9), as both groups exhibited similar results (n = 22 participants in total). We found that contrast thresholds varied systematically and consistently with acoustic noise level (correlation r = .97, p = .0025; Figure 1D). Specifically, thresholds decreased with increasing noise level, and statistical analysis revealed a significant effect of Noise Level on contrast threshold (ANOVA, F(3, 63) = 5.93, p = .003). Subsequent analysis between softest and loudest noise levels confirmed that thresholds differed significantly between levels (Figure 1D, scatter plot; sign test, p < .001). The psychometric slope did not vary systematically with acoustic noise level (r = −.02, p = .95). Across participants, the effect of Noise Level on slope was not significant (F = 2.3, p = .08). These results demonstrate that visual perception is significantly and systematically affected by the presence of continuous acoustic noise in a manner that depends on noise amplitude.

For transient sounds, the critical feature is believed to be the reliable temporal association with the visual target (Lippert et al., 2007, but see Fiebelkorn et al., 2011), and we performed a control experiment to test whether variations in the intensity of transient sounds has an additional effect beyond the simple appearance of the sound. To this end, we varied the intensity of click sounds synchronized to the visual target in a separate paradigm (n = 14 participants). Comparing contrast thresholds across click intensities revealed no correlation with click level (r = .03, p = .96), and statistical analysis did not show a significant effect of Click Level on contrast thresholds (ANOVA, F(3, 39) = 0.1, p = .98) or the slope of the psychometric curve (correlation: r = −.35, p = .6; ANOVA, F = 0.6, p = .6). This suggests that the mechanism mediating the multisensory influence of the click sound on detection thresholds is already engaged by the sound’s appearance and is not additionally enhanced by the click’s amplitude.

Together, these results demonstrate that multisensory facilitation of visual perception by sounds can be induced to similar degrees by either a target-synchronized, transient sound or a continuously present, stationary noise. In both paradigms, we found quantitatively similar results: Sounds consistently reduced contrast thresholds without affecting the slope of psychometric curves, and the relative reduction of contrast thresholds was comparable in both paradigms (−6.2 ± 2% for click sounds; −5.9 ± 1.9% for lowest vs. highest noise level; percentage values relative to participants’ average contrast threshold). Although previous work has started to elucidate some of the mechanisms involved in mediating visual facilitation by transient sounds, it remains unclear which and whether the same mechanisms underlie the here reported shift of psychometric curves by continuous acoustic noise. The second experiment was performed to directly address this.

Noise Enhances Visual Evoked Responses: Experiment 2

We repeated the “continuous noise” paradigm during the acquisition of EEG signals (n = 11 participants; n = 9 with usable data). Given previous work showing an effect of sensory noise on rhythmic brain activity (Ward et al., 2010; Kitajo et al., 2007; Mori & Kai, 2002) and given previous work highlighting the impact of transient sounds on alpha-band rhythms (Romei, Gross, et al., 2012; Thut...
et al., 2012; Romei et al., 2009), we specifically expected an impact of noise on slow rhythms over occipital sites.

