



Short-Term Audiovisual Spatial Training Enhances Electrophysiological Correlates of Auditory Selective Spatial Attention

Christina Hanenberg¹, Michael-Christian Schlüter¹, Stephan Getzmann² and Jörg Lewald^{1*}

¹ Faculty of Psychology, Ruhr University Bochum, Bochum, Germany, ² Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany

OPEN ACCESS

Edited by:

Isabelle Peretz,
Université de Montréal, Canada

Reviewed by:

Claude Alain,
Rotman Research Institute (RRI),
Canada
Benjamin Rich Zendel,
Memorial University of Newfoundland,
Canada

*Correspondence:

Jörg Lewald
Joerg.Lewald@rub.de

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 23 December 2020

Accepted: 09 June 2021

Published: 01 July 2021

Citation:

Hanenberg C, Schlüter M-C,
Getzmann S and Lewald J (2021)
Short-Term Audiovisual Spatial
Training Enhances
Electrophysiological Correlates
of Auditory Selective Spatial Attention.
Front. Neurosci. 15:645702.
doi: 10.3389/fnins.2021.645702

Audiovisual cross-modal training has been proposed as a tool to improve human spatial hearing. Here, we investigated training-induced modulations of event-related potential (ERP) components that have been associated with processes of auditory selective spatial attention when a speaker of interest has to be localized in a multiple speaker (“cocktail-party”) scenario. Forty-five healthy participants were tested, including younger (19–29 years; $n = 21$) and older (66–76 years; $n = 24$) age groups. Three conditions of short-term training (duration 15 min) were compared, requiring localization of non-speech targets under “cocktail-party” conditions with either (1) synchronous presentation of co-localized auditory-target and visual stimuli (audiovisual-congruency training) or (2) immediate visual feedback on correct or incorrect localization responses (visual-feedback training), or (3) presentation of spatially incongruent auditory-target and visual stimuli presented at random positions with synchronous onset (control condition). Prior to and after training, participants were tested in an auditory spatial attention task (15 min), requiring localization of a predefined spoken word out of three distractor words, which were presented with synchronous stimulus onset from different positions. Peaks of ERP components were analyzed with a specific focus on the N2, which is known to be a correlate of auditory selective spatial attention. N2 amplitudes were significantly larger after audiovisual-congruency training compared with the remaining training conditions for younger, but not older, participants. Also, at the time of the N2, distributed source analysis revealed an enhancement of neural activity induced by audiovisual-congruency training in dorsolateral prefrontal cortex (Brodmann area 9) for the younger group. These findings suggest that cross-modal processes induced by audiovisual-congruency training under “cocktail-party” conditions at a short time scale resulted in an enhancement of correlates of auditory selective spatial attention.

Keywords: cognitive training, auditory selective spatial attention, cocktail-party effect, dorsolateral prefrontal cortex, sound localization

INTRODUCTION

Numerous lines of human and animal research have provided clear evidence that the representation of sound sources in space can be modulated by vision. In particular, in the so-called ventriloquism effect the perceived sound location is shifted toward a spatially disparate, temporally coincident visual event (e.g., Klemm, 1910; Thomas, 1941; Jackson, 1953; Howard and Templeton, 1966; Jack and Thurlow, 1973; Lewald et al., 2001; Lewald and Guski, 2003; for review, see Radeau, 1994). Moreover, exposure to a consistent audiovisual spatial disparity over a certain period of time can induce a systematic shift in sound localization such that the representation of the auditory space is shifted to that of the visual space (Helmholtz, 1867; Stratton, 1896, 1897; Held, 1955; Kalil and Freedman, 1967; Canon, 1970, 1971; Radeau and Bertelson, 1977, 1978; Recanzone, 1998; Lewald, 2002b). These cross-modal adaptive changes, which can emerge over short time scales from seconds to minutes (cf. Bosen et al., 2018), have been termed ventriloquism after-effect. Inspired by animal studies, which demonstrated similar (though more long-term) plasticity of auditory and visual neural representations (Knudsen and Knudsen, 1985, 1989; Brainard and Knudsen, 1993; Knudsen, 1999; Hyde and Knudsen, 2000, 2002; Zheng and Knudsen, 2001), experiments on this phenomenon led to the conception that vision calibrates human auditory spatial perception (Recanzone, 1998; Lewald, 2002b). Results obtained in blind and blindfolded sighted humans as well as in patients with visual-field loss demonstrating specific alterations of sound localization were in accordance with this view (e.g., Zwiers et al., 2001a,b; Lewald, 2002a,c, 2013; Lewald et al., 2009a,b, 2013; Feierabend et al., 2019).

In terms of beneficial effects of multisensory learning (for review, see Shams and Seitz, 2008), approaches of sensory training have been developed, in which auditory stimuli were presented in spatio-temporal alignment with visual stimuli. For example, in patients with pure hemianopia, who suffer from a loss of one half of the visual field due to brain damage while having sufficient audiospatial performance (Lewald et al., 2009b, 2013), neuro-rehabilitative audiovisual training or even auditory unimodal stimulation have been shown to induce long-lasting improvements of visual functions in the anopic hemifield (Bolognini et al., 2005; Passamonti et al., 2009; Lewald et al., 2012). Conversely, cross-modal approaches of sensory training have also demonstrated improvements of auditory functions, suggesting that persons with hearing impairments could benefit from them (Zahorik et al., 2006; Strelnikov et al., 2011; Majdak et al., 2013; Kuk et al., 2014; Grasso et al., 2016; Cai et al., 2018). In the audiospatial domain, audiovisual training has been shown to significantly increase the accuracy of localization of single sound sources under monaural (Strelnikov et al., 2011) and binaural conditions in healthy adults (Cai et al., 2018). In particular, auditory-visual training was found to induce a stronger improvement in sound localization compared to auditory-only training and a significant reduction of front-back confusion for both, trained and untrained sound positions (Cai et al., 2018).

While previous approaches to improve audiospatial performance by sensory training used single sound sources presented in isolation (e.g., Cai et al., 2018), spatial hearing in everyday life requires more complex functions of selective spatial attention, since auditory objects of interest have to be detected and localized among several distractor sound sources. Listening in such a “cocktail-party” situation (Cherry, 1953) is a remarkable ability of the human auditory system, which allows to orient the focus of attention to a sound source of interest in noisy, multi-speaker scenarios (for review, see Bregman, 1990; Bronkhorst, 2015). Such conditions of listening can be challenging already for normal-hearing people, but become substantially more difficult in hearing-impaired persons and at older age, resulting in serious restrictions of communication and social interaction in everyday life (Lewald and Hausmann, 2013; Getzmann et al., 2014, 2015c; see also Pichora-Fuller et al., 2017). This leads to the question of whether hearing performance under these conditions could be improved by training interventions. On the basis of the previous sensory training approaches mentioned above, it seems reasonable to assume that audiovisual stimulation could be an effective tool in this respect. Thus, in the present study a bimodal spatial training was developed in order to enhance brain functions associated with selective spatial attention under multiple-speaker conditions. Two types of sensory training were employed and compared with a control condition: (1) an audiovisual-congruency training, in which auditory targets were presented in spatiotemporal alignment with light stimuli, and (2) a visual-feedback training, in which correct responses on target location were indicated by light flashes. We hypothesized that audiovisual-congruency training may result in more effective learning due to the specific enhancement of multimodal brain circuits by audiovisual spatiotemporal alignment (Shams and Seitz, 2008). Due to findings on auditory short-term or so-called rapid learning (e.g., Alain et al., 2007, 2010), we expected relatively short training periods of about 15 min to enhance earlier evoked responses. For example, Alain et al. (2010) observed rapid improvements in performance within the first few of ten training blocks of trials (<1 h) in a sound identification task of two spectrally identical consonant vowel syllables, differing in voice onset time. The improvements in performance were accompanied by rapid physiological changes in the human auditory system (N1, P2, N2, and LPC).

