TITLE

# Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: a frequency-tagging study

AUTHORS:

Amra Covic[1,2*], Christian Keitel[3*^], Emanuele Porcu[4], Erich Schröger[1], & Matthias M Müller[1]

AFFILIATIONS:

1 – Institut für Psychologie, Universität Leipzig, Neumarkt 9-19, 04109 Leipzig, Germany

2 – Institut für Medizinische Psychologie und Medizinische Soziologie, Universitätsmedizin Göttingen, Georg-August-Universität, 37973 Göttingen, Germany

3 – Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, 58 Hillhead Street, G12 8QB Glasgow, UK

4 – Institut für Psychologie, Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2 Gebäude 23, 39106 Magdeburg

* Joint first authors, equal contributions

^ corresponding author: christian.keitel@glasgow.ac.uk

KEYWORDS:

spatial attention, selective attention, multisensory integration, audio-visual synchrony, brain oscillation, neural rhythm, steady-state response (SSR), EEG, brain-computer interface (BCI)

Word count: Abstract (212), Introduction (1,122), Discussion (1,774)

1 Table, 4 Figures, 74 References

**ABSTRACT**

The neural processing of a visual stimulus can be facilitated by both, attending to its position or a co-occurring auditory tone. Using frequency-tagging we investigated whether facilitation by spatial attention and audio-visual synchrony rely on similar neural processes. Participants attended to one of two flickering Gabor patches (14.17 and 17 Hz) located in opposite lower visual fields. Gabor patches further "pulsed" (i.e. showed smooth spatial frequency variations) at distinct rates (3.14 and 3.63 Hz). Frequency-modulating the auditory stimulus at the pulse-rate of one visual stimulus established audio-visual synchrony. Flicker and pulsed stimulation elicited stimulus-locked rhythmic electrophysiological brain responses that allowed tracking the neural processing of simultaneously presented stimuli. These steady-state responses (SSRs) were quantified in the spectral domain to examine visual stimulus processing under conditions of synchronous vs. asynchronous tone presentation and when respective stimulus positions were attended vs. unattended. Strikingly, unique patterns of effects on pulse- and flicker driven SSRs indicated that spatial attention and audiovisual synchrony facilitated early visual processing in parallel and via different cortical processes. We found attention effects to resemble the classical top-down gain effect facilitating both, flicker and pulse-driven SSRs. Audio-visual synchrony, in turn, only amplified synchrony-producing stimulus aspects (i.e. pulse-driven SSRs) possibly highlighting the role of temporally co-occurring sights and sounds in bottom-up multisensory integration.

## INTRODUCTION

Behavioral goals, as well as the physical properties of sensory experiences, shape how neural processes organize the continuous and often rich influx of sensory information into meaningful units. One such process, selective attention, serves to prioritize currently behaviorally relevant sensory input while attenuating irrelevant aspects (Posner et al., 1980; Treisman and Gelade, 1980). In a visual search display, for example, items matching the color or orientation of a pre-defined target stimulus undergo prioritized processing relative to other items (Treisman and Gelade, 1980; Wolfe, 1994; Wolfe et al., 1989). Another process exploits the spatial and temporal structure of dynamic sensory input, extracting regularities either in the visual modality alone (Alvarez and Oliva, 2009; Lee, 1999) or, by cross-referencing co-occurrences across sensory modalities (Fujisaki and Nishida, 2005). In fact, aforementioned visual search can be drastically improved by presenting a spatially uninformative tone pip that coincides (repeatedly) with a sudden change in target appearance in a dynamic search array (Van der Burg et al., 2008). This pop-out effect has been ascribed to a gain in relative salience of the target stimulus caused by the unique integration of auditory and visual information. The impression of a multisensory object hereby hinges on the temporal precision of coinciding unisensory inputs, also termed audio-visual synchrony, a critical cue for multisensory integration (Werner and Noppeney, 2011). Consecutive synchronous co-occurrences of the same auditory and visual stimulus components further increase the likelihood of multisensory integration (Parise, 2012). Generalizing this multisensory effect to our everyday experience of dynamic cluttered visual scenes, Talsma et al (2010) put forward that multisensory objects tend to involuntarily attract attention towards their position. As a consequence, they would gain an automatic processing advantage over unisensory stimuli. In a task that requires a sustained focus of attention on a specific position in the visual field multisensory stimuli may then act as strong

distractors (Krause et al., 2012) because they withdraw common processing resources from the task-relevant focus of attention.

Interestingly, this influence seems to work both ways: As Alsius et al. (2005) have shown focusing on a visual task impedes the integration of concurrent but irrelevant visual and auditory input. This effect has been related to the concept of the temporal binding window, a period during which co-occurring attended visual and auditory stimuli are most likely to be integrated (Colonius and Diederich, 2012). The window can expand for stimuli appearing at attended locations but remains unaffected (or contracts) when spatial attention is averted (Donohue et al., 2015).

Both phenomena - the involuntary orientation of spatial attention towards multisensory events as well as impeded multisensory integration when maintaining focused attention - have largely been studied in isolation (Talsma et al., 2010). We frequently encounter situations, however, in which the two biases can act concurrently. Additionally, they may fluctuate between having conjoined and conflicting effects over time depending on whether attended positions and multisensory events overlap or diverge in the visual field.

This complex interplay therefore warranted a dedicated investigation in a paradigm that allowed contrasting both cases directly. In the present study, we manipulated trial by trial whether participants attended to a dynamic audio-visual synchronous stimulus while leaving a concurrently presented asynchronous stimulus unattended or vice versa.

We probed early cortical visual processing by tagging stimuli with distinct temporal frequencies (Norcia et al., 2015; Regan, 1989). This frequency-tagged stimulation elicited periodic brain responses, termed steady-state responses (SSRs). SSRs index continuous processing of individual stimuli in multi-element displays and have been demonstrated to indicate the allocation of spatial attention (Kim et al., 2007; Müller et al., 1998a; Walter et al., 2012) as well as audio-visual synchrony, (Jenkins et al., 2011; Keitel and Müller, 2015; Nozaradan et al., 2012).

Crucially, employing frequency-tagging allowed us to tease apart the relative facilitating effects of both factors as follows: Our paradigm featured two Gabor patches, one per lower visual hemifield, that each displayed two rhythmic physical modulations: As in classical frequency-tagging experiments they were flickering on and off at different rates (14.17 and 17 Hz, respectively). Additionally, spatial frequencies of the Gabor patches modulated smoothly at slower rates (3.14 and 3.62 Hz, respectively), which gave the impression of a pulsation-like movement (see *Figure 1*). The latter modulation was exploited to introduce audio-visual synchrony with a concurrently presented tone that carried a frequency modulation with the same temporal profile as one of the visual stimuli (Giani et al., 2012; Hertz and Amedi, 2010 for similar approaches; see Keitel and Müller, 2015). Participants were then cued randomly on each trial to attend to one of the two stimulus positions, while one of the two Gabor patches pulsed in synchrony with the tone. This paradigm enabled comparisons of SSR-indexed visual processing between four cases of Gabor patch presentation: attended synchronous (*A+S+*), attended asynchronous (*A+S-*), unattended synchronous (*A-S+*) and unattended asynchronous (*A-S-*).

We expected our data to replicate well-described gain effects of top-down cued spatial attention on flicker-driven SSRs (Keitel et al., 2013; Kim et al., 2007; Müller et al., 1998a). Further, we assumed that these gain effects extend to pulsation-driven SSRs, because spatial attention should prioritize any information presented at an attended location.

Secondly, we hypothesized that in line with previous findings (Nozaradan et al., 2012) audio-visual synchrony produced gain effects on SSRs. In contrast to attentional gain, results of an earlier investigation suggested that synchrony-related gain effects may be specific to pulsation-driven SSRs. Using a paradigm similar to the present study, Keitel and Müller (2015) found that an SSR component with a frequency of twice the pulsation rate was exclusively susceptible to synchrony-related gain effects. At this rate, the stimulation presumably contained strong transients critical for establishing audio-visual synchrony

(Werner and Noppeney, 2011). If that were the case the current paradigm was expected to produce similarly selective effects. Alternatively, however, if audio-visual synchrony simply attracted spatial attention then synchrony-related facilitation should mirror the pattern of attention-related gain effects on pulse- and flicker-driven SSRs. More specifically, synchrony alone should produce gain effects for flicker-driven SSRs.

Comparable patterns of attention- and synchrony-related facilitation would further point towards an account in which they may draw upon similar resources and therefore interact in facilitating visual processing: An attended stimulus would benefit less from audio-visual synchrony compared with an unattended synchronous stimulus, because attention has already been allocated to its position. However, if attention- and synchrony-related facilitation relied on distinct neural resources, they were assumed to have independent additive effects on SSRs.

The latter finding could then be cast in a framework in which spatial attention biases are conveyed top-down via a fronto-parietal cortical network (Corbetta and Shulman, 2002), whereas audio-visual synchrony may have been established bottom-up via direct cortico-cortical connections or subcortical relays (Lakatos et al., 2009; van Atteveldt et al., 2014).