We first analyzed the neural signature of the visual response evoked by the target Gabor patches (Figure 2). To this end, we used a differential signal to overcome the rather low signal evoked by small and low contrast stimuli. This revealed clear visual evoked responses (negative deflection about 210- to 270-msec poststimulus) over the uli. This revealed clear visual evoked responses (negative rather low signal evoked by small and low contrast stimuli. To this end, we used a differential signal to overcome the rather low signal evoked by small and low contrast stimuli. This revealed clear visual evoked responses (negative deflection about 210- to 270-msec poststimulus) over the occipital ROI, with a topography highlighting an occipital–temporal origin (Figure 2B). The amplitude of these responses increased with increasing target contrast (Figure 2C), resulting in a highly significant effect of Contrast on evoked amplitude (ANOVA, $F(7, 56) = 7.0, p < .0001$). Post hoc tests revealed significantly nonzero amplitudes ($t$ tests, $p < .05$) for contrasts above 12%. Importantly, these evoked responses were only present on trials where participants correctly indicated the side of target presentation, hence suggesting that they actually saw the target (termed “seen” trials; Figure 2D). Across participants, the amplitude of the evoked response was significantly larger in “seen” compared with “not seen” trials and virtually absent in the latter (mean $\pm \text{SEM}$: $-0.86 \pm 0.13 \mu V$ vs. $-0.07 \pm 0.04 \mu V$; paired $t$ test, $p < .001$; pooling all contrast values). This effect was also present when accounting for the variation of response amplitude with contrast, as the difference between seen and not-seen trials persisted when pooling only high-contrast trials (>20% contrast; $-1.3 \pm 0.12 \mu V$ vs. $-0.15 \pm 0.02 \mu V$; $p = .022$), ruling out that differences between “seen” and “not seen” trials simply arise from higher contrast levels being more likely to be perceived. The effect also persisted when calculated based on the same number of trials (randomly selected) for “seen” and “not seen” conditions. This suggests that the evoked responses are a good proxy to the neural response in visual cortices contributing to the participant’s percept.

We then compared the amplitude of the visual response across levels of the auditory noise. We hypothesized that the effect of auditory noise on this evoked response should resemble the noises’ perceptual effect. Thus, we expected a larger visual response for higher noise levels. To test this, we quantified whether and in which ROI the amplitude of the (negative) evoked response increased proportionally with the reduced perceptual threshold. To bootstrap the actual correlation between response amplitude and thresholds, we used a shuffling procedure to establish a significance level under our null hypothesis of no relation between amplitude and noise level. This analysis revealed that the neural–perceptual correlation between visual evoked amplitude and perceptual thresholds was weak and not significant over fronto-central ROIs (median: $r = .20$ and .23, $p > .05$) and stronger and significant over temporal ($r = .5, p < .05$) and occipital electrodes ($r = .48, p < .01$; Figure 2E). Over occipital regions, evoked responses were stronger (more negative) when auditory noise level was higher (Figure 2F). When tested at the single participant level, these correlations reached significance (at $p < .05$, bootstrap) for 80% of the participants (seven of nine) for the occipital ROI. Overall, we found that auditory noise lowers both perceptual contrast thresholds and increases occipital response amplitudes for the same stimuli.

**Auditory Noise Influences Oscillatory Perceptual Gating: Experiment 2**

We then tested whether the reduction of contrast thresholds and the increase of response amplitudes could be attributed to a specific oscillatory mechanism. Specifically, we asked whether and in which ROI the presence of the acoustic background noise affected oscillatory neural activity in the pretarget period. For each ROI, we quantified the correlation of oscillatory power with increasing noise level relative to a bootstrap baseline. This revealed that, across ROIs and frequency bands, only two bands showed a significant effect. For alpha (8–12 Hz) and beta (14–20 Hz; Figure 3A), the oscillatory power was significantly lower when the auditory noise level was higher. This was significant for alpha over centro-occipital ROIs ($p < .05$ for central and $p < .01$ for occipital ROIs), and a topographic display revealed a localization over parieto-occipital sites (Figure 3B). For beta, this was significant over fronto-central ROIs ($p < .05$), and the topography revealed a prominence of this effect at slightly lateral central sites. The continuous auditory background noise hence systematically affected the baseline oscillatory activity over occipital and parietal sites such that alpha power was systematically reduced with increasing noise level.

