

# Sustained selective intermodal attention modulates processing of language-like stimuli

Christian Keitel · Erich Schröger · Katja Saupe ·  
Matthias M. Müller

Received: 13 October 2010 / Accepted: 28 March 2011 / Published online: 19 April 2011  
© Springer-Verlag 2011

**Abstract** Intermodal attention (IA) is assumed to allocate limited neural processing resources to input from one specific sensory modality. We investigated effects of sustained IA on the amplitude of a 40-Hz auditory (ASSR) and a 4.3-Hz visual steady-state response (VSSR). To this end, we concurrently presented amplitude-modulated multi-speech babble and a stream of nonsense letter sets to elicit the respective brain responses. Subjects were cued trialwise to selectively attend to one of the streams for several seconds where they had to perform a lexical decision task on occasionally occurring words and pseudowords. Attention to the auditory stream led to greater ASSR amplitudes than attention to the visual stream. Vice versa, the VSSR amplitude was greater when the visual stream was attended. We demonstrate that IA research by means of frequency tagging can be extended to complex stimuli as used in the current study. Furthermore, we show not only that IA selectively modulates processing of concurrent multisensory input but that this modulation occurs during trial-by-trial cueing of IA. The use of frequency tagging may be suitable to study the role of IA in more naturalistic setups that comprise a larger number of multisensory signals.

**Keywords** Human EEG · Intermodal attention · Sustained attention · VSSR · ASSR · Speech

## Introduction

The human perceptual system incorporates at least two means to deal with concurrent multisensory input. On the one hand, it is able to integrate information across sensory modalities (cf. Calvert and Thesen 2004). On the other hand, if different sensory modalities provide conflicting information, a crossmodal competition for further processing arises (Ghatan et al. 1998; Berman and Colby 2002; Shomstein and Yantis 2004). It is very likely that an intermodal attention (IA) mechanism resolves competition in favour of input from the task-relevant modality. The effect of IA on event-related potentials (ERPs) has been observed by comparing the processing of a stimulus that is presented in the attended modality to its processing when attention is directed to another modality (Alho et al. 1992, 1994; Woods et al. 1992; de Ruiter et al. 1998; Eimer and Schröger 1998; Talsma and Kok 2001, 2002).

Recently, frequency tagging has been introduced to overcome fundamental constraints of the ERP approach such as the incapability to separate responses to simultaneously presented stimuli. A frequency-tagged stimulus elicits a continuous electrophysiological brain response, the steady-state response (SSR) that has the same temporal frequency as the driving stimulus and yields a measure of stimulus processing over time (Regan 1989). Two simultaneously presented stimuli which are tagged with different frequencies will drive distinct SSRs.

In the visual domain, it is well established that attention to a flickering stimulus increases the amplitude of its corresponding visual SSR (VSSR) relative to responses to ignored stimuli (Morgan et al. 1996; Müller et al. 1998; Toffanin et al. 2009; Quigley et al. 2010; Andersen et al. 2011). More ambiguous results with regard to attention effects have been reported for auditory SSRs (ASSRs)

---

C. Keitel · E. Schröger · K. Saupe · M. M. Müller (✉)  
Institut für Psychologie, Universität Leipzig,  
Seeburgstraße 14-20, 04103 Leipzig, Germany  
e-mail: m.mueller@rz.uni-leipzig.de

(Linden et al. 1987; Ross et al. 2004; Bidet-Caulet et al. 2007; Skosnik et al. 2007; Müller et al. 2009). So far, a higher ASSR amplitude has been observed in studies only when subjects had to attend to the frequency of the amplitude modulation (AM) of the auditory signal that produces the ASSR (Linden et al. 1987; Bidet-Caulet et al. 2007).