Here we focused on training-induced modulations of event-related potential (ERP) components that have been associated with processes of auditory selective spatial attention in “cocktail-party” situations, using a multiple-speaker sound-localization task (Lewald et al., 2016; Hanenberg et al., 2019). In particular, we expected a training-induced increase of the N2 component. Localization of predefined auditory target stimuli in multiple-distracter environments has been shown to result in a substantially stronger N2 component of the ERP compared with single-source localization (Lewald and Getzmann, 2015). More generally, the N2 component has been regarded as a neural correlate of processes of cognitive control and orienting attention (Potts, 2004; Folstein and Van Petten, 2008), and has been related to conflict processing and suppression of irrelevant information in the auditory domain (Falkenstein et al., 2002;

Bertoli et al., 2005; Getzmann et al., 2015c; Rueda et al., 2015). Even in patients suffering from anxiety, attention training resulted in larger N2 amplitudes and a better ability of avoiding “threats,” which the authors interpret as higher demand of orienting attention away from the “threat” (Dennis and Chen, 2007; Eldar and Bar-Haim, 2010). Regarding age, the N2 is usually reduced in older, compared with younger, adults (e.g., Anderer et al., 1996; Wascher and Getzmann, 2014). This has been interpreted to reflect an age-related decline in inhibitory control over concurrent speech information (Getzmann et al., 2015b,c), in line with the more general inhibitory deficit hypothesis (Hasher and Zacks, 1988; for review, see Gazzaley and D’Esposito, 2007). It has been shown that explicit training of divided attention led to increased N2 amplitudes while also improving older adults’ ability to allocate their attention (Zendel et al., 2016). Therefore, as we were interested in potential age-related differences in training effects on ERPs (as also demonstrated for the P300 or Late positive component, LPC; e.g., Yang et al., 2018), groups of younger (19–29 years) and older participants (66–76 years) were compared.

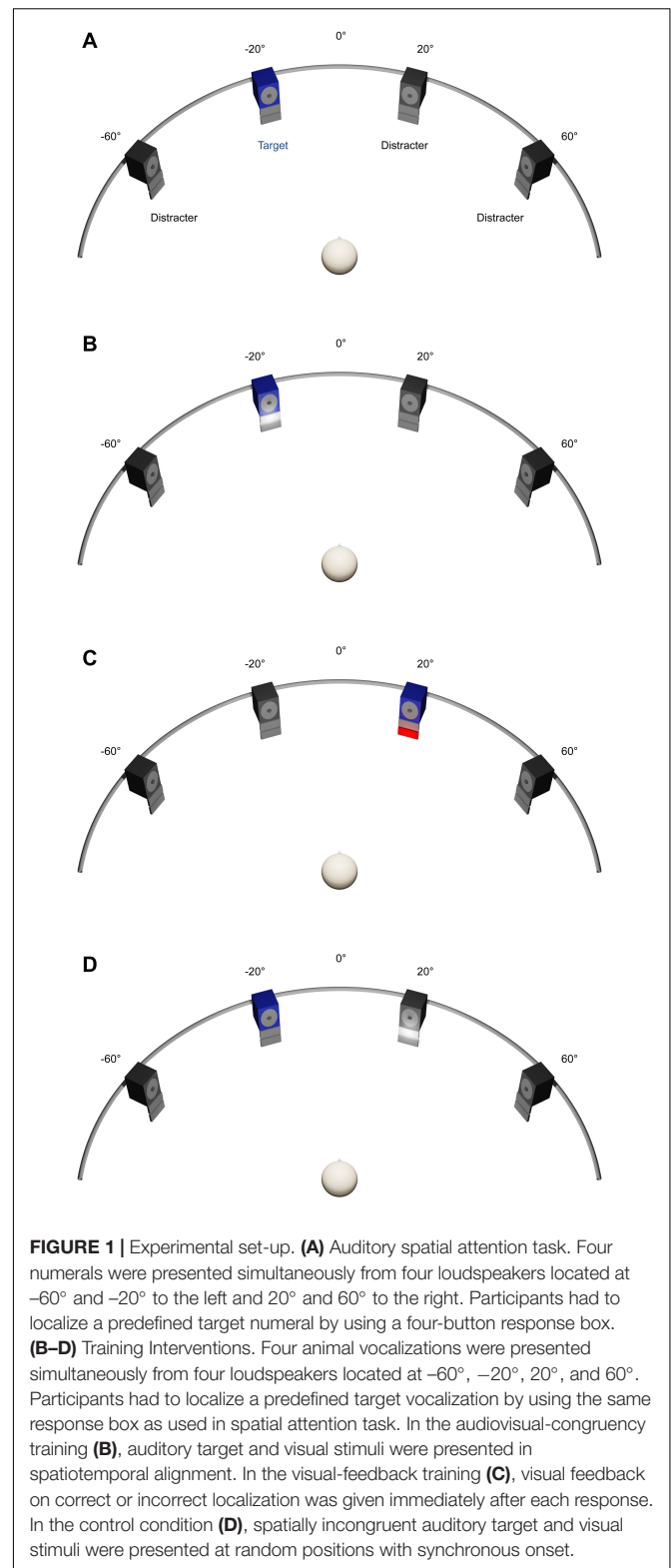
MATERIALS AND METHODS

Participants

Forty-five adult participants took part in this study, assigned to either a younger ($n = 21$; 12 women; mean age 25.0 years, SE 0.7 years, age range 19–29 years) or older group ($n = 24$; 12 women; mean age 71.0 years, SE 0.7 years, age range 66–76 years). Data from eight further participants (three younger and five older) were excluded from the analyses since the participants responded in less than 50% of all trials. All participants spoke German fluently and wrote with their right hand. Audiometric thresholds of younger participants were normal (mean across 11 pure-tone frequencies ≤ 25 dB hearing level), while the older group showed mild impairments (≤ 40 dB hearing level; 0.125–8 kHz; Oscilla USB100, Inmedico, Lystrup, Denmark), within the normal range of age-related hearing loss. This study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the British Medical Journal (July 18, 1964) and was approved by the Ethical Committee of the Leibniz Research Centre for Working Environment and Human Factors, Dortmund. All participants gave their written informed consent for participation.

Experimental Setup

As in previous studies, the experiments were carried out in a dimly illuminated sound-proof and echo-reduced room ($5.0 \times 3.3 \times 2.4$ m³; Getzmann et al., 2014). Participants were seated on a comfortable chair, positioned with equal distances to the front- and side-walls of the room. At a distance of 1.5 m from the participant’s head, a semicircular array with four active broadband loudspeakers (SC5.9; Visaton, Haan, Germany; housing volume 340 cm³) was arranged in the horizontal plane at -60° and -20° to the left, and at 20° and 60° to the right of the participant’s median plane (Figure 1). The participant’s head position was stabilized by a chin rest. In the median plane of the



participant’s head, a red light-emitting diode (LED) (diameter 3 mm, luminous intensity 0.025 mcd), mounted at the central position of the semicircular array, served as fixation target during

testing, while no fixation was required during training. For audiovisual-congruency and visual-feedback training (see section “General Procedure”), an array consisting of a red ($4 \times 7 \text{ cm}^2$; 70 cd/m^2) and a white LED screen ($4 \times 7 \text{ cm}^2$; 200 cd/m^2) was mounted below each loudspeaker. The two screens were mounted below each other, with the white screen in the upper position, immediately adjacent to the lower edge of the loudspeaker housing (see **Figure 1**).

Auditory Stimuli

The auditory stimuli used for testing were as in previous studies (for further details, see Lewald et al., 2016). Four different German one syllable numerals, spoken by four different speakers (duration about 600 ms) were presented with synchronous stimulus onset from the four loudspeaker positions. The numerals used were “eins” (1), “vier” (4), “acht” (8), and “zehn” (10), each of them spoken by two male and two female native German speakers. Each numeral was presented from each of the four loudspeakers with equal frequency of occurrence at an overall sound pressure level of 66 dB(A), as was measured at the position of the participant’s head, using a sound level meter with a 1/2” free-field measuring microphone (Types 2226 and 4175, Brüel & Kjaer, Naerum, Denmark).

Stimuli used for training have already been used previously (for details, see Lewald, 2016). Sounds consisted of animal vocalizations taken from a sound library (Marcell et al., 2000). Four different vocalizations (denoted “birds chirping,” “dog barking,” “frog,” “sheep” in Marcell et al., 2000) were presented with synchronous stimulus onset from the four loudspeakers. Stimulus durations were adjusted to about 600 ms by cutting the original sound files. Stimuli were digitized at 48 kHz sampling rate and 16-bit resolution and converted to analog form via a PC-controlled soundcard (Terrasoniq TS88 PCI, TerraTec Electronic, Nettetal, Germany). Each animal vocalization was presented from each of the four loudspeakers with equal frequency of occurrence. Stimuli were presented at an overall sound pressure level of 70 dB(A), measured as for the stimuli used for testing (see above).