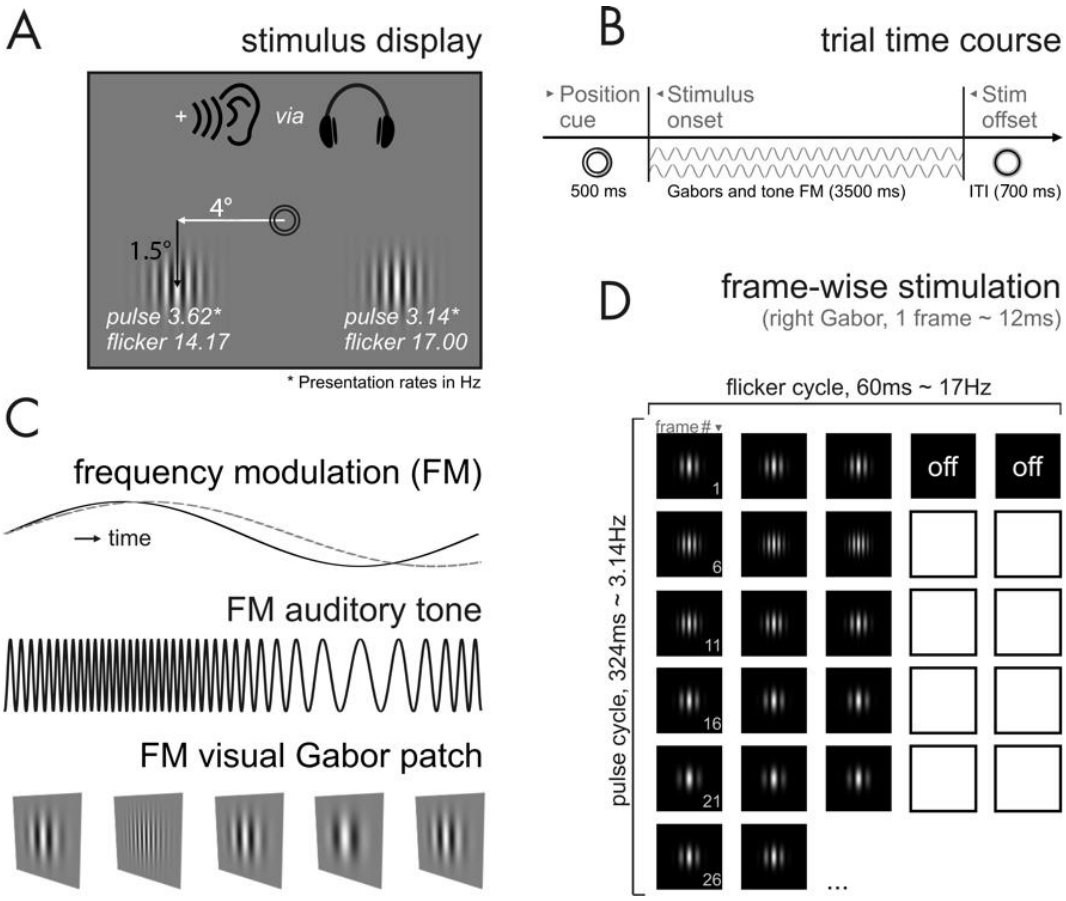


**Figure 1** Stimulation details. (A) On-screen stimulus display comprising central fixation rings and one Gabor patch per lower left and right visual hemifield. All items not to scale. Participants received auditory stimulation via headphones. (B) Schematic trial time course. An instructive position cue allocates attention to the left or right stimulus. Subsequent ongoing Gabor-patch and tone stimulation are represented by grey sinusoids. (C) A common frequency modulation (FM; solid black line) of auditory tone pitch and the spatial frequency of one of the two Gabor patches produces a synchronous pulsing audio-visual percept. Concurrently, the spatial frequency of the other Gabor patch modulates at a slightly different frequency (dashed grey line), thus rendering it asynchronous to the tone. (D) Frame-by-frame visual stimulation for the right Gabor patch. The illustration shows the first 27 frames of each trial. Note the emphasis on the on–off cycles leading to a 17-Hz flicker along the horizontal axis (white boxes depict further off-frames) and one full cycle of the spatial frequency modulation leading to a 3.14-Hz 'pulsation' along the vertical axis.

## METHODS

### Participants

We collected data from 14 participants with normal or corrected-to-normal vision and normal hearing. Participants gave informed written consent prior to experiments. None reported a history of neurological diseases or injury. They received course credit or a small monetary compensation for participation. The experiment was conducted in accordance with the Declaration of Helsinki and the guidelines of the ethics committee of the University of Leipzig.

Two participants showed excessive eye movements during EEG recordings and were thus excluded. Data of 12 participants aged 18 – 31 years (all right-handed, 9 female) entered analyses.

### Stimulation

Stimuli were presented on a 19-inch cathode ray tube screen positioned 0.8 m in front of participants. The screen was set to a refresh rate of 85 frames per second and a resolution of 1024 x 768 pixel (*width* x *height*). Visual experimental stimulation consisted of two monochrome Gabor patches with a diameter of ~3° of visual angle, one located in the lower left and the other one located in the lower right visual field at eccentricities of 4° from vertical and 1.5° from horizontal meridians (see *Figure 1a*). Stimuli were presented against a grey background (RGB: 128,128,128; luminance = 30 cd/m$^2$). Two black concentric circles (.4° of visual angle outer eccentricity, RGB: 0, 0, 0) in the center of the display served as fixation point.

Both Gabor stimuli underwent two independent periodic changes in the course of a trial: (1) The left patch presentation followed a cycle of 4 on-frames and 2 off-frames (2/1 on/off-ratio) resulting in a 17 Hz flicker. The right patch flickered at a rate of 14.2 Hz achieved by repetitive cycles of 3 on-frames and 2 off-frames (3/2 on/off-ratio). (2) While flickering, the

183 spatial frequency of the Gabor patches oscillated between a maximum of 2 Hz/° and a
184 minimum of 1 Hz/° at a rate of 3.14 Hz for the right patch and 3.62 Hz for the left patch.
185 Periodic spatial frequency changes gave the impression of alternating contractions and
186 relaxations that led to the percept of pulsing Gabor patches over time (*Figure 1b-d*). Pulse
187 frequencies were chosen based on pilot experiments that served to determine a trade-off
188 frequency range in which pulsing was readily perceptible, yet, still allowed driving periodic
189 frequency-following brain responses (SSRs).
190 In addition to the visual stimuli we presented a tone with a center frequency of 440 Hz
191 binaurally via headphones. The frequency of the tone was rhythmically modulated following
192 sinusoidal excursions from the center frequency (10% maximum excursion = ±44 Hz). On
193 each trial the modulation rate exactly matched the pulse rate of one of the two Gabor
194 patches. Common rhythmic changes over time resulted in sustained audio-visual synchrony
195 (see e.g. Schall et al., 2009).
196 Prior to the experiment, we employed the method of limits (Leek, 2001) to approximate
197 individual hearing thresholds using one of the experimental stimuli, a 3.14-Hz frequency
198 modulated tone (see e.g. Herrmann et al., 2014; Keitel and Müller, 2015). In our
199 implementation, participants listened to a series of 10 tone sequences with a maximum
200 duration of 15 s per sequence. Tone intensity changed during each sequence while
201 alternating between log-linear decreases and increases across sequences. Participants were
202 instructed to indicate by button press when they stopped or started hearing respective
203 tones. Cross-referencing button response times with tone intensity functions yielded
204 individual estimates of psychophysical hearing thresholds, i.e. sensation levels (SL). In the
205 experiment, acoustical stimulation was presented at an intensity of +35 dB SL.
206

**Procedure and Task**

Participants were seated comfortably in an acoustically dampened and electromagnetically shielded room and directed gaze towards the fixation ring on the computer screen. At the beginning of each trial, participants were cued to attend exclusively to the left or the right visual stimulus. To this end, a green semi-circle appeared inside the fixation ring for 500 ms either on the right or the left internal side of the ring to indicate the relevant Gabor patch. Subsequently, the two pulsing Gabor patches and the tone were presented for 3500 ms. At the end of each trial, the fixation ring remained on screen for an extra 700 ms allowing participants to blink before the next trial started (*Figure 1b*).

Participants were instructed to respond to occasionally occurring luminance changes of the cued Gabor patch (= targets) while ignoring similar events in the other patch (= distractors). During such events, Gabor patch luminance faded out to a minimum of 50% and back in within a 300 ms interval. Targets and distractors occurred in 50% of trials and up to 3 times in one trial with a minimum interval of 800 ms between subsequent onsets. Behavioral responses were recorded as space-bar presses on a standard keyboard. The responding hand was changed halfway through the experiment with the starting hand counterbalanced across participants.

We manipulated the two factors *attended position* (left vs. right Gabor patch) and audio-visual *synchrony* between attended Gabor patch and tone (synchronous vs. asynchronous) in a fully balanced design. Trials of the resulting four conditions – attended synchronous (*A+S+*), attended asynchronous (*A+S-*), unattended synchronous (*A-S+*) and unattended asynchronous (*A-S-*) – were presented in a pseudo-randomized order. Note that the tone was always in sync with one of the two Gabor patches. Therefore, in the two conditions in which the tone was out of sync with the attended Gabor patch, it was in sync with the unattended patch.