As of today, few studies have employed SSRs to investigate the effects of selective IA on stimulus processing (e.g. Ross et al. 2004; Talsma et al. 2006; Saupé et al. 2009a). Even fewer have tapped the full advantages of SSRs by delivering frequency-tagged stimuli to different modalities and obtain simultaneous measures of sustained multisensory processing (Saupé et al. 2009b; de Jong et al. 2010; Gander et al. 2010). In line with the IA account, Saupé et al. (2009b) as well as Gander et al. (2010) found a systematic modality-specific enhancement of SSR amplitudes in an audiovisual design when the corresponding modality was attended and the other modality was ignored. Interestingly, these findings were contrasted by results from de Jong et al. (2010). They studied intermodal selective and divided attention during detection and discrimination performance and reported the greatest effect of attention on VSSR amplitude in a divided attention condition and no effect on ASSR amplitude at all. In all of these studies, stimulus configurations were sine tones and geometric objects or single letters. Subjects were cued prior to each block which modality they had to attend and were instructed to perform tasks that varied with the task-relevant modality. The blockwise switching between different tasks (Saupé et al. 2009b; de Jong et al. 2010; Gander et al. 2010) or especially differing task demands (de Jong et al. 2010; Gander et al. 2010) might have led to varying levels of arousal between experimental conditions. Griskova et al. (2007) found that higher arousal reduced ASSR amplitudes. Thus, in the above-mentioned studies, the effects of attention and arousal on ASSR amplitudes might have been confounded.

The present study was intended to overcome these limitations. We aimed to shed further light on the question of whether stimulus processing can be modulated by sustained selective IA. However, we introduced substantial changes to previously employed experimental protocols. Our paradigm was the first to cue subjects trial-by-trial which modality to attend to. SSR amplitude differences between attend and ignore conditions could thus hardly be explained by arousal effects and would be more likely related to selective attention. As ERP research revealed that block-cueing and trial-by-trial cueing do not necessarily yield the same effects (Näätänen et al. 2002), it had to be determined whether IA influences SSRs at all.

We frequency-tagged a new class of concurrently presented language-like auditory and visual stimuli to elicit

corresponding ASSRs and VSSRs. Multi-speech babble served as the auditory stimulus and random letter sets comprised the visual stream (see “Methods”). Importantly, these stimuli enabled a uniform lexical decision task in both modalities. Subjects were instructed to discriminate occasionally occurring words from pseudowords in the to-be-attended stream and respond to words only.

Our results clearly show facilitated stimulus processing for attended auditory or visual streams compared with when they were unattended.

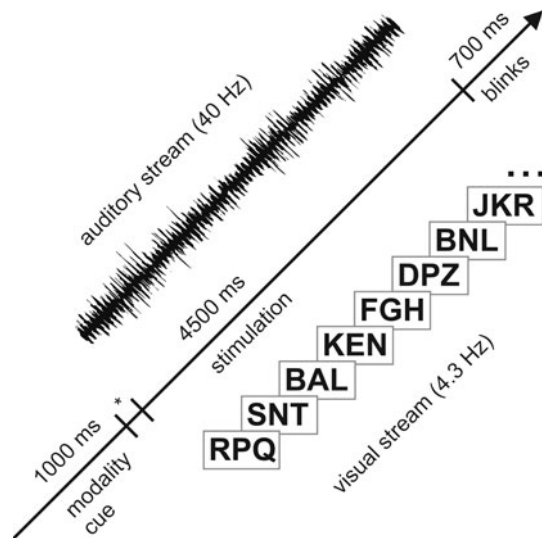
## Methods

### Subjects

Sixteen subjects participated in the experiment. Three subjects were excluded from further data analysis as one subject did not show ASSR amplitudes above general noise level, data of another subject were contaminated with muscle artefacts and a third subject could not perform the visual task. The remaining subjects (mean age = 24.6 years, SD = 4.1, seven women) had normal or corrected-to-normal vision and normal hearing. None of them reported a history of a neurological disease or injury. According to the Declaration of Helsinki, written informed consent was obtained from each subject prior to the experiment. Subjects received course credits or monetary payment for participation.

### Stimuli and procedure

Subjects were seated comfortably in an acoustically dampened and electromagnetically shielded chamber and were instructed to fixate a red dot ( $0.4^\circ$  of visual angle) at the centre of a monitor (refresh rate = 60 Hz, distance = 1.4 m). Each trial started with the presentation of a cue for 1,000 ms followed by a short period of 200 ms in which only the fixation dot was shown (see Fig. 1). The cue indicated which one of the two stimulus streams (auditory stream: unintelligible multi-speech-babble; visual stream: random letter strings; see below) subjects had to attend to. A pulsating tone (frequency-modulated sine with a centre frequency of 500 Hz) served as a cue to attend to the auditory stream. Attention to the visual stream was cued by a red pulsating circle around the fixation dot. Subsequently, auditory and visual stimulus streams started simultaneously and were concurrently presented for 4,500 ms. Finally, to minimise eye movements and blinks during stimulation, the fixation point was replaced by an ‘X’ for 700 ms, which indicated that subjects were now allowed to blink before the next trial started (Fig. 1).