General Procedure

Following a within-participant repeated-measures crossover design, each participant was tested in three sessions conducted on different days, with intervals of at least 1 week and maximally 3 weeks in between. Each session comprised four blocks, one training block (288–316 trials, duration 15.0–16.5 min), and three identical blocks of testing (prior to, immediately after, and 1 h after training) with an auditory spatial-attention task (see below), each consisting of 288 trials (duration 15 min). The first (pre-training) test block was used as baseline measurement, immediately followed by the training block. Training blocks differed in the type of training conducted at each of the three testing days: a task of localizing a pre-defined target sound (animal vocalization) among three distracters was combined with (1) synchronous presentation of co-localized auditory and visual stimuli (*audiovisual-congruency*), (2) immediate visual feedback on correct or incorrect localization responses (*visual-feedback*), or (3) presentation of spatially incongruent auditory-target and

visual stimuli presented at random positions with synchronous onset (*control condition*). The sequence of conditions was counterbalanced across participants in each of the two groups (one-sample Kolmogorov–Smirnov test; $D \leq 0.24$, $p \geq 0.16$).

The second test block (post-training) was started immediately following the training, and the third test block (1-h post-training) began 60 min after the end of the training block. Between pre-training and training-blocks, the instructions given prior to experiments related to the specific training condition (see section “Training Conditions”) were briefly repeated. Between post-training and 1-h post-training blocks, participants were allowed to rest, remaining seated in the experimental chair. Prior to the experiment, all participants were informed that they would receive a type of audiovisual training in each session.

Auditory Spatial Attention Task

The auditory spatial attention task used for testing was similar to previous studies (Lewald et al., 2016; Hananberg et al., 2019). In each trial, participants had to localize a predefined target numeral out of three distracter numerals by pressing one out of four response buttons (**Figure 1A**). The buttons were semi-circularly arranged on a response box, representing the four possible target positions (i.e., far left, mid left, mid right, and far right). Participants were instructed to use the right index finger for responding. For each participant, one numeral (1, 4, 8, or 10) was defined as target, which was kept constant for all measurements. Targets were counterbalanced across participants. During testing, participants had to keep the eyes open and to fixate on the central LED to reduce artifacts due to eye movements and alpha activity in the EEG. Participants were instructed to respond as accurately as possible within about 2 s after stimulus offset. Each trial lasted 3.125 s. Target position, distracter positions, and speakers changed between trials following a pseudo-random order (for details, see Lewald et al., 2016). The timing of the stimuli and the recording of the participants’ responses were controlled by custom-written software. Before the beginning of the experiment, participants completed about ten practice trials. They did not receive any feedback on their performance in the auditory spatial attention task during the experiment.

Training Conditions

In all sessions, participants were instructed to indicate target locations using the response box, as described for the auditory spatial attention task (see above). Target sounds used with training were assigned to target numerals during test blocks, that is, when the target numeral was “1” during testing, the target animal vocalization during training was always “birds chirping,” “4” was combined with “dog barking,” “8” with “frog,” and “10” with “sheep.” In the *audiovisual-congruency training* condition (**Figure 1B**), the target sound appeared simultaneously with illumination of the white LED screen (duration 600 ms) in 288 of a total of 316 trials (trial duration 3.125 s). The LED screen was mounted immediately below the target loudspeaker. To keep constant the participants’ spatial attention and preventing them from relying only on visual cues, we included 28 catch trials, in which target and LED screen appeared at incongruent positions. Accordingly, participants were informed that white

LED screens did not reliably predict the position of the target and were instructed to rely on audition, not vision. Incorrect responses were indicated by flashing of a second, red LED screen (duration 600 ms) at the actual target position immediately after button pressing. The red LED screen was mounted immediately below the white LED screen. Participants were briefed about this procedure prior to training.

In the *visual-feedback training* condition (Figure 1C), the sound-localization task and the presentation of auditory stimuli were as in the audiovisual-congruency training condition, but without presentation of spatio-temporally congruent visual stimuli. Thus, there was no use of catch trials, thus resulting in a total of 288 trials. For visual feedback, each response was immediately followed by flashing of one of the two LED screens mounted below the loudspeaker that emitted the target sound (duration 600 ms). Correct responses were indicated by white light, incorrect responses by red light at the actual target position. Prior to training, the participant was informed about the visual feedback.

The *control* condition (Figure 1D) was similar to the audiovisual-congruency training condition (288 trials). However, the white LED screen appeared at random positions, always diverging from the auditory target (duration 600 ms). Participants were instructed to localize the predefined target sound while ignoring the light flashes. In this condition, no feedback was provided after pressing a button.

Analysis of Behavioral Data

In order to investigate potential effects of the different training paradigms on performance, absolute error was taken as the main measure of localization accuracy, in addition to the percentage of correct responses. The rationale for using this measure and the computation of absolute error has been described previously in detail (Lewald, 2016). In short, the participants' responses were assigned to the azimuth indicated by the position of the response button (-60° ; -20° ; 20° ; 60°), and the unsigned deviation of the response from the actual target azimuth was taken as absolute error. Responses to targets presented at $\pm 60^\circ$ were excluded from analyses since these data did not provide information on errors to more eccentric positions (Lewald, 2016). Absolute errors were normalized with reference to the pre-training block in the same way as described below for ERP data.

EEG Recording and ERP Analysis

The continuous EEG was sampled at 1 kHz using a QuickAmp-72 amplifier (Brain Products, Gilching, Germany) and 58 Ag/AgCl electrodes, with electrode positions based on the International 10–10 system. Horizontal and vertical electro-oculograms were recorded from four additional electrodes positioned around the left and right eyes. The ground electrode was placed on the center of the forehead, just above the nasion. Two additional electrodes were placed on the left and right mastoids. Electrode impedance was kept below 5 k Ω . The raw data were band pass filtered off-line (cut-off frequencies 0.5 and 25 Hz; slope 48 dB/octave), re-referenced to the average of 58 channels (56 EEG and 2 mastoid electrodes), and segmented into 2000-ms stimulus-locked epochs covering the period from -200 to

1800 ms relative to sound onset. Data were corrected for ocular artifacts using the procedure proposed by Gratton et al. (1983). Individual epochs exceeding a maximum-minimum difference of 200 μ V were excluded from further analysis, using the automatic artifact rejection implemented in the BrainVision Analyzer software (Version 2.0; Brain Products, Gilching, Germany). The remaining epochs were baseline corrected to a 200-ms pre-stimulus window and averaged for each participant, separately for each training condition (audiovisual-congruency; visual feedback; control) and each test block (pre-training; post-training; 1-h post-training).

Peaks of five primary ERP components (P1, N1, P2, N2, and LPC) were defined as the maximum positivity or negativity within particular latency windows of specific waveforms after sound onset (P1: 10–110 ms at FCz; N1: 60–160 ms at Cz; P2: 155–255 ms at FCz; N2: 240–340 ms at Cz; LPC: 400–700 ms at Pz). The choice of electrode positions was based on previous knowledge of the ERPs topographical scalp distribution (e.g., Anderer et al., 1996; Folstein and Van Petten, 2008; Martin et al., 2008) and confirmed by visual inspection of the grand average waveforms. As we were interested in effects of training on successful sound localization, only trials with correct responses were included in ERP analyses. To take account of placebo and learning effects, ERP data (peak amplitudes and latencies) were normalized by subtraction of pre-training results (baseline) using the formulae:

$$\text{baseline-normalized post-training value} =$$

$$\text{post-training value} - \text{pre-training value}$$

and

$$\text{baseline-normalized 1-h post-training value} =$$

$$1\text{-h post-training value} - \text{pre-training value}.$$

The baseline-normalized data were submitted to $3 \times 2 \times 2$ analyses of variance (ANOVAs) with training condition (audiovisual-congruency; visual feedback; control) and block (post-training; 1-h post-training) as within-participant factors and group (younger; older) as between-participants factor. In addition, to investigate effects of training on ERP amplitudes and latencies, baseline-normalized data were analyzed using one-sample *t*-tests against zero.