232 In total, we presented 600 trials (= 150 trials per condition) divided into 10 blocks (~5 min
233 each). Before the experiment, participants performed training for at least one block. After
234 each training and experimental block, they received feedback on the average hit rate and
235 reaction time.
236
237 **Behavioral data recording and analyses**
238 Responses were considered a 'hit' when the space bar was pressed between 200 to 1000 ms
239 after target onset. We further defined correct rejections as omitted responses to distractor
240 stimuli. Based on these data, we calculated the response accuracy as the ratio of correct
241 responses (number of hits and correct rejections) to the total number of targets and
242 distractors for each condition and participant. Accuracies were subjected to a two-way
243 repeated measures analysis of variances (ANOVA) with factors of *attended position* (left vs.
244 right Gabor patch) and *synchrony* (synchronous vs. asynchronous). Response speed,
245 quantified as median reaction times, was analyzed accordingly.
246 For all repeated measures ANOVAs conducted in this study effect sizes are given as $\eta^2$ (eta-
247 squared). Where applicable, the Greenhouse–Geisser (GG) adjustment of degrees of
248 freedom was applied to control for violations of sphericity (Greenhouse and Geisser, 1959).
249 Original degrees of freedom, corrected p-values ($P_{GG}$) and the correction coefficient epsilon
250 ($\varepsilon_{GG}$) are reported.
251 Further Post-hoc tests – two-tailed t-tests for paired comparisons or against zero – were
252 applied where necessary. We applied the Holm-Bonferroni procedure to correct *p*-values
253 ($P_{HB}$) for multiple comparisons (Holm, 1979).
254

255 **Electrophysiological data recording**

256 EEG was recorded from 64 scalp electrodes that were mounted in an elastic cap using a
257 BioSemi ActiveTwo system (BioSemi, Amsterdam, Netherlands) set to a sampling rate of 256
258 Hz. Lateral eye movements were monitored with a bipolar outer canthus montage
259 (horizontal electrooculogram). Vertical eye movements and blinks were monitored with a
260 bipolar montage positioned below and above the right eye (vertical electrooculogram). From
261 continuous data, we extracted epochs of 3500 ms starting at audio-visual stimulus onset. In
262 further preprocessing, we excluded (1) 50% of epochs per condition (= 75) that
263 corresponded to trials containing transient targets and distractors (= brief luminance
264 fadings) as well as (2) epochs with horizontal and vertical eye movements exceeding 25 µV
265 (= 2.5° of visual angle), or containing blinks. To correct for additional artefacts, such as single
266 noisy electrodes, we applied the 'fully automated statistical thresholding for EEG artefact
267 rejection'(Nolan et al., 2010). This procedure corrected or discarded epochs with residual
268 artefacts based on statistical parameters of the data. Artefact correction employed a
269 spherical-spline-based channel interpolation. For each participant FASTER interpolated up to
270 4 electrodes (median = 2) across recordings and an average of up to 5.6 electrodes
271 (minimum = 1.9, median = 3.6) per epoch. Note that epochs with more than 12 artefact-
272 contaminated electrodes were excluded from further analysis. In total, we discarded an
273 average of 15% of epochs per participant and condition. Subsequently, data were re-
274 referenced to average reference and averaged across epochs for each condition and
275 participant, separately. Basic data processing steps such as extraction of epochs from
276 continuous recordings and re-referencing made use of EEGLAB (Delorme and Makeig, 2004)
277 in combination with custom routines written in MATLAB (The Mathworks, Natick, MA).
278

279 **Electrophysiological data analyses**

280 In our analyses we focused on two neural markers that have been repeatedly demonstrated
281 to index attentional modulation: SSR amplitudes (Morgan et al., 1996; Müller and Hubner,
282 2002; Quigley and Müller, 2014) and SSR inter-trial phase coherence (ITC, Kashiwase et al.,
283 2012; Kim et al., 2007; Porcu et al., 2013). Both measures also reflect effects of audio-visual
284 synchrony on early visual processing (Nozaradan et al., 2012). Approaches to derive
285 amplitudes and inter-trial phase coherence differ and thus are described separately below.
286 Both approaches required spectral decompositions of EEG time series for which we used the
287 Fieldtrip toolbox (Oostenveld et al., 2011).

288

289 *SSR power*

290 Artefact-free epochs were truncated to segments of 3000 ms that started 500 ms after
291 audio-visual stimulation onset and averaged separately for each EEG sensor, experimental
292 condition and participant. The first 500 ms were omitted in order to exclude event-related
293 potentials to stimulus onset from spectral analyses. From de-trended (i.e. linear trend
294 removed) 3000 ms segments we quantified power (= squared amplitude) spectra by means
295 of Fourier transforms. *Figure 2a* illustrates that our stimulation was effective in driving
296 distinct SSRs: Power spectra pooled across all 64 scalp electrodes and experimental
297 conditions showed clear peaks at the stimulation rates. Notably, spectra revealed strong
298 harmonic responses at twice the pulse frequencies (6.28 and 7.24 Hz). We included these
299 pulse-driven harmonics in further analyses because fundamental and harmonic responses
300 have been hypothesized to reflect different aspects of stimulus processing (Kim et al., 2011;
301 Pastor et al., 2007; Porcu et al., 2013) and showed modulation by synchrony in a previous
302 study (Keitel and Müller, 2015).
303 Grand-average topographical distribution of pulse-driven as well as flicker-driven SSR power
304 averaged over conditions showed widespread maxima at parieto-occipital electrode sites

(scalp maps in *Figure 2a*) that are typically observed in experiments with lateralized flicker stimulation (see e.g. Keitel et al., 2013).

For each participant and condition, SSR amplitudes were averaged across a cluster of 15 electrodes covering parieto-occipital maxima (Oz, O1, O2, Iz, I1, I2, POz, PO3, PO4, PO7, PO8, P7, P8, P9, P10; as indicated in left-most scalp map in *Figure 2a*) and normalized by taking the decadic logarithm, then multiplying it by 20, to yield dB-scaled values (termed log-power in the following).

SSR log power was subjected to four-way repeated measures analysis of variances (ANOVAs) with factors of driving *stimulus position* (left vs. right hemifield), *attention* (attended vs. unattended), *synchrony* (synchronous vs. asynchronous) and *SSR component* (pulse 1f, pulse 2f and flicker 1f).

The factor *stimulus position* had no effect on SSR log power and did not show any interaction with the other factors (see *Results*). This afforded collapsing normalized power across left and right stimuli, i.e. across pulse frequency following ('pulse 1f') 3.14 Hz and 3.62 Hz, pulse frequency doubling ('pulse 2f') 6.28 and 7.24 Hz, as well as flicker frequency following ('flicker 1f') 14.17 and 17.00 Hz SSRs, respectively, in subsequent analyses.

*SSR inter-trial phase coherence*

We computed inter-trial phase coherence (Cohen, 2014) based on Fourier transforms of artefact-free single trial epochs, truncated to 3000 ms segments (as described above for SSR amplitude analyses) according to:

$$ITC(f) = \left|\frac{1}{N}\sum_{n=1}^{N}\frac{c_n(f)}{|c_n(f)|}\right| \quad [1]$$

where $c_n(f)$ is the complex Fourier coefficient of trial n at frequency *f* and |.| indicates the absolute value. Inter-trial phase coherence as a measure of SSR modulation has been introduced to SSR analyses more recently (Kim et al., 2007; Nozaradan et al., 2012) and SSR amplitude and phase coherence have demonstrated different sensitivities to top-down

influences on sensory processing (Kashiwase et al., 2012; Porcu et al., 2013). SSR Inter-trial phase coherence can be visualized as spectra that typically display narrow peaks at stimulation frequencies and higher order harmonics (Nozaradan et al., 2012; Ruhnau et al., 2016).

Similar to SSR amplitudes, ITCs showed broad topographic maxima at parieto-occipital electrode sites. Condition-averaged ITC spectra pooled across the 15-electrode cluster as described above (see *SSR amplitudes*) revealed distinct peaks at the six frequencies of interest (*Figure 2b*).

Pooled ITCs were subjected to a four-way ANOVA with a design identical to SSR amplitude analyses. Note that ITCs were normalized by taking the natural logarithm prior to statistical evaluation.

*Power of the ongoing EEG and SSRs*

As depicted in *Figure 2c*, SSRs have very low signal-to-noise ratios when being evaluated on the basis of averaged single-trial power spectra. Instead, these spectra accentuate the typical $1/f^x$ profile of power decreasing towards higher frequencies as well as peaks in the vicinity of 10 Hz that are consistent with alpha rhythmic brain activity. In turn, these features are much attenuated in SSR 'evoked' power and ITC spectra (*Figures 2a* and *b*).

*Joint analyses of SSR amplitude and inter-trial phase coherence modulation*

As laid out in the Results section, both of our manipulations, spatial attention and audio-visual synchrony, revealed distinct patterns of effects on SSR amplitudes and ITCs. To further characterize and compare these effects we computed an index that expressed attention- and synchrony-related amplitude and ITC modulations for each subject and SSR frequency component *f* (pulse 1f, pulse 2f and flicker 1f) according to:

$$AMI_f = \frac{Amp_f^{att} - Amp_f^{ign}}{Amp_f^{att} + Amp_f^{ign}} \qquad [2]$$

where *Amp* represents SSR amplitudes pooled across left and right stimuli. This attention modulation index (AMI) expressed the net gain effect of attention (*att* = attended, *ign* = ignored). A similarly scaled synchrony modulation index (SMI) was computed by contrasting SSR amplitudes between in-sync and out-of-sync conditions. We were thus able to compare both indices directly. Entering ITCs instead of SSR amplitudes into formula (1) yielded ITC-based AMIs and SMIs.

ANOVAs carried out for SSR amplitudes and ITC revealed that *attention* and *synchrony* influenced SSRs additively, i.e. no interaction between these factors was found (see *Results*). This finding justified collapsing AMIs across synchrony conditions and SMIs across attention conditions for each SSR component, separately, in the following analyses. As an example, we pooled the AMIs expressing the gain between synchronous conditions (*A+S+* vs *A-S+*) and asynchronous conditions (*A+S-* vs *A-S-*).