Oscillations in the alpha band have been linked with the neural and perceptual excitability of visual areas (Lange, Oostenveld, & Fries, 2013; Romei, Thut, Mok, Schyns, & Driver, 2012; VanRullen, Busch, Drewes, & Dubois, 2011; Romei et al., 2008). We directly confirmed this for the present data (Figure 3C). The alpha power in the occipital ROI was significantly lower when participants correctly indicated the target (“seen” trials) compared with when they did not (“not seen” trials; paired $t$ test, $p = .006$ when using all trials for each condition, $p = .04$ when using the same number of trials for each condition). This effect also persisted when analyzed separately for each noise level, hence discounting effects of noise on perception or oscillatory power. No significant difference was found for the beta band in the same region ($p = .5$). This effect of alpha power on visual detection was strongest about 200 msec before target (Figure 3C, right) in agreement with previous studies (Mathewson et al., 2012; Busch, Dubois, & VanRullen, 2009). This suggests that auditory noise impacts visual perception by affecting the sustained level of oscillatory power in a frequency band that is considered to play an important role in shaping visual perception.

We finally asked whether the noise affects oscillatory coherence between spatially separate sites (between ROIs). As for power, we quantified the correlation of oscillatory coherence with noise level relative to a bootstrap

Gleiss and Kayser 705
baseline in a prestimulus epoch (VanRullen, 2011). This revealed systematic coherence changes with noise level in alpha and beta bands between some ROIs, with reduced coherence for higher noise intensity. Figure 4 shows the significant coherence effects in the beta band (at least p < .01, bootstrap). Given partly equivalent results for oscillatory power, some of these coherence effects may mainly be because of changes in local power (Palva & Palva, 2012), and for some of these coherence effects, an explanation by power cannot be ruled out. However, the change in beta coherence between temporal and occipital ROIs is likely to be independent of local changes in power for two reasons. First, neither of these ROIs revealed a significant effect in beta power as shown above. Second, an additional bootstrap analysis directly confirmed that correlations of coherence with noise level occur independently of power effects within our set of participants (cf. Methods for details). The beta-band coherence between temporal and occipital ROIs was significantly reduced with increasing noise level (r = −.95, p = .04; Figure 4), suggesting that acoustic noise alters not only alpha power in temporal and occipital regions but also the beta-band coherence between these.

**DISCUSSION**

**Multisensory Facilitation by Transient and Continuous Sounds**

Multisensory processing comprises both the voluntary integration of feature-based and object-specific information provided by each modality (e.g., audio-visual speech or shape-related information) and the more automatic binding of spatio-temporally coincident events (Stein & Stanford, 2008; Ernst & Bulthoff, 2004). Many studies have shown that multisensory inputs can enhance perceptual performance in a variety of low-level tasks involving the detection or the perceived intensity of faint stimuli. For example, a simultaneous tone can improve the detection of dimly flashed lights (Noesselt et al., 2010; Teder-Salejarvi, Di Russo, McDonald, & Hillyard, 2005; Odgaard, Arieh, & Marks, 2003; McDonald, Teder-Salejarvi, & Hillyard, 2000), facilitate the search for dynamic visual targets (Van der Burg et al., 2008), or enhance the perceived brightness of light (Stein, London, Wilkonson, & Price, 1996). Similar effects occur in other combinations of modalities (Thorne & Debener, 2008; Odgaard et al., 2004; Lovelace, Stein, & Wallace, 2003) and may be further enhanced when highly behaviorally relevant stimuli are involved (e.g., looming sounds; Romei, Murray, Cappe, & Thut, 2013; Cappe, Thelen, Romei, Thut, & Murray, 2012). Although the exact mechanisms underlying the multisensory facilitation in these paradigms are still a matter of debate, proposed mechanisms rely on the transient nature of the multisensory auxiliary stimulus, which shapes perception of simultaneous and briefly delayed stimuli across modalities (Fiebelkorn et al., 2011). Transient auxiliary events were suggested to effectively reduce the temporal uncertainty of sensory inputs (Chen et al., 2011; Lippert et al., 2007), to induce a transient boosting of sensory energies (Zannoli, Cass, Mamassian, & Alais, 2012; Andersen & Mamassian, 2008; Downar, Crawley, Mikulis, & Davis, 2000), or to enhance momentary sensory saliency by capturing attention (Chen et al., 2011; Spence & Santangelo, 2009; Spence, Senkowski, & Roder, 2009).