Cortical Source Localization

The cortical sources of the effect of training on ERP amplitudes were localized using standardized low-resolution brain electromagnetic tomography (sLORETA) (Pascual-Marqui, 2002), which is part of the LORETA-KEY software package (v20171101) of the KEY Institute for Brain-Mind Research, Zurich, Switzerland¹. Data were baseline corrected to a 200-ms pre-stimulus window for each participant, separately for each training and each test block. Then, data obtained in the

¹www.uzh.ch/keyinst/loreta

pre-training block were subtracted from data obtained in post-training blocks, and the resulting baseline-normalized data for training conditions were contrasted against the related baseline-normalized data for the control condition (paired groups, test $[A-A2] = [B-B2]$, with $A = \text{post/training}$, $A2 = \text{pre/training}$, $B = \text{post/control}$, $B2 = \text{pre/control}$). We employed sLORETA within 5-ms time windows around the individual ERP peak-amplitude values of P1, N1, P2, N2, and LPC (using all electrodes) for each participant, with the individual latencies taken from the ERP analyses described above.

RESULTS

Behavioral Data

Prior to the main analyses, performance levels were tested against chance-level in order to ensure that the task has been adequately solved. Also, performance in pre-training blocks was compared between younger and older groups to assess effects of age. Both groups showed high levels of percentages of correct responses, differing significantly from chance-level (25%; younger group: $t[20] = 28.55$, $p < 0.001$; older group: $t[23] = 13.09$, $p < 0.001$). For pre-training trials, an ANOVA including the within-participant factor condition (audiovisual-congruency; visual feedback; control) and the between-participants factor group (younger; older) revealed a significant main effect of group ($F[1,43] = 10.30$, $p = 0.003$, $\eta_p^2 = 0.19$), indicating a higher percentage of correct responses in younger ($M = 84.3\%$, $SE 2.4\%$), than older ($M = 69.9\%$, $SE 3.6\%$), participants. Also, prior to training, absolute errors were significantly smaller in the younger ($M = 11.7^\circ$, $SE 3.1^\circ$), than in the older ($M = 26.5^\circ$, $SE 2.9^\circ$, $F[1,43] = 11.94$, $p = 0.001$, $\eta_p^2 = 0.22$), group. There were no significant differences between training conditions within groups in the pre-training blocks, neither in the percentages of correct responses (all $F \leq 0.60$, $p \geq 0.55$), nor in absolute errors (all $F \leq 0.70$, $p \geq 0.50$).

In order to investigate potential effects of the different training paradigms on performance, absolute error was taken as the main measure of localization accuracy. A $3 \times 2 \times 2$ ANOVA on baseline-normalized absolute errors with training condition (audiovisual-congruency; visual feedback; control) and block (post-training; 1-h post-training) as within-participant factors and group (younger; older) as between-participants factor did not indicate main effects or interactions (all $F \leq 3.13$, $p \geq 0.084$). However, across conditions, blocks, and groups, baseline-normalized absolute errors were significantly below zero ($t[44] = -3.98$, $p = 0.0003$), thus indicating general improvement in accuracy that was independent of the type of training condition.

ERPs

In both groups, sound onset elicited a prominent response at vertex position Cz, mainly consisting of a positive deflection (P1), a negative deflection (N1), a second positive deflection (P2), a second negative deflection (N2), and a third positive deflection (LPC; **Figure 2**). The P2 and N2 waves were less prominent in amplitude in older, than younger, participants,

which is in line with previous results (e.g., Getzmann et al., 2015b). Mean latencies (with reference to sound onset) were 116 ms ($SE 3$ ms) for N1, 213 ms ($SE 3$ ms) for P2, 296 ms ($SE 5$ ms) for N2 waves, and 570 ms ($SE 5$ ms) for LPC (averaged across groups and conditions). Baseline-normalized data (see section “EEG Recording and ERP Analysis”) were submitted to $3 \times 2 \times 2$ ANOVAs, with training condition (audiovisual-congruency; visual feedback; and control) and block (post-training; 1-h post-training) as within-participant factors and group (younger; older) as between-participants factor, to detect potential impacts of the different types of training on amplitude and latency of ERP components, indicating short-term processes of auditory spatial attention.

P1 Wave

For baseline-normalized P1 peak amplitudes at electrode position Cz, the ANOVA indicated a significant interaction of condition and block ($F[1,43] = 3.19$, $p = 0.046$, $\eta_p^2 = 0.07$), but no further main effects or interactions (all $F \leq 3.00$, $p \geq 0.99$). However, t -tests against zero, conducted separately for training conditions and blocks, did not reveal significant results (all $t[44] \leq 1.33$, $p \geq 0.19$; one-sample t -test, two-tailed; Bonferroni-adjusted $\alpha = 0.008$). Also, paired *post hoc* t -tests, comparing baseline-normalized peaks obtained after audiovisual-congruency training or visual-feedback training with the control condition did not reveal any significant differences (all $t[44] \leq -1.61$, $p \geq 0.12$; Bonferroni-adjusted $\alpha = 0.007$). For baseline-normalized P1 peak latencies, there were no significant interactions or main effects (all $F \leq 1.17$, $p \geq 0.29$).

N1 Wave

For baseline-normalized N1 peak amplitudes at electrode position Cz, the ANOVA indicated a main effect of block ($F[1,43] = 9.90$, $p = 0.003$, $\eta_p^2 = 0.19$), but no further main effects or interactions (all $F \leq 1.73$, $p \geq 0.18$). Baseline-normalized amplitudes in post-training ($M = 0.49 \mu\text{V}$, $SE = 0.17$; $t[44] = 2.84$, $p = 0.007$), but not in 1-h post-training, blocks ($M = 0.08 \mu\text{V}$, $SE = 0.20$; $t[44] = 0.39$, $p = 0.70$) were significantly in the positive range (one-sample t -tests, two-tailed; Bonferroni-adjusted $\alpha = 0.025$), thus indicating reduction in N1 amplitude immediately after training. Baseline-normalized peak amplitudes were more positive in post-training blocks than in 1-h post-training blocks ($t[44] = 3.21$, $p = 0.002$; paired t -test, two-tailed).

For baseline-normalized N1 peak latencies, the ANOVA indicated a significant block \times group interaction ($F[1,43] = 6.99$, $p = 0.011$, $\eta_p^2 = 0.14$), as well as a main effect of group ($F[1,43] = 7.45$, $p = 0.009$, $\eta_p^2 = 0.15$). No further main effects or interactions were obtained (all $F \leq 1.10$, $p \geq 0.34$). N1 peak latencies significantly differed between groups in the post-training (younger group: $M = 3.49$ ms, $SE = 1.56$ ms; older group: $M = -5.26$ ms, $SE = 1.97$; $t[43] = -3.42$, $p = 0.001$), but not in the 1-h post-training, block (younger group: $M = 0.00$, $SE = 1.42$; older group: $M = -2.36$ ms, $SE = 1.60$ ms; $t[43] = -1.09$, $p = 0.28$; Bonferroni-adjusted $\alpha = 0.025$). However, t -tests against zero for baseline-normalized N1 latencies (conducted separately for blocks and groups) did not reveal significant results (younger