**Fig. 1** Schematic illustration of one trial. 40-Hz amplitude-modulated multi-speech babble served as the auditory stimulus. The visual stream of letter sets was presented with a rate of 4.3 Hz. In both streams words/pseudowords were occasionally included (not shown for the auditory stream, note the presentation of the German word ‘BAL’-‘KEN’ [transl. ‘beam’] in the RSVP). At the beginning of each trial, a modality cue was presented for 1,000 ms followed by a 200-ms period (*Asterisk*) during which only the fixation point was shown. Afterwards, auditory and visual stimulus streams were presented concurrently for 4,500 ms. Trials ended with a 700-ms period during which subjects were allowed to blink

Subjects were instructed to perform a lexical decision task, namely to discriminate between words and pseudowords in the attended stream. To this end, words/pseudowords were embedded in both stimulus streams in 50% of all trials and up to two times in one trial with a minimum interval of 1,000 ms between subsequent onsets. Subjects responded to word occurrences via button press. The responding hand was changed halfway through the experiment with the starting hand counterbalanced across subjects.

For each condition—*attend vision* and *attend audition*—180 trials were presented in 6 consecutive blocks. Trial presentation order was randomised across conditions. Subjects started individual blocks by pressing a button. Prior to the experiment, they were familiarised with the word/pseudoword stimuli and the task for at least one block. After each training and experimental block, subjects received feedback regarding average hit rate and reaction time.

Visual stimuli were successively shown sets of three random black letters each (selected from the German alphabet with vowels excluded, font size 3.5°) on a white rectangle. The obtained rapid serial visual presentation (RSVP) with a rate of 4.3 Hz subtended a visual field of  $6 \times 4^\circ$  in the centre of the screen in front of a grey background. Identical letters were not allowed to be

presented at the same spatial location in consecutive letter sets. Each presentation cycle consisted of 14 frames (233 ms) during which one set was only shown for the first seven frames (117 ms). This particular rate of presentation was chosen on the basis of pilot behavioural data that had been collected prior to the experiment as it provided an appropriate task difficulty in yielding an average accuracy of about 80% (for calculation of accuracy see below).

The auditory stream was presented binaurally via headphones with an intensity of 65 dB sound pressure level. It consisted of an amplitude-modulated multi-speech babble (MSB) that was generated from samples of eight different speakers (four women). Each speaker provided four different 20-s samples extracted randomly from continuous speech. These samples were low-pass filtered with a cut-off frequency of 4,000 Hz and normalised in amplitude by a root mean square procedure to match intensities. The subsequent combination of individual samples yielded a signal with 32 simultaneously speaking voices. This signal was then multiplied by a 40-Hz sine, which produced an amplitude modulation with a depth of 100%.

For the lexical decision task, we chose word stimuli from a fixed set of two-syllabic German nouns (cf. Muller-Gass et al. 2007, see Tab. 1). Pseudowords were generated by replacing the second syllable of one word (e.g. BAL-KEN, transl.: beam) by one of another word (e.g. MANTEL, transl.: coat) to form a word-like entity with no semantic content in the German language (BAL-TEL). Auditory word/pseudoword stimuli, spoken by a female speaker, were adjusted to 42% of the intensity of the amplitude-modulated (AM) MSB to provide a task difficulty similar to the visual task ( $\sim 80\%$  accuracy on average). Subsequently, they were algebraically added to the AM-MSB. Visual word/pseudoword stimuli were embedded as syllables in two successive cycles of the RSVP (see Fig. 1) (Table 1).