A separate line of research has emphasized potential benefits of noise for the nervous system (McDonnell & Ward, 2011) and has shown that auxiliary noise can enhance perception both within a sensory modality and across modalities (Lugo et al., 2012; Lugo, Doti, & Faubert, 2008; Manjarrez et al., 2007; Harper, 1979). Some studies have directly interpreted this phenomenon in the context of stochastic resonance, whereby intermediate noise levels are particularly effective in facilitating the transfer in complex systems (Moss, Ward, & Sannita, 2004; Wiesenfeld & Moss, 1995). Other work suggests that similar kinds of multisensory facilitation may also occur in the context of more structured continuous stimuli such as music (Jausovec & Habe, 2004). Regardless of interpretation, our present results extend these findings on noise-induced benefits and provide a direct comparison with multisensory benefits induced by transient stimuli. First, we characterize a multisensory noise-induced perceptual benefit across a larger group of participants using a 2AFC procedure, thus ruling out changes in decision criterion or response bias as possible explanations (see, e.g., Manjarrez et al., 2007; Odgaard et al., 2003). Second, by directly comparing both kinds of auditory facilitation in the same task design, we show that both transient sounds and continuous noise can facilitate visual perception in the same qualitative manner (shift of psychometric curve without change in slope) and to the same quantitative degree. Thus, it seems that both transient sounds and continuous acoustic noise hence can induce highly similar functional facilitation of visual perception.
Of the several proposed explanations for the auditory facilitation of visual detection (uncertainty reduction, boosting of sensory energy, and attention-like enhancement), only the attention-related explanation seems applicable to both transient and continuous stimuli. Attention-related processes shape sensory perception, and their functional impact is well characterized using psychometric curves. Attention is often viewed as shifting psychometric functions without changing their slope (Cameron, Tai, & Carrasco, 2002; Carrasco, Penpeci-Talgar, & Eckstein, 2000; Lee, Itti, Koch, & Braun, 1999 but see Cohen & Maunsell, 2011), and the normalization model of attention predicts exactly such a shift under conditions of small stimuli and a large attention field, as it was the case here (Herrmann, Montaser-Kouhsari, Carrasco, & Heeger, 2010; Reynolds & Heeger, 2009). With this in mind, a parsimonious explanation is that both the appearance of a click sound and an increased level of acoustic noise both engage attention-related processes and thereby facilitate visual detection. This interpretation could also explain why changing the amplitude of the click sound had no additional effect beyond its basic presence: The attentional resource attracted by a sudden transient sound was already fully engaged for each of the intensities considered in our experiment. Yet, it is difficult to determine whether the functional mechanisms actually implementing the multisensory enhancement are indeed similar to those involved in mediating attention simply from behavioral data (Perez-Bellido, Soto-Faraco, & Lopez-Moliner, 2013; Chen et al., 2011 and discussions therein).

One cannot rule out that changes in noise level affected the level of participant’s arousal or vigilance. Arousal, at least as mediated by classical neuromodulatory mechanisms, changes on time scales slower than those that can account for effects by transient sounds (Lee & Dan, 2012; Jones, 2003), hence may differentially affect our two paradigms. In addition, although participants were instructed to focus on the visual task and to ignore sounds, it may be that the degree of attentional deployment varied with noise. Given that focused attention can control the selective processing of sensory information, for example, by virtue of oscillatory mechanisms (Schroeder & Lakatos, 2009; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008) as discussed below, it remains unclear to what degree the effect of acoustic noise on visual perception persists in everyday scenarios or could be exploited in more general cognitive paradigms (Soderlund et al., 2007).