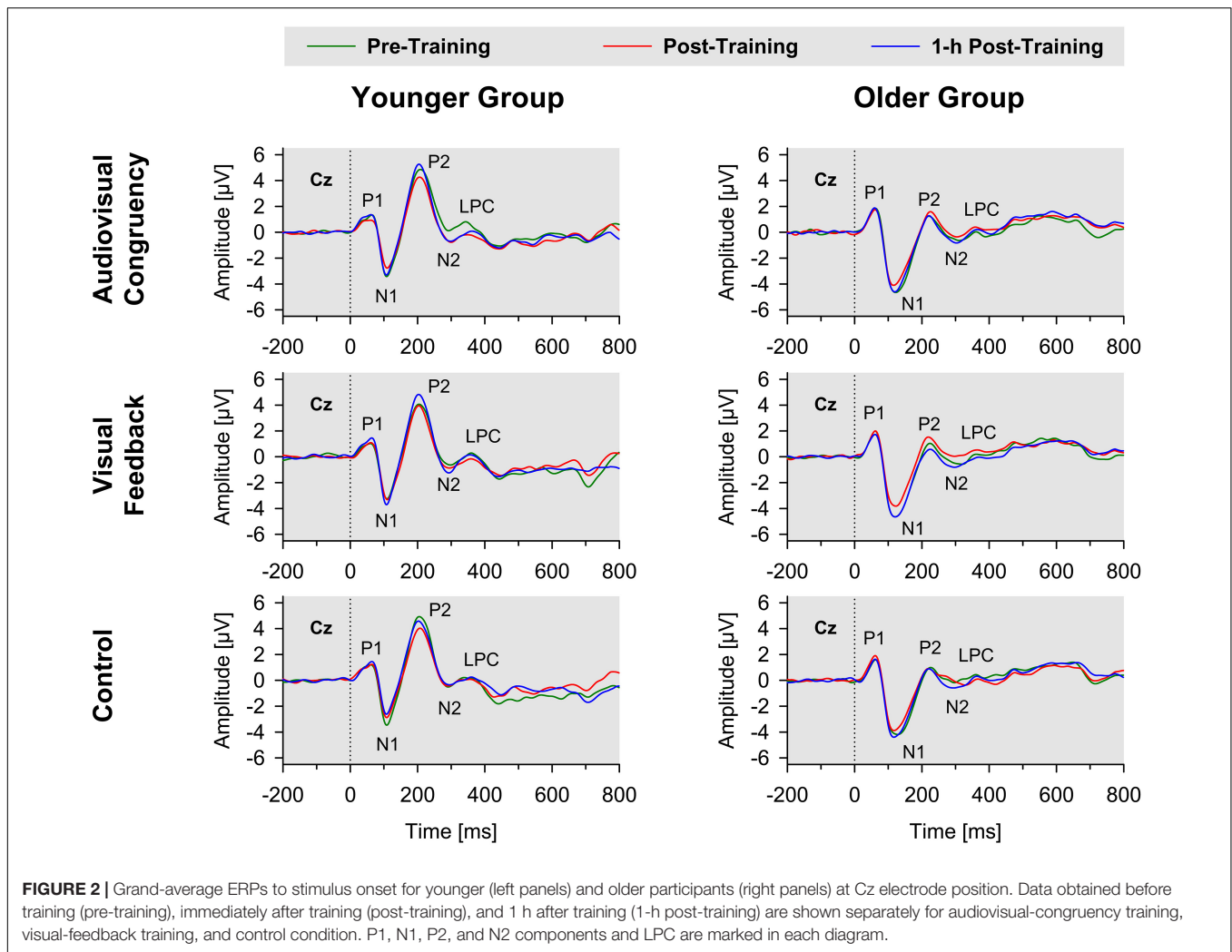


FIGURE 2 | Grand-average ERPs to stimulus onset for younger (left panels) and older participants (right panels) at Cz electrode position. Data obtained before training (pre-training), immediately after training (post-training), and 1 h after training (1-h post-training) are shown separately for audiovisual-congruency training, visual-feedback training, and control condition. P1, N1, P2, and N2 components and LPC are marked in each diagram.

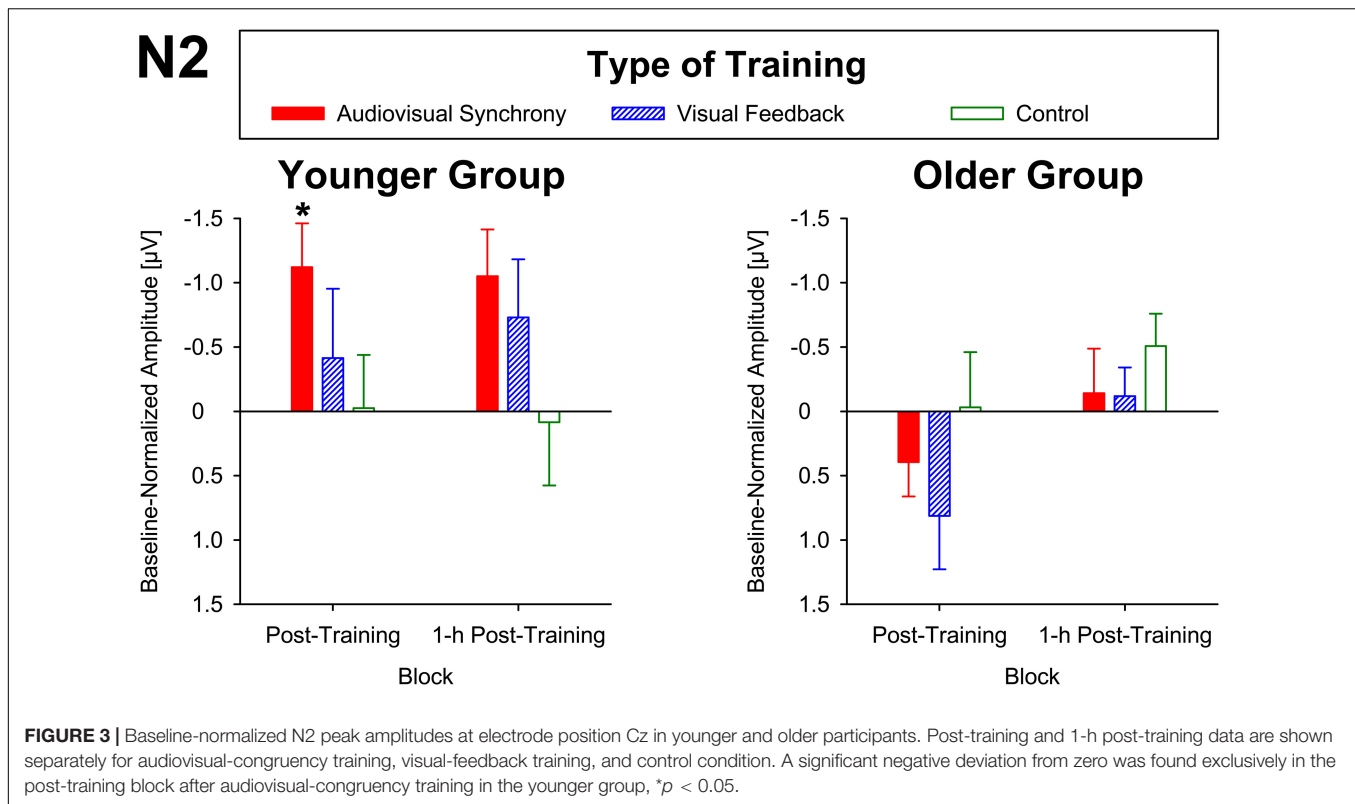
group: all $t[20] \leq 2.24$, $p \geq 0.04$; older group: all $t[23] \leq -2.67$, $p \geq 0.014$; Bonferroni-adjusted $\alpha = 0.0125$).

P2 Wave

For baseline-normalized P2 peak amplitudes at electrode position FCz, there was a significant interaction of block \times group ($F[1,43] = 9.24$, $p = 0.004$, $\eta_p^2 = 0.18$), but no further main effects or interactions (all $F \leq 2.54$, $p \geq 0.12$). As revealed by *post hoc* testing, amplitudes in the younger, but not in the older, group were significantly lower in post-training, than in 1-h post-training, blocks (younger group: post-training $M = -0.73 \mu V$, $SE = 0.41 \mu V$, 1-h post-training $M = 0.24 \mu V$, $SE = 0.36 \mu V$, $t[20] = -3.27$, $p = 0.004$; older group: post-training $M = 0.34 \mu V$, $SE = 0.29 \mu V$, 1-h post-training $M = 0.04 \mu V$, $SE = 0.28 \mu V$, $t[23] \leq 1.03$, $p > 0.31$; two paired *t*-tests, one for each group, two-tailed; Bonferroni-adjusted $\alpha = 0.0125$). However, *t*-tests against zero for baseline-normalized P2 peak amplitudes showed neither significant results when conducted separately for each block (all $t[44] \leq -0.62$, $p \geq 0.54$; one-sample *t*-test, two-tailed; Bonferroni-adjusted $\alpha = 0.025$), nor when conducted separately for each group (younger group: all $t[20] \leq -1.77$, $p > 0.09$; older

group: all $t[23] \leq 1.17$, $p > 0.26$; one-sample *t*-tests, two-tailed; Bonferroni-adjusted $\alpha = 0.0125$).

For baseline-normalized P2 peak latencies, the ANOVA indicated a significant interaction of training condition \times group ($F[1,43] = 5.70$, $p = 0.005$, $\eta_p^2 = 0.12$), as well as a main effect of block ($F[1,43] = 7.44$, $p = 0.009$, $\eta_p^2 = 0.15$), with no further main effects or interactions (all $F \leq 1.24$, $p \geq 0.27$). *Post hoc* testing revealed that baseline-normalized peak latencies were significantly shorter than zero in the 1-h post-training block ($M = -4.38$ ms, $SE = 1.11$ ms; $t[44] = -3.94$, $p < 0.0001$), but not in the post-training block ($M = -0.52$ ms, $SE = 1.44$ ms; $t[44] = -0.36$, $p = 0.72$; one-sample *t*-test, two-tailed; Bonferroni-adjusted $\alpha = 0.025$). Baseline-normalized latencies, analyzed separately for training conditions and groups, were significantly shorter than zero in the younger, but not in the older, group for the audiovisual-congruency condition in the 1-h post-training block, but not for the other conditions and the post-training block (younger group: $M = -10.19$ ms, $SE = 3.08$ ms, $t[20] = -3.31$, $p = 0.003$; all other $t[20] \leq -2.67$, $p > 0.015$; older group: all $t[23] \leq -3.11$, $p \geq 0.005$; one-sample *t*-tests, two-tailed; Bonferroni-adjusted $\alpha = 0.004$).