Additionally, we applied a Bayesian inference approach because in contrast to the classical frequentist inference it allows determining the amount of evidence in favor of the null hypothesis ($H_0$: no interaction) explicitly. To this end, we estimated Bayes factors (Rouder et al., 2012), i.e. the plausibility of a specific model given the data, for SSR power and ITC for the following two models: an additive model A + S (*attention* plus *synchrony*) and a model additionally containing an interactive term A + S + (A * S). The analysis was performed by means of the function *anovaBF* provided by the R (version 3.3.0; R Core Team, 2013) package *Bayes factor* v0.9.12–2 (Morey et al., 2015). We adopted the Jeffrey-Zellner-Siow (JZS) prior with a standard scaling factor *r* of .707 (Rouder et al., 2012; 2009; Schönbrodt and Wagenmakers, 2015). A default of 10,000 iterations were used for the Monte-Carlo sampling. Participants were considered as random factor. Importantly, Bayesian modelling

favored the additive model (see *Results*) and further justified calculating AMIs and SMIs. Results were robust against changing scaling factors.

Finally, AMIs and SMIs were entered into a three-way ANOVA with factors of *SSR component* (pulse 1f, pulse 2f, and flicker 1f), *gain type* (attention vs synchrony) and *gain measure* (SSR amplitude vs ITC).

## RESULTS

### Behavioral data

Participants detected luminance fadings more accurately when attending to left Gabor patches (main effect *attended stimulus*: $F(1,11) = 32.30$, $P < 0.001$, $\eta^2 = 0.579$; see *Table 1*). Accuracy remained unaffected by in-sync vs. out-of-sync tone presentation (main effect *synchrony*: $F(1,11) < 1$). The interaction of both factors was not significant ($F(1,11) < 1$). Reaction times decreased slightly when participants performed the task on in-sync Gabor patches (main effect *synchrony*: $F(1,11) = 9.27$, $P < 0.05$, $\eta^2 = 0.061$; see Table 1) but was comparable between left and right stimuli (main effect *attended location*: $F(1,11) < 1$). As for accuracy, the interaction of both factors remained negligible ($F(1,11) < 1$).

**Table 1** Average behavioral performance in the visual fading detection task (N = 12).

| Attended Stimulus | | Left | | Right | |
|---|---|---|---|---|---|
| Synchrony | | *S+* | *S-* | *S+* | *S-* |
| Proportion correct (%) | *M* | 85.6 % | 84.2 % | 76.4 % | 76.8 % |
| | *±SEM* | 2.2 % | 2.0 % | 2.4 % | 2.7 % |
| Reaction time (ms) | *M* | 674 | 662 | 667 | 662 |
| | *±SEM* | 14 | 16 | 16 | 13 |

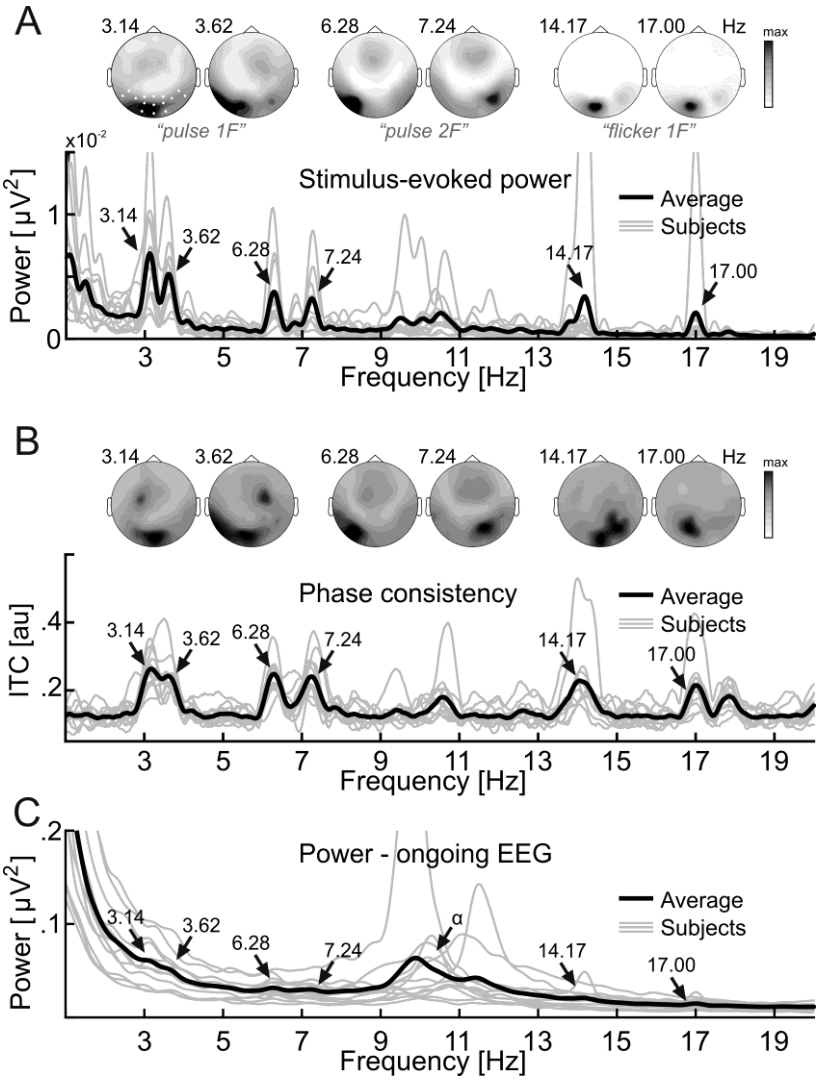


**Figure 2** Stimulus-driven steady-state responses (SSRs) – spectra and scalp maps. (A) SSR power extracted from spectral decomposition of trial-averaged EEG waveforms, thus "stimulus-evoked". Scalp maps show topographical distributions of power for the pulse-frequency following (*pulse 1f*), pulse-frequency doubling (*pulse 2f*) and flicker-frequency following (*flicker 1f*) SSR components driven by left and right stimuli respectively. White dots in left-most scalp map highlight the uniform sensor cluster used in all data analyses. Spectra below depict condition-averaged individual power spectra (grey lines) and, superimposed in black, the grand-average spectrum. Arrows indicate peaks that correspond to the respective driving frequencies (in Hz). (B) Same as (A) but for SSR inter-trial phase consistency (ITC). (C) Power spectra based on averaged spectral decompositions of single trials for comparison. Note that this approach emphasizes spectral characteristics of the ongoing EEG, such as the alpha rhythm (see peaks around 10 Hz, denoted *α*), over SSRs given our stimulation.

**EEG data**

We focused our analyses on SSR amplitudes and inter-trial phase coherence values (ITCs) to evaluate effects of spatial attention and audio-visual synchrony on early visual stimulus processing. Each stimulus drove three spectrally distinct SSR components: one at the frequency of stimulus pulsation, another one at twice the pulsation rate and a third following stimulus flicker (i.e., pulse 1f, pulse 2f and flicker frequencies, respectively). Note that for statistical evaluation, all data were collapsed across left and right stimuli and thus across specific frequencies within each category of SSR components: For example, data of 3.14 Hz (right stimulus) and 3.62 Hz (left stimulus) pulsation-driven SSRs were pooled to yield aggregate pulse 1f SSR amplitudes and ITCs.

*SSR power*

SSR power decreased with increasing stimulus presentation rate (main effect *SSR component*: $F(2,22) = 55.76$, $P_{GG} < 0.001$, $\varepsilon_{GG} = 0.90$, $\eta^2 = 0.301$; also see *Figure 3*) as has been documented extensively before (Keitel and Müller, 2015; Porcu et al., 2014). *Figure 3c* underlines that amplitudes further varied with the allocation of attention towards stimuli (main effect *attention*: $F(1,11) = 24.15$, $P < 0.001$, $\eta^2 = 0.094$) and were affected by audio-visual *synchrony* ($F(1,11) = 71.01$, $P < 0.001$, $\eta^2 = 0.067$). Amplitudes were comparable for left and right stimuli (main effect *stimulus position*: $F(1,11) < 1$). A significant *SSR component x synchrony* interaction ($F(2,22) = 37.03$, $P_{GG} < 0.001$, $\varepsilon_{GG} = 0.56$, $\eta^2 = 0.057$) warranted a closer investigation of synchrony effects on specific SSR components. The crucial *attention x synchrony* interaction ($F(1,11) = 1.12$, $P = 0.313$, $\eta^2 < 0.001$) as well as other interaction terms remained non-significant (maximum $F(2,22) = 2.94$, $P = 0.074$, $\eta^2 = 0.009$ for the *stimulus position x SSR component interaction*).

Regarding the absence of an *attention x synchrony* interaction, Bayesian inference confirmed that the additive model (A + S, Bayes factor Bf ± error = 170254 ± 1.26%, tested against

random effects model) was more plausible than the interaction model (A + S + A*S, Bf = 32747 ± 2.65%) given our data ($Bf_{additive}$ / $Bf_{interactive}$ = 5.20 ± 2.94%).

The *SSR component x synchrony* interaction originated from overall differences in the effect of synchrony (in-sync minus out-of-sync) on each SSR component that was most pronounced for pulse 2f components and virtually absent for flicker 1f responses (see *Figure 4a*). Specific contrasts confirmed that pulse 2f SSRs were more susceptible to synchrony effects than pulse 1f components (t(11) = 4.19, $P_{HB}$ < 0.05). Pulse 1f components in turn showed stronger modulation than flicker 1f components (t(11) = 5.02, $P_{HB}$ < 0.05). Lastly, pulse 2f components carried greater synchrony effects than flicker 1f components (t(11) = 7.83, $P_{HB}$ < 0.05).