### Oscillatory Gating of Visual Perception

Previous work has investigated the cortical mechanisms underlying the multisensory facilitation by transient stimuli. Several neuroimaging studies revealed the modulation of early visual cortex by the presentation of brief sounds (Mercier et al., 2013; Naue et al., 2011; Noesselt et al., 2010; Werner & Noppeney, 2010; Martuzzi et al., 2007; Watkins, Shams, Josephs, & Rees, 2007; Molholm et al., 2002). In addition, a series of elegant studies has scrutinized the perceptual and neural mechanisms underlying the excitation of visual cortices by transient sounds. Using TMS, Romei and colleagues have shown that brief sounds can facilitate the excitability of visual cortex and render it more susceptible to the induction of visual phosphenes (Romei, Gross, et al., 2012; Romei et al., 2009; Romei, Murray, Merabet, & Thut, 2007), whereas additional studies showed that both the spatio-temporal alignment as well as acoustic properties of the sound can modulate this effect (Spierer, Manuel, Bueti, & Murray, 2013; Bolognini, Senna, Maravita, Pascual-Leone, & Merabet, 2010; Romei et al., 2009). Importantly, this auditory influence on visual cortices had a direct neural underpinning in oscillatory activity, as acoustic stimulation induced phase-locked alpha-band oscillations over parieto-occipital areas (Romei, Gross, et al., 2012). The posterior alpha rhythm originates from networks in visual areas (Sauseng et al., 2009; Klimesch, Sauseng, & Hanslmayer, 2007), and the state of alpha oscillations is a good proxy to the excitability of visual areas (Thut et al., 2012; Banerjee, Snyder, Molholm, & Foxe, 2011; Busch & VanRullen, 2010; van Dijk, Schoffelen, Oostenveld, & Jensen, 2008; Thut, Nietzel, Brandt, & Pascual-Leone, 2006). Auditory facilitation of visual cortex by transient sounds hence may be explained by sound-induced changes in processes that can shape the sampling of the visual environment (Lange et al., 2013; Thut et al., 2012; VanRullen et al., 2011).

Our present findings suggest that the facilitation of visual perception by continuous sounds could be explained by very similar mechanisms. Previous work relating occipital oscillations and visual perception has shown that reduced levels of alpha power are predictive of improved or faster visual perception (Lange et al., 2013; VanRullen et al., 2011; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009; Thut et al., 2006). We here found that alpha power (before any visual stimulus) was reduced by the presence of acoustic noise, with higher noise leading to a stronger alpha reduction. Thereby, our results show that continuous noises can impact on very specific neural mechanisms that affect the gain of visual cortical circuits and facilitate visual perception. Changes in visual cortical excitability indexed by alpha band activity have also been implicated in mediating visuospatial attention (Jensen, Bonnefond, & VanRullen, 2012; Romei, Thut, et al., 2012; Thut et al., 2006; see Foxe & Snyder, 2011), a notion that fits well with the interpretation that our behavioral results reflect changes in attentional or arousal states. Together with previous work, our results show that both transient and stationary continuous sounds can induce functionally similar enhancements of visual perception and do so by involving those oscillatory mechanisms that have been implicated in shaping visual perception in relation to attention and general changes in cortical excitability.

A point worth keeping in mind is that we did not directly investigate the dependency of perception on oscillatory phase. It has been shown that stimulus detection is...
affected not only by oscillatory power but also by the precise temporal state of the oscillation (i.e., its phase) before the appearance of the stimulus (Romei, Gross, et al., 2012; Busch et al., 2009; Mathewson et al., 2009; see Henry & Oleser, 2012; Ng, Schroeder, & Kayser, 2012, for similar results in the auditory system). In the paradigm involving transient sounds (presented on a silent background), the prestimulus oscillatory phase is in a spontaneous state and most likely influenced by endogenous attentional or top–down processes. In the acoustic noise paradigm, in contrast, occipital alpha oscillations are affected by the acoustic input, and previous work has shown that sounds in general can impact the precise timing of occipital oscillations by means of phase resets or induced evoked potentials (Fiebelkorn et al., 2013; Mercier et al., 2013; Naue et al., 2011; Lakatos et al., 2009). These previous studies used transient sounds so it remains unclear whether and to what degree phase resetting or entrainment occurs for stationary and rather unstructured sounds such as white noise. However, at least attended and meaningful sounds can induce phase-entrained oscillations in visual areas (Luo, Liu, & Poeppel, 2010). Recent work shows a clear relation between the deployment of attentional resources, the selective enhancement or entrainment of cortical oscillations, and the cross-modal mechanisms that permit such an interaction (Schroeder & Lakatos, 2009; Lakatos et al., 2008). A differential impact of attention between the transient and noise paradigms used here and the selective entrainment of visual phase by attended sounds may add another level of variability not investigated in the current study.