#### Behavioural data

Responses were considered correct when a button press occurred from 750 to 1,300 ms after the onset of a word stimulus in the *attend audition* condition and from 300 to 1,000 ms after the onset of a word stimulus in the *attend vision* condition. Time windows were derived from the distribution of reaction times across subjects. Two types of false alarms were defined in equal time windows: Type I as responses to pseudoword stimuli in the same modality and type II as responses to word stimuli in the other modality. Behavioural data were analysed by means of two accuracy measures that were obtained for both experimental conditions according to the following formula (adapted from Macmillan and Creelman 2005, p. 7):

**Table 1** Fixed set of German word and corresponding pseudoword stimuli as used in the lexical decision task

Word (translation)	Pseudoword
BAL-KEN (beam)	BAL-TEL
DIS-KUS (discus)	DIS-PEL
DOK-TOR (doctor)	DOK-TEN
GAR-TEN (garden)	GAR-BUS
GÜR-TEL (belt)	GÜR-BEL
HUS-TEN (cough)	HUS-BIS
KAS-TEN (box)	KAS-TAR
KOL-BEN (piston)	KOL-PER
KOR-KEN (cork)	KOR-TOR
KÖR-PER (body)	KÖR-DIS
MAN-TEL (coat)	MAN-ZEN
MUS-KEL (muscle)	MUS-PER
PUL-VER (powder)	PUL-MOR
SIL-BER (silver)	SIL-TEN
WIM-PER (eyelash)	WIM-KUS
WUR-ZEL (root)	WUR-GEN
ZIR-KUS (circus)	ZIR-TEL

Separation of syllables for presentation is exemplified

$$\text{Accuracy} = \frac{(N_{\text{hits}} + N_{\text{correct rejections}})}{(N_{\text{hits}} + N_{\text{correct rejections}} + N_{\text{misses}} + N_{\text{false alarms}})} * 100\% \quad (1)$$

Intramodal accuracy was calculated using type I false alarms as  $N_{\text{false alarms}}$  to quantify the difficulty of the task within the attended modality. Crossmodal accuracy expressed the distractibility of subjects' attention towards the effectively ignored sensory stream and was calculated using type II false alarms as  $N_{\text{false alarms}}$  in Eq. 1. Both accuracy measures were tested separately between *attend audition* and *attend vision* conditions by means of paired *t* tests. Reaction times were not analysed since differences between conditions were unavoidable due to the differences in word/pseudoword presentation between sensory streams and thus considered not informative.

### EEG recording and analysis

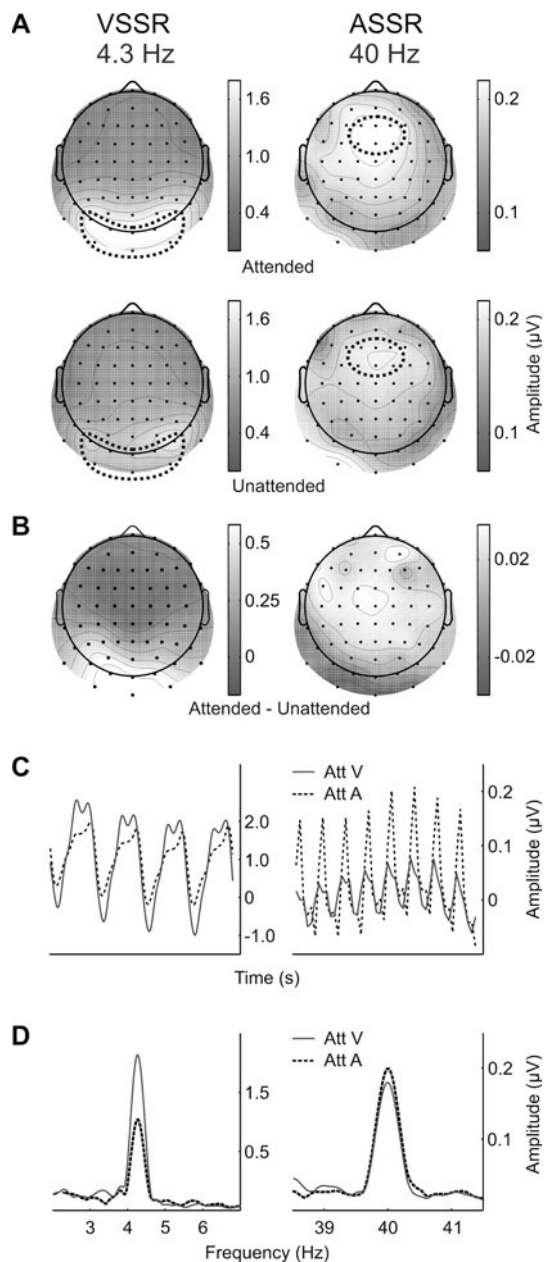
Brain electrical activity was recorded from 64 Ag/AgCl scalp electrodes and amplified by a BioSemi ActiveTwo amplifier (BioSemi, Amsterdam, Netherlands) that was set to a sampling rate of 256 Hz. Electrodes were mounted in a nylon cap following an extended version of the international 10–20 system (Oostenveld and Praamstra 2001). Lateral eye movements were monitored with a bipolar outer canthus montage. Vertical eye movements and blinks were monitored with a bipolar montage positioned below