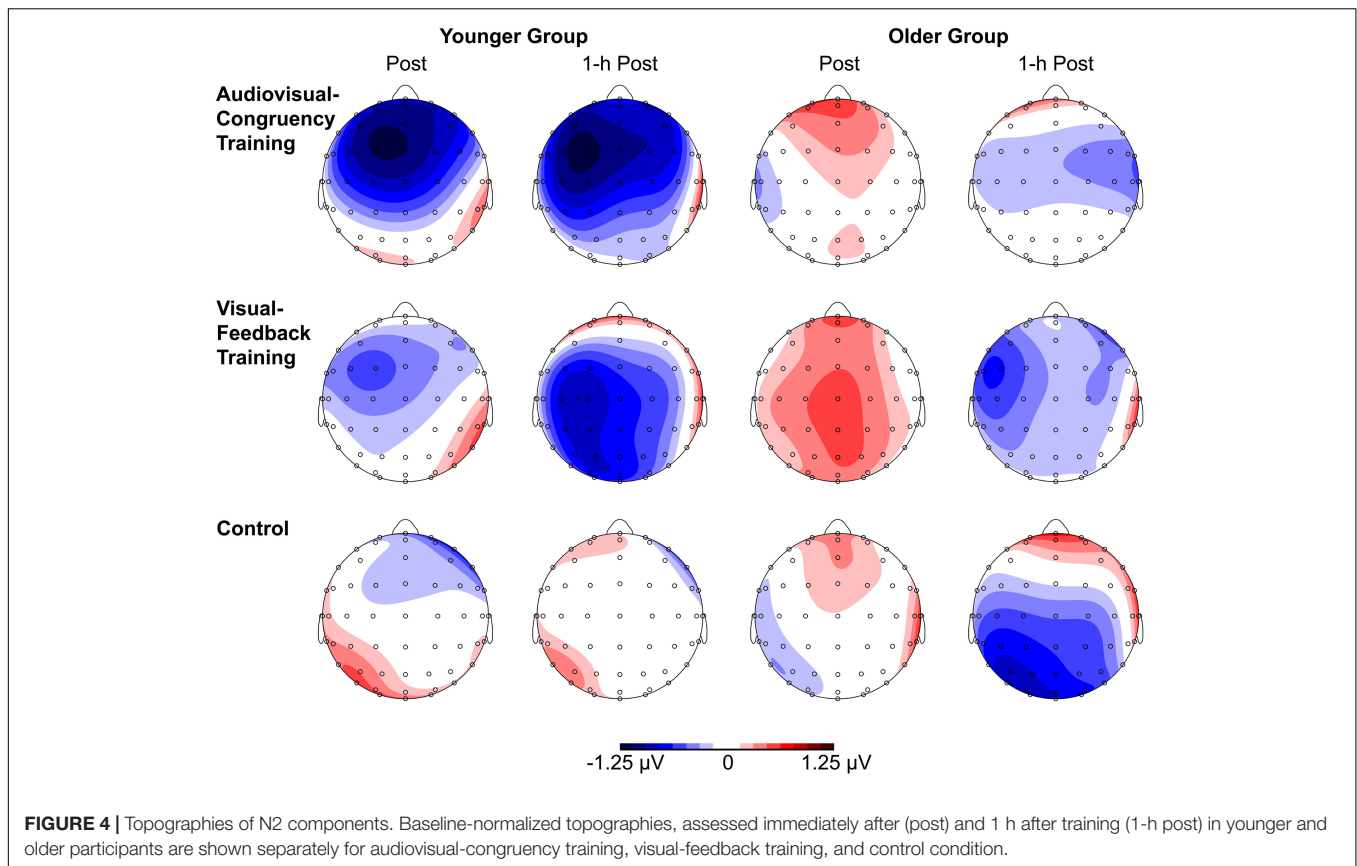
N2 Wave

For baseline-normalized N2 peak amplitudes at electrode position Cz, the ANOVA indicated a training condition \times group interaction ($F[2,86] = 3.34$, $p = 0.04$, $\eta_p^2 = 0.07$; **Figure 3**), but merely non-significant numerical trends for the factors block ($F[1,43] = 3.70$, $p = 0.06$, $\eta_p^2 = 0.08$) and group ($F[1,43] = 3.79$, $p = 0.06$, $\eta_p^2 = 0.08$). No further main effects or interactions were found (all $F \leq 2.78$, $p \geq 0.1$). Averaged across post-training and 1-h post-training blocks, baseline-normalized peak amplitudes were significantly larger than zero for the younger participants after audiovisual-congruency training ($t[20] = -3.39$, $p = 0.003$; two one-sample t -tests, two-tailed; Bonferroni-adjusted $\alpha = 0.008$), but not for the other conditions (all $|t[20]| \leq -1.24$, $p \geq 0.23$). Neither, there were any significant results for the older group with all conditions (all $t[23] \leq 1.27$, $p \geq 0.22$; one-sample t -test, two-tailed; Bonferroni-corrected $\alpha = 0.025$). In line with that, baseline-normalized N2 peak amplitudes in the audiovisual-congruency condition were significantly larger than in the visual-feedback training and control conditions exclusively for the younger group and for the post-training block, ($M = -1.12 \mu\text{V}$, $SE = 0.34 \mu\text{V}$; $t[20] = -3.29$, $p = 0.004$; all other $|t[20]| \leq -0.77$, $p \geq 0.45$; one-sample t -test, two-tailed; Bonferroni-adjusted $\alpha = 0.008$), indicating a specific increase in N2 amplitude after audiovisual-congruency training. In the 1-h post-training block, baseline-normalized N2 peak amplitudes in the audiovisual-congruency condition were numerically larger than in the visual-feedback training and control conditions, again, exclusively for younger participants ($M = -1.05 \mu\text{V}$, $SE = 0.36 \mu\text{V}$; $t[20] = -2.90$, $p = 0.009$;

all other $|t[20]| \leq -1.62$, $p \geq 0.12$; one-sample t -test, two-tailed; Bonferroni-adjusted $\alpha = 0.008$). *Post hoc* testing did not reveal any significant differences in the older group (all $|t[23]| \leq -2.00$, $p \geq 0.06$; one-sample t -test, two-tailed; Bonferroni-adjusted $\alpha = 0.008$). In line with these findings on N2 peak amplitudes, baseline-normalized topographies showed a fronto-central/parietal negativity after audiovisual-congruency training exclusively in the younger group, with its maximum in the left hemisphere (**Figure 4**).

For baseline-normalized peak latencies, the ANOVA revealed a significant main effect of block ($F[1,43] = 7.10$, $p = 0.011$, $\eta_p^2 = 0.14$). No further main effects or interactions were obtained (all $F \leq 2.09$, $p \geq 0.13$). Averaged across training conditions and groups, N2 latencies were significantly longer in post-training ($M = 6.28$ ms, $SE = 3.20$ ms), than 1-h post-training, blocks ($M = -0.06$ ms, $SE = 2.27$ ms; $t[44] = 6.34$, $p = 0.009$). However, there were no significant differences from zero (all $t[44] \leq 1.96$, $p \geq 0.056$; one-sample t -test, two-tailed; Bonferroni-corrected $\alpha = 0.025$).

To estimate potential effects of the sequence of the training conditions on these results, a four-factor ANOVA on the dependent variable baseline-normalized N2 amplitude was performed with the within-participant factors day of testing (first, second, or third) and block, and the between-participant factors sequence (all six permutations of sequences of audiovisual-congruency training, visual-feedback training, and control condition) and group. Neither in the younger, nor in the older group, the sequence of conditions had an impact on the N2 amplitude. The ANOVA indicated a significant interaction of day



of testing and group ($F[2,66] = 4.55, p = 0.014, \eta_p^2 = 0.12$), but no effects involving the factor sequence (all $F \leq 1.48, p \geq 0.17$). A *post hoc* inspection of the interaction of day of testing and group using two one-factor ANOVAs separately for each group suggested a non-significant trend of an increase of baseline-normalized N2 amplitude between the first and the last day of testing in the younger group ($F[2,40] = 3.22, p = 0.050, \eta_p^2 = 0.14$), but not in older participants ($F[2,46] = 0.23, p = 0.80$, Bonferroni-corrected $\alpha = 0.025$).