Finally, we found that auditory noise reduced the beta coherence between occipital and other regions. Whereas some of the coherence effects may be explained by the noise-induced modulation of local power, the coupling between occipital and temporal regions seemed to be independent of power effects. Beta coherence has been associated with long-range coupling (Womelsdorf & Fries, 2007; von Stein, Chiang, & Konig, 2000) and may be specifically involved in mediating higher level control over early sensory areas (Siegel, Donner, Oostenveld, Fries, & Engel, 2008; Gross et al., 2004). Importantly, recent studies reported that the perceptual outcome of multisensory stimuli depends on the strength of long-range beta coupling (Keil, Muller, Ihssen, & Weisz, 2012; Hipp, Engel, & Siegel, 2011). Enhanced beta coherence may favor the interaction between sensory stimuli, whereas a decoupling may favor the separate processing of unisensory information. Our results can therefore be interpreted as a decoupling of sensory networks by more intense acoustic noise, which improves visual perception by ensuring that other and irrelevant stimuli do not interfere with visual processing.

Given the unknown origin of the temporal coherence seen in the present paradigm, it remains difficult to make predictions about the exact areas that supposedly would be functionally decoupled. The study by Keil et al. (2012) found strong beta effects in superior temporal cortex, a region known to be a site of multisensory binding (Werner & Noppeney, 2010; Dahl, Logothetis, & Kayser, 2009; Calvert, 2001). This would suggest a reduced coupling between higher-order multisensory regions and parieto-occipital cortex during acoustic noise. It is interesting to speculate how such changes in top–down coupling would affect the multisensory impact of a transient sound, which perceptually seems to induce a similar multisensory benefit. Future studies could, for example, address whether and how the multisensory impact of continuous and transient sounds interacts. Along this line, it would also be interesting to see what effect a gap in a continuous noise can have on visual perception. Interruptions of sounds capture attention (Kayser, Petkov, Lippert, & Logothetis, 2005) hence may have similar effects as the transient appearance of a sound.

Conclusion

Studying auditory facilitation of visual perception, we compared the influence of brief and target-synchronized sounds with that of continuous and stationary acoustic noise. Both acoustic stimuli induced comparable shifts of psychometric curves, and together with previous work, our results show that both kinds of multisensory facilitation prominently involve occipital alpha rhythms. This suggests that auditory facilitation of visual perception by widely different types of sounds may arise from the same and possibly attention-related oscillatory mechanisms, reflecting a general improvement of basic sensory perception by cross-modal events. These results are all the more important as continuous sounds and noises are omnipresent in our environment, hence unspecific multisensory interactions may be more widespread than previously thought.

Acknowledgments

This work was supported by the Max Planck Society and profited from support of the Bernstein Centre for Computational Neuroscience, Tübingen, funded by the German Federal Ministry of Education and Research (BMBF; FKZ: 01GQ1002). We are grateful to Benedict Ng for help during initial phases of the experiment.

Reprint requests should be sent to Christoph Kayser, Institute of Neuroscience and Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, United Kingdom, or via e-mail: christoph.kayser@glasgow.ac.uk.

REFERENCES