*SSR inter-trial phase coherence*

ITC showed substantial variation with audio-visual *synchrony* (F(1,11) = 39.48, *P* < 0.001, $\eta^2$ = 0.113) and the allocation of *attention* (F(1,11) = 23.43, *P* < 0.001, $\eta^2$ = 0.139) but no effect of *SSR component* (F(2,22) = 2.24, *P* = 0.130, $\eta^2$ = 0.026) or *stimulus position* (F(1,11) < 1). A significant *SSR component x synchrony* interaction (F(2,22) = 16.16, $P_{GG}$ < 0.001, $\varepsilon_{GG}$ = 0.54, $\eta^2$ = 0.064) indicated that some SSR components were more susceptible to effects of audio-visual synchrony than others (*Figure 3b and d*). Remaining interaction terms, especially the *attention x synchrony* term (F(1,11) < 1), failed to indicate systematic effects (maximum F(1,11) = 2.80, *P* = 0.082, $\eta^2$ = 0.014 for the *attention x SSR component interaction*). Only the *synchrony x stimulus position* interaction was significant (F(1,11) = 5.05, *P* = 0.046) but explained a negligible amount of variance in the data ($\eta^2$ = 0.003) and was thus not further investigated.

Similar to SSR power, Bayesian inference supported the lack of an attention x synchrony interaction. Comparing additive (Bf = 176889 ± 1.82%) and interactive models (Bf = 32010 ± 2.98%) by means of the Bayesian approach showed evidence in favor of the additive model ($Bf_{additive}$ / $Bf_{interactive}$ = 5.53 ± 3.49%).

*Figure 4b* illustrates that the *SSR component x synchrony* interaction stemmed from greater synchrony effects (in-sync minus out-of-sync) on pulse 1f than flicker 1f components ($t(11) = 4.50$, $p_{HB} < 0.05$). Also, synchrony affected pulse 2f ITC more strongly than flicker 1f components ($t(11) = 5.06$, $p_{HB} < 0.05$). Effects between pulse 1f and 2f SSRs were comparable ($t(11) = 2.09$, $p_{HB} = 0.19$).

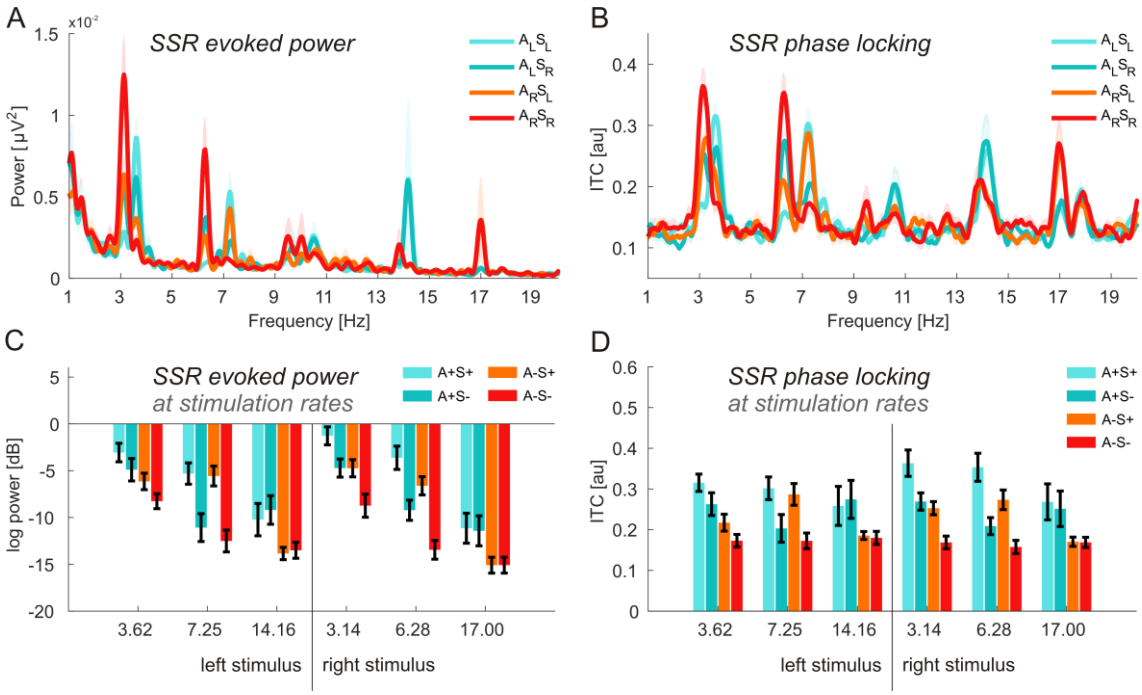


**Figure 3** SSRs by condition. (A) Condition-resolved grand-average power spectra. Colors code the four experimental conditions: Attend Left – Synchronous Left ($A_LS_L$), Attend Left – Synchronous Right ($A_LS_R$), Attend Right – Synchronous Left ($A_RS_L$) and Attend Right – Synchronous Right ($A_RS_R$). Shaded areas represent standard error of the mean (SEM). (B) Same as in (A) but for SSR inter-trial phase consistency (ITC; or "phase locking"). (C) Bars depict power (on a dB scale) at SSR component frequencies (in Hz, x-axis). Note that for the purpose of intuitive comparisons the colors now code according to the following condition labels: Attended – Synchronous (A+S+), Attended – Asynchronous (A+S-), Unattended – Synchronous (A-S+) and Unattended – Asynchronous. Error bars denote SEM. (D) Same as in C but for SSR phase locking.

*Attention- vs Synchrony-related gain effects*

As described in detail in the methods section, we computed indices that expressed SSR attention- and synchrony-related modulation of each SSR component. These modulation indices (AMIs and SMIs) allowed for a direct statistical comparison of the magnitude of attention and synchrony-related gain effects on SSR amplitudes and ITCs. As MI analyses assumed effects of attention and synchrony to be additive, further to the non-significant *attention x synchrony* interactions reported above, we estimated the plausibility of additive vs interactive models given our data by using a Bayesian approach. The estimated Bayes factors for SSR power and ITC (see above) indicated that both results were more than 5 times more likely under the additive than the interactive model.

Comparing modulation indices based on SSR amplitudes (*Figures 4c*) and SSR phase coherence (*Figure 4d*) revealed that across measures attention and synchrony modulated SSRs similarly (main effect *gain type*: $F(1,11) < 1$) while SSR inter-trial phase coherence indicated overall modulation more strongly (main effect *gain measure*: $F(1,11) = 74.30$, $P < 0.001$, $\eta^2 = 0.12$). Gain modulation varied systematically between SSR components (main effect *SSR component*: $F(2,22) = 4.32$, $P_{GG} < 0.05$, $\varepsilon_{GG} = 0.84$, $\eta^2 = 0.07$) and this effect further depended on whether attention or synchrony produced corresponding gains (*gain type x SSR component* interaction: $F(2,22) = 17.54$, $P_{GG} < 0.001$, $\varepsilon_{GG} = 0.49$, $\eta^2 = 0.16$). Remaining interaction terms were not significant ($F_{max} = 2.72$, $P_{min} = 0.09$).

For a closer investigation of the interaction, SSR amplitude and ITC-based modulation indices were pooled and tested against zero. Attention systematically modulated pulse 1f ($t(11) = 6.18$, $P_{HB} < 0.001$), and flicker 1f SSR components ($t(11) = 3.21$, $P_{HB} < 0.01$) but not pulse 2f responses ($t(11) = 2.26$, $P_{HB} > 0.05$). Synchrony, in turn, modulated pulse 1f ($t(11) = 4.44$, $P_{HB} < 0.005$) and pulse 2 SSR components ($t(11) = 7.52$, $P_{HB} < 0.001$) but not flicker 1f responses ($t(11) = 0.13$, $P_{HB} > 0.05$).

In direct comparisons, focusing on pulse 1f and pulse 2f components, systematic attention- and synchrony related gains were comparable for pulse 1f responses (t(11) = 0.98, $P = 0.34$), but differed for pulse 2f responses (t(11) = -3.73, $P_{HB} < 0.05$), with synchrony producing greater overall gain modulation ($M = 0.176$, $SEM = 0.025$) than attention ($M = 0.062$, $SEM = 0.028$). Finally, comparing the difference of attentional and synchrony-related modulation between pulse 1f ($M = 0.032$, $SEM = 0.035$) and 2f SSR components ($M = 0.115$, $SEM = 0.032$) showed that it was greater for pulse 2f responses (t(11) = 3.71, $P_{HB} < 0.005$) likely giving rise to the *gain type x SSR component* interaction.