and above the right eye. We extracted epochs that started 1,200 ms prior to and ended 4,500 ms after the onset of the two sensory stimulus streams. Epochs from trials that contained word/pseudoword stimuli or elicited false behavioural responses were excluded as their presentation in either modality might have led to transient evoked brain responses (Rohrbaugh et al. 1990). Individual epochs were rejected automatically when contaminated with blinks or eye movements that exceeded a threshold of 25  $\mu\text{V}$ . Subsequently, we performed the 'statistical control of artefacts in dense array EEG/MEG studies' (SCADS, Junghöfer et al. 2000). This procedure corrects artefacts such as noisy electrodes by using a combination of channel approximation and epoch exclusion based on statistical parameters of the data. Epochs with more than 12 contaminated electrodes were excluded from further analysis. This resulted in an average rejection rate of 17.0%, which did not differ significantly between conditions. The number of trials per condition was matched to maintain a similar signal-to-noise ratio after averaging. To this end, trials were randomly drawn from the condition with less rejected trials. Afterwards data were re-referenced to the nose lead and averaged separately for both experimental conditions. Basic data-processing steps such as extraction of epochs from the continuous recordings, re-referencing and plotting scalp isocontour voltage maps made use of EEGLab (Delorme and Makeig 2004) in combination with custom routines written in MATLAB (The Mathworks, Natick, MA).

ASSR and VSSR amplitudes at each electrode were calculated by a Fourier transform of data from a 4,000-ms time window that started 500 ms after stimulus onset in order to exclude ERPs to stimulus onset. Prior to the Fourier transform, data within the time window were detrended (removal of mean and linear trends) and multiplied by a Hanning taper. ASSR and VSSR amplitudes were quantified as the absolute value of the complex Fourier coefficients at the stimulation frequencies 4.3 and 40 Hz.

The topographical distribution of VSSR amplitude values averaged over experimental conditions was maximal at a cluster of eight occipital electrode sites (Oz, O1, O2, Iz, I1, I2, PO7, PO8; see Fig. 2a). ASSR amplitude topography revealed a maximum at a cluster of six frontal electrode sites (Fz, FCz, F1, F2, FC1, FC2; see Fig. 2a). For subsequent statistical analyses, ASSR and VSSR amplitudes were averaged over electrodes of the corresponding clusters. The processing steps described above yielded respective 4.3-Hz VSSR and 40-Hz ASSR amplitudes from both conditions. For each subject, we normalised amplitudes by the individual mean across conditions to account for variance in absolute amplitude between subjects. Differences in normalised amplitudes of each frequency





**Fig. 2** **a** Grand average scalp isocontour voltage maps show amplitude maxima at frontocentral sites (40-Hz ASSR) and at occipital sites (4.3-Hz VSSR) in both attention conditions. *Dashed lines* indicate electrode clusters chosen for further analyses. **b** Difference maps (attended—unattended, calculated from **a**) show a frontocentral attention effect on ASSR amplitude and an occipital attention effect on VSSR amplitude. Note the different scales in **a** and **b**. **c** 4.3-Hz VSSR and 40-Hz ASSR waveforms of one representative subject, extracted by moving window averages for both experimental conditions, attend vision (Att V, *solid grey line*) and attend audition (Att A, *dashed black line*). **d** Grand average Fourier spectra show peaks at the respective stimulation frequencies for conditions attend vision and attend audition. As can be seen in **c** and **d**, attention to one modality leads to greater amplitudes of the corresponding SSRs when compared to attention to the other modality