LPC

For baseline-normalized LPC peak amplitudes at electrode position Pz, the ANOVA did not reveal any main effects or interactions (all $F \leq 1.71, p \geq 0.19$). For baseline-normalized LPC peak latencies no main effects or interactions were obtained (all $F \leq 2.71, p \geq 0.11$).

Cortical Source of Electrical Activity at the Time of the N2

The cortical source of the enhancing effect of training on the N2 amplitude found for the younger group in the post-training block of the audiovisual-congruency condition was localized using sLORETA. Data obtained in the pre-block were subtracted from data obtained in the post-block and the resulting difference values were contrasted against the related difference values for the control condition. The analysis revealed an enhancement of electrical activity induced by audiovisual-congruency training at

a focal peak location at MNI coordinates $X = 25$ mm, $Y = 50$ mm, $Z = 40$ mm ($t = 6.45, p = 0.005$, two-tailed), in right superior frontal gyrus (Brodmann area, BA 9) in the anterior region of superior frontal sulcus (SFS) (**Figure 5**). Results for the five brain structures with the highest t -scores revealed by this contrast are summed up in **Table 1**.

Post hoc, related analyses using sLORETA were conducted for all ERP components, conditions, blocks, and groups, in order to control for the specificity of this result obtained at the time of the N2 for the younger group in the post-training block of the audiovisual-congruency condition. As shown in **Table 2**, all results except that described above failed to reach significance level (all $|t| \leq 4.92, p \geq 0.06$, two-tailed). Imaging data for all analyses are shown in **Supplementary Figure 1**.

DISCUSSION

These results demonstrated an age-specific effect of audiovisual spatial training on neurophysiological correlates of auditory selective spatial attention in a simulated “cocktail-party” scenario. After audiovisual-congruency training, but not after visual-feedback training and the control condition, the N2 peak amplitude was increased. This effect was observed in younger, but not older, participants with a non-significant decline within 1 h after training. At the time of the N2, electrical

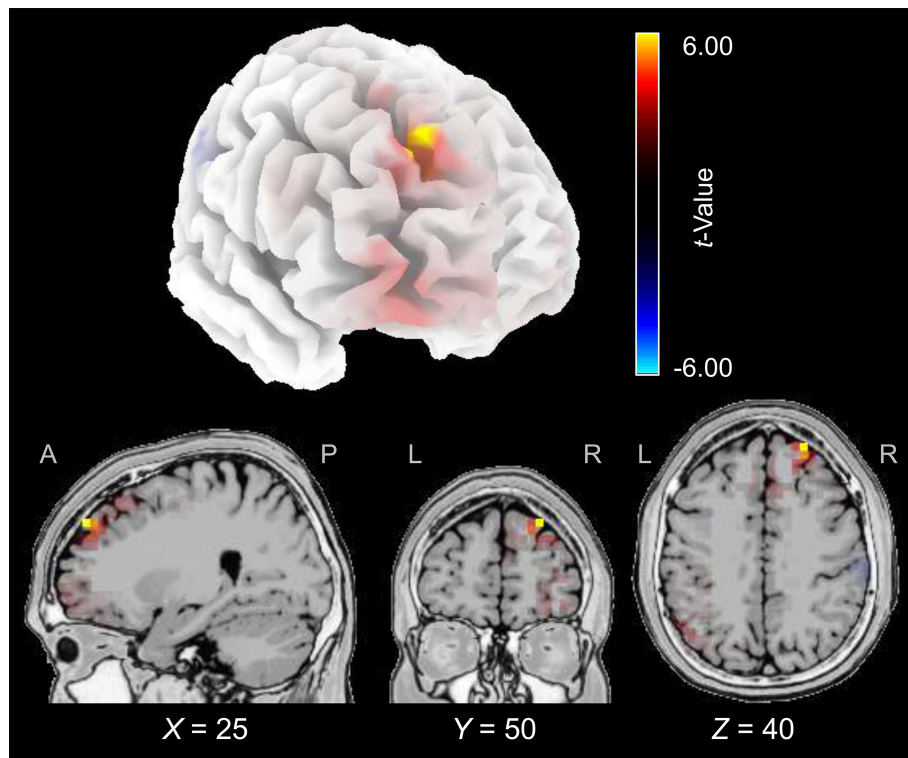


FIGURE 5 | Electrical brain activity at the time of the N2 after audiovisual-congruency training in the younger group. Electrical activity as revealed by sLORETA in the post-block was normalized with reference to pre-training activity and the resulting difference values were contrasted against the related difference values for the control condition. Data were mapped onto a standard 3-D brain template (top) or onto sagittal, coronal, and horizontal slices (T2 MNI-template “Colin 27” of sLORETA) positioned at X, Z, and Y coordinates (MNI) as given in the figure (bottom). Color coding shows t-values, with warm colors indicating increase of electrical activity and cold colors indicating decrease of electrical activity after training intervention (peak activity: $t = 6.45$, $p = 0.005$).

TABLE 1 | Brain structures revealed by the contrast of baseline-normalized activities obtained with audiovisual-congruency training vs. control condition at the time of the N2.

t-value	p-value	Structure	X	Y	Z
6.45	0.005**	Right Superior Frontal Gyrus	25	50	40
4.15	0.245	Left Postcentral Gyrus	-40	-30	55
3.77	0.411	Right Middle Frontal Gyrus	30	45	40
3.39	0.638	Left Angular Gyrus	-50	-75	30
3.31	0.688	Left Middle Temporal Gyrus	-45	-75	10

Peak activities immediately after training in the younger group are shown for the five structures with highest t-scores. Coordinates (X, Y, and Z) are in MNI space and refer to the voxel with the maximum t-score in each structure. Significant effect in bold, ** $p < 0.01$.

imaging revealed an increase in activity after audiovisual-congruency training that was located in right dorsolateral prefrontal cortex (DLPFC) (BA 9), in the anterior region of SFS. These findings provided, on the one hand, clear evidence that specific training characteristics, namely temporal synchrony and spatial congruency of auditory and visual target stimuli, can enhance ERP correlates of auditory spatial attention. On the other hand, we failed to find a related specificity for behavioral performance, which was generally improved after all training and control conditions.

Enhanced N2 Amplitude After Audiovisual-Congruency Training

In younger participants, audiovisual-congruency training induced a specific post-intervention enhancement of the N2, rather than P1, N1, and P2 components or LPC, for which no consistent changes were observed. This finding is in alignment with previous studies suggesting the N2 to be a decisive correlate of auditory selective spatial attention in “cocktail-party” scenarios (Gamble and Luck, 2011; Getzmann et al., 2015b; Lewald and Getzmann, 2015; Lewald et al., 2016; Hanenberg et al., 2019). An enhanced N2 was present not only immediately following training, but also in the 1-h-post-block, with a non-significant tendency to decrease as a function of time after training. Thus, it seems as if there were longer-lasting modulations in cortical processing after audiovisual-congruency training, which might be due to processes of neural plasticity induced by performing multisensory target localization.

An effect of cognitive training on the N2 has not been reported thus far in the context of auditory spatial attention. Previous studies focusing on improvements of working memory reported increases in N2 amplitude following *n*-back training in healthy participants (Covey et al., 2019) as well as in Multiple Sclerosis patients (Covey et al., 2018). Thus, on the basis of the limited data available at present, it remains open

TABLE 2 | Contrast of training vs. control conditions, as revealed by sLORETA.