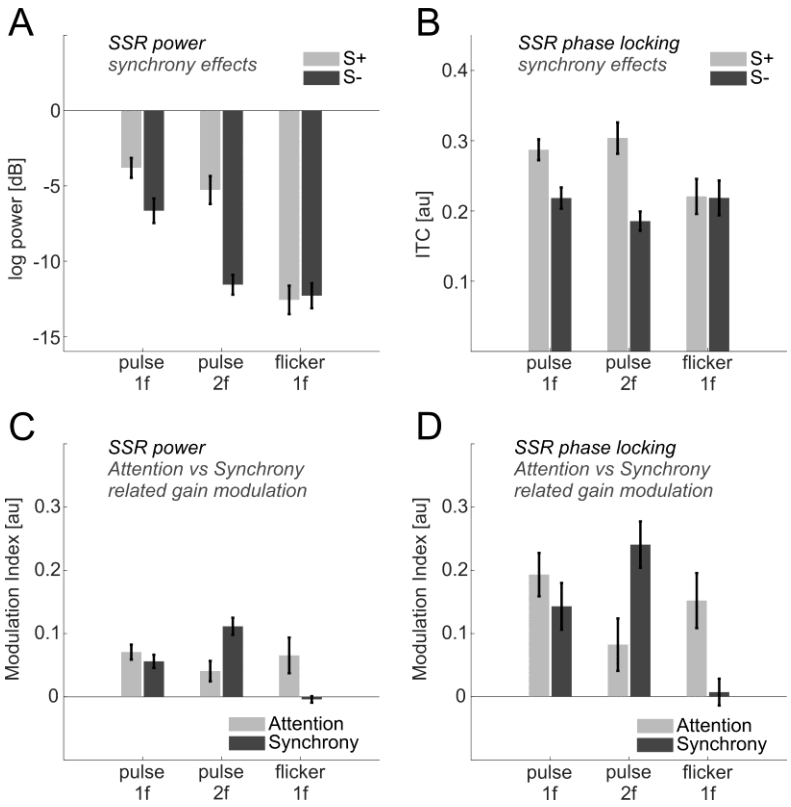


**Figure 4** Spotlight on synchrony effects and comparisons of attention- and synchrony-related gain. (A) SSR power for all three SSR components of interest (*pulse 1f*, *pulse 2f* and *flicker 1f*) separated by whether the driving visual stimulus was in sync (*S+*, light grey) or out-of-sync (*S-*, dark grey) with the tone. Note the synchrony-related gain in *pulse 1f* and *pulse 2f* components that is absent in *flicker 1f* responses. (B) Same as in (A) but for SSR inter-trial phase consistency (ITC or "phase locking"). (C) Bars indicate SSR power modulation by attention (light grey) and synchrony (dark grey) for all three SSR components of interest.

Note that the flicker 1f component is exclusively modulated by attention. (D) Same as in (C) but for SSR phase locking (ITC). Error bars in all plots display standard error of the mean.

## DISCUSSION

The role of top-down attention in multisensory binding and, conversely, bottom-up multisensory influences on attentional orienting have been studied largely independent of each other (Talsma et al., 2010). The present study was designed to bridge this gap. Specifically, we studied situations in which participants attended to the position of one of two pulsing and flickering stimuli providing it with a top-down processing advantage over the other stimulus. Additionally, a tone pulsing in synchrony with either the attended or unattended stimulus was introduced to produce a strong multisensory bottom-up bias in visual processing. EEG-recorded SSRs driven by stimulus flicker and pulsation allowed us to test whether and how spatial attention and audio-visual synchrony acted and possibly interacted to facilitate cortical visual stimulus processing.

We evaluated two commonly used SSR measures, evoked power and inter-trial phase coherence (ITC) to quantify modulations in stimulus processing. Both measures widely agree on patterns of effects and will thus be considered jointly as SSRs in the following. It is worth mentioning however that effects on ITC were more pronounced in general possibly qualifying it as the more sensitive measure for future research.

Briefly summarizing the results, spatial attention facilitated pulse- *and* flicker-driven SSRs. In contrast, synchrony specifically facilitated pulse-driven SSRs only with greater effects on pulse 2f components while leaving flicker 1f components unaffected. Most importantly, attention and synchrony produced independent additive gain effects. We confirmed that, given our data, an additive model of both influences was more plausible than assuming

552 interactive effects. These findings replicate results from an earlier study using a related
553 paradigm. In that study we tested concurrent influences of feature-based attention and
554 audio-visual synchrony on two spatially super-imposed Gabor patches (Keitel and Müller,
555 2015).
556
557 **Spatial attention facilitates processing of all stimulus aspects**
558 The described effects of spatial attention are in line with numerous studies demonstrating
559 sensory gain effects on SSR-indexed cortical visual processing (Müller et al., 1998a; Störmer
560 et al., 2014; Walter et al., 2015). Interestingly, our results show that spatial attention has
561 comparable effects on SSRs driven by two different but simultaneous rhythmic changes in
562 stimulus appearance: a relatively fast on-off flicker and a slow-paced sinusoidal spatial
563 frequency modulation. These results support the notion that spatial attention prioritizes all
564 aspects of sensory information within its focus (Andersen et al., 2008; Keitel and Müller,
565 2015) as is central to psychological (Treisman and Gelade, 1980; Wolfe, 1994) and
566 neurophysiological models of attention (Bundesen et al., 2015; Reynolds and Heeger, 2009) .
567 Note that participants performed better in the visual detection task when they attended to
568 the left stimulus. This effect could be due to a left-hemifield advantage as has been
569 described previously for rapid serial visual presentation paradigms (Śmigasiewicz et al.,
570 2014; Verleger et al., 2011).Nevertheless, SSR analyses did not show differences in stimulus
571 processing between left and right stimulus positions. It is therefore unlikely that an
572 imbalance in task difficulty influenced our neural measures of stimulus processing.
573
574 **Synchrony selectively facilitates stimulus aspects relevant for multisensory integration**
575 Facilitation of visual processing by audio-visual synchrony has largely been studied using
576 transient stimuli (Busse et al., 2005; Talsma et al., 2009). So far, only a few studies have
577 demonstrated similar effects while employing dynamic ongoing stimulation (Keitel and

Müller, 2015; Nozaradan et al., 2012; Schall et al., 2009). Prolonged exposure to synchronous sensory input, however, can be a vital factor in multisensory integration because it improves the estimate of temporal correlations between visual and auditory stimuli over time (Parise and Ernst, 2016). This is important in situations with multiple concurrent stimuli as studied here because even unrelated visual and auditory events can occur simultaneously on occasion.

Our study corroborates this role of ongoing audio-visual synchrony. Interestingly, synchrony-related gain effects were thereby restricted to SSR components that reflected stimulus pulsing (pulse 1f and 2f), i.e. those rhythmic modulations that produced the impression of synchrony.

With regard to the pulse 1f component, which was inherently linked to the auditory signal, we expected and observed an effect of synchrony. The visual stimulus dynamics were, in fact, either matched or not matched with the spectral profile of the auditory stimulus, thus providing either maximal or minimal temporal correlation.

Less intuitively, the pulse 2f component showed greater synchrony modulations than the pulse 1f SSR. In line with Keitel et al. (2015), who employed a stimulus with the same properties, the pulse 2f modulation was accounted for by the transients elicited by the stimulus at twice the stimulation frequencies during maximum up- and down-slopes of the sinusoidal modulation, or alternatively its extrema, i.e. peaks and troughs.

A possible neural process underlying synchrony-related modulation of both pulse-driven components is cross-modal phase resetting. It has been considered as the primary channel for multisensory interactions between early sensory cortices (Lakatos et al., 2009; van Atteveldt et al., 2014). Unlike neurons in higher order cortices, which are intrinsically multisensory ( and hence sensitive to combined multisensory information) neurons in early

sensory cortices are primarily sensory specific, but crucially sensitive to temporal information conveyed also by non-specific modalities. As underlined by Lakatos et al. (2008), appropriately timed inputs in one modality can aid in processing a stimulus presented in a different modality. In our case these connections may support phase stability of visual SSRs by providing a cross-modal temporal scaffold (Kayser et al., 2010; Lakatos et al., 2009). As a consequence, the temporal precision of cortical stimulus representations increases, which awards them a processing advantage (Chennu et al., 2009).

Although our results are broadly in line with Nozaradan et al. (2012), who firstly measured synchrony effects on SSRs, it is worth noting a discrepancy: In contrast to our findings the authors reported an effect on a flicker-driven SSR with a frequency of 10 Hz, while establishing synchrony with auditory beats at either 2.1 or 2.4 Hz. These differences may be accounted for by the fact that the authors presented only one visual stimulus centrally. In this setup, gain effects cannot not unambiguously be ascribed to synchrony, or alternatively, altered attentional demands between synchronous and asynchronous conditions.

**Facilitatory effects of spatial attention and synchrony add up**

We found that attended and unattended stimulus experienced comparable gain through synchrony. Vice versa, synchronous and asynchronous stimuli were similarly facilitated when their position was attended. Remarkably, these findings point towards a dual reign of attention and audio-visual synchrony in early sensory cortices, suggesting that both influences can work independently and in parallel. This result seemingly contradicts previous studies (Alsius et al., 2005; Fairhall and Macaluso, 2009) that showed an interdependence between attention and multisensory interactions; however, this discrepancy can be reconciled by examining the experimental paradigm employed in the current study.

Unlike previous experiments, in which mutual input from different senses was essential for successful behavioral performance, our task (the detection of luminance changes) was functionally independent from audio-visual synchrony. Here, visual task performance did not require additional auditory information. Our paradigm might thus have promoted the independence between attention and audio-visual interactions triggering two concurrent, but distinct processes: On the one hand, performing the detection task required a sustained goal-driven deployment of spatial attention, while on the other hand merging the audio-visual signals was most likely a stimulus-driven process, triggered by the high temporal correlation between audio-visual signals.

For these two processes to co-occur independently, the involvement of distinct neural pathways can be assumed. Various aspects of attention and its influence on perception have been related to a number of anatomical networks (Shipp, 2004). Those include a dorsal fronto-parietal network, which entails the intra-parietal sulcus (IPS) in posterior parietal cortex, a portion of the precentral supplemental motor area, the so-called frontal eye fields (FEF) and early sensory areas, such as visual cortex (Corbetta and Shulman, 2002). It is one of the most comprehensively investigated cortical network implicated in the control of attention (Corbetta et al., 1998) and was likely involved in deploying the attentional resources necessary to perform in our behavioral task.

On the other side, auditory influences on visual processing could have been conveyed by two candidate routes that have been suggested as a results of earlier invasive electrophysiological and anatomical studies in the animal brain: (1) feed-forward projections between thalamus and early sensory cortices (Cappe et al., 2009), (2) lateral projections between early sensory cortices (Falchier et al., 2002). From our data alone, we

cannot say which pathway was critical in the investigated situation but both of them are anatomically distinct from the attention network and could explain our results.