**Table 2** Behavioural accuracy in the lexical decision task (SEM = standard error of the mean)

	Intramodal accuracy		Crossmodal accuracy	
	Mean (%)	SEM (%)	Mean (%)	SEM (%)
Attend vision	77.7	$\pm 3.0$	89.4	$\pm 1.4$
Attend audition	83.1	$\pm 1.4$	87.1	$\pm 1.1$

between *attend audition* and *attend vision* were tested by means of two-tailed paired *t* tests.

## Results

We found no systematic differences in difficulty between experimental conditions as indexed by similar intramodal accuracies ( $t(12) = 1.86$ ;  $P = 0.09$ ). Crossmodal accuracies of comparable magnitude indicated that subjects were able to maintain attention to the cued stream equally in either condition ( $t(12) = -1.66$ ;  $P = 0.12$ ) (Table 2).

SSR waveforms of one representative subject show the typical sinusoidal waveform (see Fig. 2c). Waveforms were extracted by means of a moving window average (e.g. Morgan et al. 1996) across the analysis time window (from 500 to 4,500 ms after stimulus onset) and were averaged across the cluster of six frontocentral electrodes for the 40-Hz ASSR and across the cluster of eight occipital electrodes for the 4.3-Hz VSSR.

Grand average spectra averaged across the same electrode clusters depict VSSR and ASSR amplitudes for both experimental conditions (see Fig. 2d). Waveforms and spectra show that attention to the respective modality led to greater amplitudes in the corresponding SSR. The 40-Hz ASSR amplitude was greater in the *attend audition* condition than in the *attend vision* condition ( $t(12) = 2.94$ ;  $P < 0.02$ ). Vice versa, the 4.3-Hz VSSR amplitude was enhanced compared with the *attend audition* condition when subjects had to attend to the RSVP of letter sets ( $t(12) = 4.68$ ;  $P < 0.001$ ).

## Discussion

The present experiment was motivated by inconsistent results from recent studies, which employed frequency tagging to investigate the effects of sustained selective IA on stimulus processing during concurrent audiovisual stimulation. We were able to corroborate earlier findings from Saupé et al. (2009b) and Gander et al. (2010) in a considerably modified paradigm that made use of a new

class of speech/language-like stimuli. Here, attention to amplitude-modulated multi-speech babble led to a significant enhancement of ASSR amplitude relative to when subjects attended to the visually presented letter sets. Likewise, VSSR amplitude was enhanced when the RSVP stream became task-relevant as compared with when subjects attended to the auditory stream.

Given similar behavioural performance in both conditions and the trialwise cueing, it is highly unlikely that SSR amplitude modulations can be explained by a difference in arousal between conditions (Griskova et al. 2007). Moreover, we used the identical set of words/pseudowords in auditory and visual tasks with only the task-relevant presentation modality varying between attention conditions. Thus, given our trial-by-trial cueing design, a cued shift of IA was not confounded with a switch in task sets (Töllner et al. 2008) because the lexical decision task was identical in both modalities. Only the task-relevant modality of word/pseudoword presentation varied between conditions.

Findings from previous studies (Ross et al. 2004; Bidet-Caulet et al. 2007; Saupé et al. 2009b) suggested that the AM frequency of the auditory stream needs to be task-relevant to find an attentional modulation of the ASSR amplitude. Contrary to that suggestion, we observed clear effects of attention although monitoring AM frequency would not help subjects to solve the task since it remained unchanged throughout the presentation of word/pseudoword stimuli. We assume that auditory stream processing involved higher-order cortices as would be necessary to perform the lexical decision task. Possibly, sources beyond primary auditory cortex contributed to the measured ASSR (Gutschalk et al. 1999). Cortices that correspond to these sources have been demonstrated to receive stronger attentional modulation (Johnson and Zatorre 2005), especially when words had to be processed (Grady et al. 1997).