It should be mentioned that our data analyses and interpretation of results depend on the implicit assumption that attention and synchrony effects follow similar time courses and, once established, remain constant through the course of trials. At least for, spatial attention we know that gain effects reach asymptote after ~500 msec and keep level for several seconds (Müller et al., 1998b). A time course for synchrony-related gain instead has not been established yet. This uncertainty notwithstanding, we restricted our analyses to a period starting 0.5 sec after stimulus onset. We were confident that this time frame would allow for enough audio-visual coincidence to be detected to establish synchrony. Nevertheless, the comparison of temporal profiles of attention- and synchrony related gain remains an interesting subject for future studies.

As a final remark, Talsma et al. (2010) suggested that bottom-up multisensory integration benefits a given stimulus the most when competition within one sensory modality is high, e.g. when the visual field is cluttered. Our situation promotes only minimal competition with one stimulus presented to each hemifield. Inter-hemispheric competition is introduced relatively late in the visual processing hierarchy (Schwartz et al., 2007). Moreover, attentional resources seem to split more readily between than within visual hemifields (Franconeri et al., 2012; Störmer et al., 2013; Walter et al., 2015). It would thus be interesting to test how synchrony-related gain effects vary with the amount of competition by placing more than one stimulus within each visual hemifield.

679 **Conclusion**

680 We investigated the concurrent effects of spatial attention and audio-visual synchrony on
681 early cortical visual stimulus processing. Our paradigm allowed us to test both influences in
682 isolation as well as their combined effects. We found that attention-related and synchrony-
683 related facilitation add up when an audio-visual synchronous stimulus is attended. Further,
684 attention facilitated pulse- and flicker-driven neural responses while synchrony only
685 targeted pulse-driven responses, i.e. those coding stimulus dynamics that were relevant for
686 multisensory integration. Consequentially, the present results favor an account in which
687 spatial attention and audio-visual synchrony convey their influences independently via
688 different neural processes and possibly along different neural pathways. At least for
689 situations similar to the one studied here, this finding implies that facilitation through
690 synchrony cannot simply be modelled as a sustained attraction of spatial attention.

691

692 **Acknowledgments**

693 Work was supported by the Deutsche Forschungsgemeinschaft (grant no. MU972/21-1).
694 Data presented here were recorded at the Institut für Psychologie, Universität Leipzig. The
695 authors appreciate the assistance of Renate Zahn in data collection. Experimental
696 stimulation was realized using Cogent Graphics developed by John Romaya at the Laboratory
697 of Neurobiology at the Wellcome Department of Imaging Neuroscience, University College
698 London.

699

700 **Conflict of interest:** The authors declare that they have no conflict of interest.

701

702

## References


Alsius, A., Navarra, J., Campbell, R., Soto-Faraco, S., 2005. Audiovisual Integration of Speech Falters under High Attention Demands 15, 839–843. doi:10.1016/j.cub.2005.03.046

Alvarez, G.A., Oliva, A., 2009. Spatial ensemble statistics are efficient codes that can be represented with reduced attention. Proc. Natl. Acad. Sci. U.S.A. 106, 7345–7350. doi:10.1073/pnas.0808981106

Andersen, S.K., Hillyard, S.A., Müller, M.M., 2008. Attention Facilitates Multiple Stimulus Features in Parallel in Human Visual Cortex 18, 1006–1009. doi:10.1016/j.cub.2008.06.030

Bundesen, C., Vangkilde, S., Petersen, A., 2015. Recent developments in a computational theory of visual attention (TVA). Vision Research 116, 210–218. doi:10.1016/j.visres.2014.11.005

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 102, 18751–18756. doi:10.1073/pnas.0507704102

Cappe, C., Rouiller, E.M., Barone, P., 2009. Multisensory anatomical pathways. Hearing Research 258, 28–36. doi:10.1016/j.heares.2009.04.017

Chennu, S., Craston, P., Wyble, B., Bowman, H., 2009. Attention Increases the Temporal Precision of Conscious Perception: Verifying the Neural-ST2 Model. PLOS Computational Biology 5, e1000576. doi:10.1371/journal.pcbi.1000576

Cohen, M. X. (2014). Analyzing neural time series data: theory and practice. MIT Press.

Colonius, H., Diederich, A., 2012. Focused attention vs. crossmodal signals paradigm: deriving predictions from the time-window-of-integration model. Frontiers in Integrative Neuroscience 6. doi:10.3389/fnint.2012.00062

Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber, M.R., Petersen, S.E., Raichle, M.E., Van Essen, D.C., Shulman, G.L., 1998. A common network of functional areas for attention and eye movements. Neuron 21, 761–773. doi:10.1016/S0896-6273(00)80593-0

Corbetta, M., Shulman, G.L., 2002. CONTROL OF GOAL-DIRECTED AND STIMULUS-DRIVEN ATTENTION IN THE BRAIN. Nature Reviews Neuroscience 3, 215–229. doi:10.1038/nrn755

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009

Donohue, S.E., Green, J.J., Woldorff, M.G., 2015. The effects of attention on the temporal integration of multisensory stimuli. Frontiers in Integrative Neuroscience 9. doi:10.3389/fnint.2015.00032

Fairhall, S.L., Macaluso, E., 2009. Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. European Journal of Neuroscience 29, 1247–1257. doi:10.1111/j.1460-9568.2009.06688.x

Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. Journal of Neuroscience 22, 5749–5759. doi:10.1002/(SICI)1096-9861(19981026)400:3<417::AID-CNE10>3.0.CO;2-O

Franconeri, S.L., Scimeca, J.M., Roth, J.C., Helseth, S.A., Kahn, L.E., 2012. Flexible visual processing of spatial relationships. Cognition 122, 210–227.

Fujisaki, W., Nishida, S., 2005. Temporal frequency characteristics of synchrony–asynchrony discrimination of audio-visual signals. Exp Brain Res 166, 455–464. doi:10.1007/s00221-005-2385-8

Giani, A.S., Ortiz, E., Belardinelli, P., Kleiner, M., Preissl, H., Noppeney, U., 2012. Steady-state responses in MEG demonstrate information integration within but not across the

753 auditory and visual senses. NeuroImage 60, 1478–1489.
754 doi:10.1016/j.neuroimage.2012.01.114
755 Greenhouse, S.W., Geisser, S., 1959. On methods in the analysis of profile data.
756 Psychometrika 24, 95–112. doi:10.1007/BF02289823
757 Herrmann, B., Schlichting, N., Obleser, J., 2014. Dynamic Range Adaptation to Spectral
758 Stimulus Statistics in Human Auditory Cortex. Journal of Neuroscience 34, 327–331.
759 doi:10.1523/JNEUROSCI.3974-13.2014
760 Hertz, U., Amedi, A., 2010. Disentangling unisensory and multisensory components in
761 audiovisual integration using a novel multifrequency fMRI spectral analysis. NeuroImage
762 52, 617–632. doi:10.1016/j.neuroimage.2010.04.186
763 Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scandinavian journal
764 of statistics. doi:10.2307/4615733
765 Jenkins, J., Rhone, A.E., Idsardi, W.J., Simon, J.Z., Poeppel, D., 2011. The Elicitation of
766 Audiovisual Steady-State Responses: Multi-Sensory Signal Congruity and Phase Effects.
767 Brain Topography 24, 134–148. doi:10.1007/s10548-011-0174-1
768 Kashiwase, Y., Matsumiya, K., Kuriki, I., Shioiri, S., 2012. Time Courses of Attentional
769 Modulation in Neural Amplification and Synchronization Measured with Steady-state
770 Visual-evoked Potentials. Journal of Cognitive Neuroscience 24, 1779–1793.
771 doi:10.1162/jocn_a_00212
772 Kayser, C., Logothetis, N.K., Panzeri, S., 2010. Visual enhancement of the information
773 representation in auditory cortex. Curr. Biol. 20, 19–24. doi:10.1016/j.cub.2009.10.068
774 Keitel, C., Andersen, S.K., Quigley, C., Müller, M.M., 2013. Independent Effects of Attentional
775 Gain Control and Competitive Interactions on Visual Stimulus Processing. Cerebral
776 Cortex 23, 940–946. doi:10.1093/cercor/bhs084
777 Keitel, C., Müller, M.M., 2015. Audio-visual synchrony and feature-selective attention co-
778 amplify early visual processing. Exp Brain Res 234, 1221–1231. doi:10.1007/s00221-015-
779 4392-8
780 Kim, Y.J., Grabowecky, M., Paller, K.A., Muthu, K., Suzuki, S., 2007. Attention induces
781 synchronization-based response gain in steady-state visual evoked potentials. Nature
782 Neuroscience 10, 117–125. doi:10.1038/nn1821
783 Kim, Y.J., Grabowecky, M., Paller, K.A., Suzuki, S., 2011. Differential Roles of Frequency-
784 following and Frequency-doubling Visual Responses Revealed by Evoked Neural
785 Harmonics. Journal of Cognitive Neuroscience 23, 1875–1886.
786 doi:10.1162/jocn.2010.21536
787 Krause, H., Schneider, T.R., Engel, A.K., Senkowski, D., 2012. Capture of visual attention
788 interferes with multisensory speech processing. Frontiers in Integrative Neuroscience 6.
789 doi:10.3389/fnint.2012.00067
790 Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of
791 Neuronal Oscillations as a Mechanism of Attentional Selection. Science 320, 110–113.
792 doi:10.1126/science.1154735
793 Lakatos, P., O'Connell, M.N., Barczak, A., Mills, A., Javitt, D.C., Schroeder, C.E., 2009. The
794 Leading Sense: Supramodal Control of Neurophysiological Context by Attention. Neuron
795 64, 419–430. doi:10.1016/j.neuron.2009.10.014
796 Lee, S., 1999. Visual Form Created Solely from Temporal Structure. Science 284, 1165–1168.
797 doi:10.1126/science.284.5417.1165
798 Leek, M.R., 2001. Adaptive procedures in psychophysical research. Perception &
799 Psychophysics 63, 1279–1292. doi:10.3758/BF03194543
800 Morey, R.D., Rouder, J.N., Jamil, T., (2015) BayesFactor: Computation of Bayes Factors for
801 Common Designs. R package. URL http://bayesfactorpcl.r-forge.r-project.org/
802 Morgan, S.T., Hansen, J.C., Hillyard, S.A., 1996. Selective attention to stimulus location
803 modulates the steady-state visual evoked potential. Proceedings of the National