In conclusion, we introduced substantial changes to earlier paradigms and found further evidence in favour of a selective IA mechanism that operates on the level of multisensory input. We demonstrated that IA is able to modulate the processing of complex stimuli in a trial-by-trial cueing paradigm. Importantly, unpredictable cueing of the to-be-attended modality is a prerequisite for using SSRs to investigate the time course of the IA shift itself (see e.g. Senkowski et al. 2008) in future studies.

Within the context of multisensory integration, a state of selective IA to one sensory modality can be considered a situation in which less integration takes place. Alsius et al. (2005, 2007) have already shown that integration is severely reduced when a concurrent, unimodal, attentional-demanding task has to be solved. Studying such “extreme” cases using naturalistic, frequency-tagged stimuli might provide a promising perspective that helps to incorporate attention and multisensory integration in a unified

framework (Lakatos et al. 2009; Koelewijn et al. 2010; Talsma et al. 2010).

**Acknowledgments** We thank Renate Zahn for help in data recording. Special thanks to Andreas Widmann for technical support. Further thanks to Anne Melzer, Philipp Ruhnau and two anonymous reviewers for comments on the manuscript. This experiment was realised using Cogent 2000 developed by the Cogent 2000 team at the FIL and the ICN and Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience. Work was supported by the Deutsche Forschungsgemeinschaft, graduate programme “Function of Attention in Cognition” (GRK 1182).

**Conflict of interest** None declared.

## References

- Alho K, Woods DL, Algazi A, Naatanen R (1992) Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalogr Clin Neurophysiol* 82:356–368
- Alho K, Woods DL, Algazi A (1994) Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology* 31:469–479
- Alsius A, Navarra J, Campbell R, Soto-Faraco S (2005) Audiovisual integration of speech falters under high attention demands. *Curr Biol* 15:839–843
- Alsius A, Navarra J, Soto-Faraco S (2007) Attention to touch weakens audiovisual speech integration. *Exp Brain Res* 183:399–404
- Andersen SK, Fuchs S, Müller MM (2011) Effects of feature-selective and spatial attention at different stages of visual processing. *J Cogn Neurosci* 23:238–246
- Berman RA, Colby CL (2002) Auditory and visual attention modulate motion processing in area MT+. *Brain Res Cogn Brain Res* 14:64–74
- Bidet-Caulet A, Fischer C, Besle J, Aguera PE, Giard MH, Bertrand O (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J Neurosci* 27:9252–9261
- Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. *J Physiol Paris* 98:191–205
- de Jong R, Toffanin P, Harbers M (2010) Dynamic crossmodal links revealed by steady-state responses in auditory-visual divided attention. *Int J Psychophysiol* 75:3–15
- de Ruiter MB, Kok A, van der Schoot M (1998) Effects of inter- and intramodal selective attention to non-spatial visual stimuli: an event-related potential analysis. *Biol Psychol* 49:269–294
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21
- Eimer M, Schröger E (1998) ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology* 35:313–327
- Gander PE, Bosnyak DJ, Roberts LE (2010) Evidence for modality-specific but not frequency-specific modulation of human primary auditory cortex by attention. *Hear Res* 268:213–226
- Ghatan PH, Hsieh JC, Petersson KM, Stone-Elander S, Ingvar M (1998) Coexistence of attention-based facilitation and inhibition in the human cortex. *Neuroimage* 7:23–29
- Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP (1997) Attention-related modulation of activity

- in primary and secondary auditory cortex. *Neuroreport* 8:2511–2516
- Griskova I, Morup M, Parnas J, Ruksenas O, Arnfred SM (2007) The amplitude and phase precision of 40 Hz auditory steady-state response depend on the level of arousal. *Exp Brain Res* 183:133–138
- Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hahnel S, Picton TW, Scherg M (1999) Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. *Clin Neurophysiol* 110:856–868
- Johnson JA, Zatorre RJ (2005) Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. *Cereb Cortex* 15:1609–1620
- Junghöfer M, Elbert T, Tucker DM, Rockstroh B (2000) Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology* 37:523–532
- Koelwijn T, Bronkhorst A, Theeuwes J (2010) Attention and the multiple stages of multisensory integration: a review of audio-visual studies. *Acta Psychol* 134:372–384
- Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The leading sense: supramodal control of neurophysiological context by attention. *Neuron* 64:419–430
- Linden RD, Picton TW, Hamel G, Campbell KB (1987) Human auditory steady-state evoked potentials during selective attention. *Electroencephalogr Clin Neurophysiol* 66:145–159
- Macmillan NA, Creelman DC (2005) *Detection theory: a user's guide*. Lawrence Erlbaum Associates, Mahwah, New Jersey
- Morgan ST, Hansen JC, Hillyard SA (1996) Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci USA* 93:4770–4774
- Müller MM, Teder-Salejarvi W, Hillyard SA (1998) The time course of cortical facilitation during cued shifts of spatial attention. *Nat Neurosci* 1:631–634
- Müller N, Schlee W, Hartmann T, Lorenz I, Weisz N (2009) Top-down modulation of the auditory steady-state response in a task-switch paradigm. *Front Hum Neurosci* 3:1
- Muller-Gass A, Roye A, Kirmse U, Saupé K, Jacobsen T, Schröger E (2007) Automatic detection of lexical change: an auditory event-related potential study. *Neuroreport* 18:1747–1751
- Näätänen R, Alho K, Schröger E (2002) Electrophysiology of attention. In: Pashler H, Wixted J (eds) *Steven's handbook of experimental psychology: methodology in experimental psychology*, vol 4. Wiley, New York, pp 601–653
- Oostenveld R, Praamstra P (2001) The five percent electrode system for high-resolution EEG and ERP measurements. *Clin Neurophysiol* 112:713–719
- Quigley C, Andersen SK, Schulze L, Grunwald M, Müller MM (2010) Feature-selective attention: evidence for a decline in old age. *Neurosci Lett* 474:5–8
- Regan D (1989) *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine*. Elsevier, New York
- Rohrbaugh JW, Varner JL, Paige SR, Eckardt MJ, Ellingson RJ (1990) Auditory and visual event-related perturbations in the 40 Hz auditory steady-state response. *Electroencephalogr Clin Neurophysiol* 76:148–164
- Ross B, Picton TW, Herdman AT, Pantev C (2004) The effect of attention on the auditory steady-state response. *Neurol Clin Neurophysiol* 2004:22
- Saupé K, Widmann A, Bendixen A, Müller MM, Schröger E (2009a) Effects of intermodal attention on the auditory steady-state response and the event-related potential. *Psychophysiology* 46:321–327
- Saupé K, Schröger E, Andersen SK, Müller MM (2009b) Neural mechanisms of intermodal sustained selective attention with concurrently presented auditory and visual stimuli. *Front Hum Neurosci* 3:58
- Senkowski D, Saint-Amour D, Gruber T, Foxe JJ (2008) Look who's talking: the deployment of visuo-spatial attention during multisensory speech processing under noisy environmental conditions. *Neuroimage* 43:379–387
- Shomstein S, Yantis S (2004) Control of attention shifts between vision and audition in human cortex. *J Neurosci* 24:10702–10706
- Skosnik PD, Krishnan GP, O'Donnell BF (2007) The effect of selective attention on the gamma-band auditory steady-state response. *Neurosci Lett* 420:223–228
- Talsma D, Kok A (2001) Nonspatial intermodal selective attention is mediated by sensory brain areas: evidence from event-related potentials. *Psychophysiology* 38:736–751
- Talsma D, Kok A (2002) Intermodal spatial attention differs between vision and audition: an event-related potential analysis. *Psychophysiology* 39:689–706
- Talsma D, Doty TJ, Strowd R, Woldorff MG (2006) Attentional capacity for processing concurrent stimuli is larger across sensory modalities than within a modality. *Psychophysiology* 43:541–549
- Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG (2010) The multifaceted interplay between attention and multisensory integration. *Trends Cogn Sci* 14:400–410
- Toffanin P, de Jong R, Johnson A, Martens S (2009) Using frequency tagging to quantify attentional deployment in a visual divided attention task. *Int J Psychophysiol* 72:289–298
- Töllner T, Gramann K, Müller HJ, Kiss M, Eimer M (2008) Electrophysiological markers of visual dimension changes and response changes. *J Exp Psychol Hum Percept Perform* 34:531–542
- Woods DL, Alho K, Algazi A (1992) Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalogr Clin Neurophysiol* 82:341–355