804 Academy of Sciences 93, 4770–4774. doi:10.1073/pnas.93.10.4770
805 Müller, M.M., Hubner, R., 2002. Can the Spotlight of Attention be Shaped Like A Doughnut?
806 Psychological Science 13, 119–124. doi:10.1111/j.0956-7976.2002.t01-1-.x
807 Müller, M.M., Picton, T.W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W.A., Hillyard, S.A.,
808 1998a. Effects of spatial selective attention on the steady-state visual evoked potential
809 in the 20–28 Hz range. Cognitive Brain Research 6, 249–261. doi:10.1016/S0926-
810 6410(97)00036-0
811 Müller, M.M., Teder-Sälejärvi, W., Hillyard, S.A., 1998b. The time course of cortical
812 facilitation during cued shifts of spatial attention. Nature Neuroscience 1, 631–634.
813 doi:10.1038/2865
814 Nolan, H., Whelan, R., Reilly, R.B., 2010. FASTER: Fully Automated Statistical Thresholding for
815 EEG artifact Rejection. Journal of Neuroscience Methods 192, 152–162.
816 doi:10.1016/j.jneumeth.2010.07.015
817 Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottereau, B.R., Rossion, B., 2015. The steady-
818 state visual evoked potential in vision research: A review. Journal of Vision 15, 4–4.
819 doi:10.1167/15.6.4
820 Nozaradan, S., Peretz, I., Mouraux, A., 2012. Steady-state evoked potentials as an index of
821 multisensory temporal binding. NeuroImage 60, 21–28.
822 doi:10.1016/j.neuroimage.2011.11.065
823 Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open source software
824 for advanced analysis of MEG, EEG, and invasive electrophysiological data.
825 Computational Intelligence and Neuroscience 2011, 156869–9.
826 doi:10.1155/2011/156869
827 Parise, C.V., 2012. Signal compatibility as a modulatory factor for audiovisual multisensory
828 integration.
829 Parise, C.V., Ernst, M.O., 2016. Correlation detection as a general mechanism for
830 multisensory integration. Nature Communications 7, 11543. doi:10.1038/ncomms11543
831 Pastor, M.A., Valencia, M., Artieda, J., Alegre, M., Masdeu, J.C., 2007. Topography of cortical
832 activation differs for fundamental and harmonic frequencies of the steady-state visual-
833 evoked responses. An EEG and PET H215O study. Cerebral Cortex 17, 1899–1905.
834 doi:10.1093/cercor/bhl098
835 Porcu, E., Keitel, C., Müller, M.M., 2014. Visual, auditory and tactile stimuli compete for early
836 sensory processing capacities within but not between senses. NeuroImage 97, 224–235.
837 doi:10.1016/j.neuroimage.2014.04.024
838 Porcu, E., Keitel, C., Müller, M.M., 2013. Concurrent visual and tactile steady-state evoked
839 potentials index allocation of inter-modal attention: A frequency-tagging study.
840 Neuroscience Letters 556, 113–117. doi:10.1016/j.neulet.2013.09.068
841 Posner, M.I., Snyder, C.R., Davidson, B.J., 1980. Attention and the detection of signals.
842 Journal of Experimental Psychology: General 109, 160–174. doi:10.1037/0096-
843 3445.109.2.160
844 Quigley, C., Müller, M.M., 2014. Feature-Selective Attention in Healthy Old Age: A Selective
845 Decline in Selective Attention? Journal of Neuroscience 34, 2471–2476.
846 doi:10.1523/JNEUROSCI.2718-13.2014
847 R Core Team (2013). R: A language and environment for statistical computing. R Foundation
848 for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-
849 project.org/.
850 Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic
851 Fields in Science and Medicine. (1989).
852 Reynolds, J.H., Heeger, D.J., 2009. The Normalization Model of Attention. Neuron 61, 168–
853 185. doi:10.1016/j.neuron.2009.01.002
854 Rouder, J.N., Morey, R.D., Speckman, P.L., Province, J.M., 2012. Default Bayes factors for

855 ANOVA designs. Journal of Mathematical Psychology 56, 356–374.
856 Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G., 2009. Bayesian t tests for
857 accepting and rejecting the null hypothesis. Psychonomic Bulletin & Review 16, 225–
858 237. doi:10.3758/PBR.16.2.225
859 Ruhnau, P., Keitel, C., Lithari, C., Weisz, N., Neuling, T., 2016. Flicker-Driven Responses in
860 Visual Cortex Change during Matched-Frequency Transcranial Alternating Current
861 Stimulation. Frontiers in Human Neuroscience 10, 440. doi:10.3389/fnhum.2016.00184
862 Schall, S., Quigley, C., Onat, S., König, P., 2009. Visual stimulus locking of EEG is modulated by
863 temporal congruency of auditory stimuli. Exp Brain Res 198, 137–151.
864 doi:10.1007/s00221-009-1867-5
865 Schönbrodt, F.D., Wagenmakers, E.J., 2015. Sequential hypothesis testing with Bayes factors:
866 Efficiently testing mean differences. Psychological ….
867 Schwartz, O., Hsu, A., Dayan, P., 2007. Space and time in visual context. Nature Reviews
868 Neuroscience 8, 522–535. doi:10.1038/nrn2155
869 Shipp, S., 2004. The brain circuitry of attention. Trends in Cognitive Sciences 8, 223–230.
870 doi:10.1016/j.tics.2004.03.004
871 Störmer, V., Cavanagh, P., Alvarez, G., 2013. The profile of multifocal attention: surround-
872 suppression between and within hemifields. Journal of Vision 13, 1283–1283.
873 doi:10.1167/13.9.1283
874 Störmer, V.S., Alvarez, G.A., Cavanagh, P., 2014. Within-Hemifield Competition in Early Visual
875 Areas Limits the Ability to Track Multiple Objects with Attention. Journal of
876 Neuroscience 34, 11526–11533. doi:10.1523/JNEUROSCI.0980-14.2014
877 Śmigasiewicz, K., Asanowicz, D., Westphal, N., Verleger, R., 2014. Bias for the Left Visual
878 Field in Rapid Serial Visual Presentation: Effects of Additional Salient Cues Suggest a
879 Critical Role of Attention. Journal of Cognitive Neuroscience 27, 266–279.
880 doi:10.1162/jocn_a_00714
881 Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay
882 between attention and multisensory integration. Trends in Cognitive Sciences 14, 400–
883 410. doi:10.1016/j.tics.2010.06.008
884 Talsma, D., Senkowski, D., Woldorff, M.G., 2009. Intermodal attention affects the processing
885 of the temporal alignment of audiovisual stimuli. Exp Brain Res 198, 313–328.
886 doi:10.1007/s00221-009-1858-6
887 Treisman, A.M., Gelade, G., 1980. A feature-integration theory of attention. Cognitive
888 Psychology 12, 97–136. doi:10.1016/0010-0285(80)90005-5
889 van Atteveldt, N., Murray, M.M., Thut, G., Schroeder, C.E., 2014. Multisensory Integration:
890 Flexible Use of General Operations. Neuron 81, 1240–1253.
891 doi:10.1016/j.neuron.2014.02.044
892 Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., Theeuwes, J., 2008. Pip and pop:
893 Nonspatial auditory signals improve spatial visual search. Journal of Experimental
894 Psychology: Human Perception and Performance 34, 1053–1065. doi:10.1037/0096-
895 1523.34.5.1053
896 Verleger, R., Śmigasiewicz, K., Möller, F., 2011. Mechanisms underlying the left visual-field
897 advantage in the dual stream RSVP task: Evidence from N2pc, P3, and distractor-evoked
898 VEPs. Psychophysiology 48, 1096–1106. doi:10.1111/j.1469-8986.2011.01176.x
899 Walter, S., Keitel, C., Müller, M.M., 2015. Sustained Splits of Attention within versus across
900 Visual Hemifields Produce Distinct Spatial Gain Profiles. Journal of Cognitive
901 Neuroscience 28, 111–124. doi:10.1162/jocn_a_00883
902 Walter, S., Quigley, C., Andersen, S.K., Mueller, M.M., 2012. Effects of overt and covert
903 attention on the steady-state visual evoked potential. Neuroscience Letters 519, 37–41.
904 doi:10.1016/j.neulet.2012.05.011
905 Werner, S., Noppeney, U., 2011. The contributions of transient and sustained response

codes to audiovisual integration. Cereb. Cortex 21, 920–931. doi:10.1093/cercor/bhq161

Wolfe, J.M., 1994. Guided Search 2.0 A revised model of visual search. Psychonomic Bulletin & Review 1, 202–238. doi:10.3758/BF03200774

Wolfe, J.M., Cave, K.R., Franzel, S.L., 1989. Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance 15, 419–433. doi:10.1037/0096-1523.15.3.419