Training Condition	Block	ERP Component	Younger Group		Older Group		
			<i>t</i> _{ext}	<i>p</i> _{ext}	<i>t</i> _{ext}	<i>p</i> _{ext}	
Audiovisual Congruency	Post-Training	P1	-3.21	0.862	2.85	0.962	
		N1	3.70	0.473	-3.78	0.437	
		P2	3.32	0.705	-3.25	0.711	
		N2	6.45	0.005**	2.24	1.000	
		LPC	-3.85	0.395	-3.08	0.919	
	1-h Post-Training	P1	4.22	0.237	-2.65	0.993	
		N1	3.64	0.355	3.74	0.416	
		P2	4.92	0.060	-4.60	0.098	
		N2	4.15	0.234	-3.59	0.578	
		LPC	3.59	0.563	3.59	0.536	
	Visual Feedback	Post-Training	P1	3.14	0.902	-4.15	0.258
			N1	3.23	0.701	3.32	0.704
			P2	3.23	0.734	-3.33	0.729
			N2	3.73	0.482	-4.04	0.290
LPC			3.37	0.703	-3.46	0.624	
1-h Post-Training		P1	3.85	0.465	-3.72	0.444	
		N1	3.57	0.405	-3.61	0.498	
		P2	3.63	0.478	-3.78	0.407	
		N2	3.52	0.570	-3.58	0.596	
		LPC	4.17	0.256	-3.62	0.554	

Extreme *t*-scores (*t*_{ext}) and related *p*-values (*p*_{ext}) resulting from comparisons of baseline-normalized electrical brain activities are shown for training conditions (audiovisual-congruency training, visual-feedback training), blocks (post-training, 1-h post-training), and ERP components (P1, N1, P2, N2, and LPC). Significant effect in bold, ***p* < 0.01.

whether audiovisual-congruency training had an effect on N2-related processes specifically involved in auditory attentional functions or rather enhanced more general cognitive processes associated with the N2.

The N2 enhancement following training was restricted to the younger group, which was an unexpected finding. The absence of an effect of training in the older group could possibly be due to the less flexible cognitive system of older adults, which is generally more occupied by task difficulty than that of younger persons (Yordanova et al., 2004; Salthouse, 2009; Gajewski and Falkenstein, 2012; Getzmann et al., 2015a,b,c; Olfers and Band, 2018). Our finding may be compatible with ERP research on age differences in speech perception. While older adults often showed reduced N2 amplitudes (Daffner et al., 2015), probably reflecting a less successful reorienting toward the stimulus of interest, younger adults showed more pronounced N2 amplitudes, reflecting an enhanced inhibitory control (Getzmann et al., 2014, 2015c). Thus, that the N2 enhancement was observed here in younger, but not older, adults could indicate a specific impact of the audiovisual-congruency training on processes of cognitive control. The question of whether related ERP enhancements, as found here for the N2 in younger participants,

can also be induced in elderly people by more intense and repeated daily training over longer periods has to be addressed by future studies.

Cortical Sources of Training-Induced N2 Enhancement

At the time of the N2, an increase of electrical activity after audiovisual-congruency training was found to be located in right superior frontal gyrus, in the anterior region of SFS (cf. Figure 5). Previous studies, focusing on the cortical correlates of selective auditory spatial attention using various methods, have revealed several cortical regions, composing a complex network. This network comprises auditory cortex, posterior superior temporal gyrus (pSTG) and planum temporale (PT), inferior parietal lobule (IPL), superior parietal lobule and precuneus, inferior frontal gyrus, frontal eye field (FEF), as well as regions of BA 9 and SFS, which were nearby the location of training-induced activity change found here (Pugh et al., 1996; Nakai et al., 2005; Zündorf et al., 2013, 2014; Kong et al., 2014; Lee et al., 2014; Lewald and Getzmann, 2015; Braga et al., 2016, 2017; Lewald, 2016, 2019; Lewald et al., 2016, 2018). In particular, the area including SFS and FEF has been related to the so-called N2ac (Lewald et al., 2016), an anterior contralateral N2 subcomponent, which has been regarded as a correlate of auditory selective spatial attention (Gamble and Luck, 2011). Generally, the SFS is well-known as auditory spatial region of dorsofrontal cortex. This region has been demonstrated in many studies to be involved in sound localization (e.g., Weeks et al., 2000; Alain et al., 2001; Zatorre et al., 2002; Lewald et al., 2008; Zündorf et al., 2016) and has been assigned to the auditory posterodorsal (spatial) pathway (Arnott et al., 2004). However, it has to be noted that the activity increase revealed here was located at a more anterior position of SFS, compared with areas described in the studies cited above, which reported positions in caudal SFS. Because of the low spatial resolution of the electrical imaging method used here, any clear-cut conclusions on the location of activity in specific subareas of DLPFC might be difficult to draw, and further studies using imaging techniques with higher spatial resolution, such as functional magnetic resonance imaging (fMRI), may have to clarify this issue.

The present results may be related to recent findings by transcranial direct current stimulation (tDCS). In a preceding study using the same “cocktail-party” task as used here, a significant enhancement of the N2 was observed after monopolar anodal tDCS of right pSTG, including PT and auditory cortex (Hanenberg et al., 2019). Also, bilateral monopolar anodal tDCS over this area has been shown to induce clear offline improvements in behavioral performance with this task (Lewald, 2019). These findings have been related to the crucial role of PT in “cocktail-party” sound localization, as had been revealed by fMRI in healthy persons (Zündorf et al., 2013) and voxel-based lesion-behavior mapping analyses in stroke patients (Zündorf et al., 2014). The PT may represent an initial stage of auditory spatial processing within the hierarchically organized posterodorsal cortical pathway, channeling information to frontoparietal areas for further

analyses (Griffiths and Warren, 2002; Krumbholz et al., 2005), including those relevant for selective spatial attention (Zündorf et al., 2014). In general alignment with this view, Hananberg et al. (2019) found a reduction of activity in right IPL at the time of the enhanced N2 after anodal tDCS over right pSTG. The IPL is connected with ipsilateral DLPFC via dorsal components of the superior longitudinal fasciculus (SLF), a white-matter bundle that is crucially involved in functions of spatial orienting and awareness, as well as attentional control (Doricchi and Tomaiuolo, 2003; Makris et al., 2005; Thiebaut de Schotten et al., 2005, 2008; Doricchi et al., 2008; Bernal and Altman, 2010; Suchan et al., 2014; Nakajima et al., 2020). Zündorf et al. (2014) showed that left-sided lesions of the SLF were associated with deficits in the “cocktail-party” task, suggesting an important role of this structure and its frontal target areas in auditory selective spatial attention. Thus, the DLPFC might be part of a temporo-parieto-frontal network concerned with auditory functions subserving “cocktail-party” listening.

The specific post-intervention increase in activity found after audiovisual-congruency training suggests that the related processes in DLPFC were specifically strengthened by bimodal stimulation. This result may be compatible with previous findings indicating multi- or supramodal properties of the dorsofrontal networks that have been usually associated with selective spatial attention in the visual modality (e.g., Slotnick and Moo, 2006; Macaluso, 2010; Lee et al., 2012; Bharadwaj et al., 2014; Lewald et al., 2018). For the monkey DLPFC, it has been suggested that neuronal processes exist for visual and auditory location information and spatial working memory (Fuster et al., 2000; Kikuchi-Yorioka and Sawaguchi, 2000; Artchakov et al., 2007; Hwang and Romanski, 2015; for review, see Plakke and Romanski, 2016), and the human DLPFC has been shown to be involved in transforming auditory and visual inputs into multimodal spatial representations that can be used to guide saccades (Tark and Curtis, 2013). The monkey DLPFC receives projections from posterior auditory cortex areas known to be involved in spatial processing and from the posterior parietal cortex (Chavis and Pandya, 1976; Rauschecker et al., 1995; Romanski et al., 1999a,b). The latter area, which also has reciprocal connections with posterior auditory cortex, has been shown to be critically concerned with auditory and visual spatial processing in human and non-human primates (e.g., Mishkin et al., 1983; Mazzoni et al., 1996; Bushara et al., 1999; Romanski et al., 1999b; Lewald et al., 2002, 2004, 2016). Thus, in conclusion, we assume that the present finding of activity enhancement in BA 9 induced by repetitive processing of spatially and temporally congruent audiovisual stimuli during training may be related to the auditory-visual bimodal properties of the dorsal attention network composed of the DLPFC region and its connections with posterior parietal areas via SLF.

Audiovisual-Congruency as a Key Factor for Training Effects

An effect of training on the N2 was found exclusively for the audiovisual-congruency condition. This result may be in alignment with the multitude of studies on audiovisual

integration, which have demonstrated bimodal enhancement by spatiotemporal alignment using several behavioral (Recanzone, 1998; Lewald et al., 2001; Lewald, 2002b; Lewald and Guski, 2003; Lovelace et al., 2003; Alais and Burr, 2004; Pages and Groh, 2013) and neurophysiological approaches (Besle et al., 2004; Stekelenburg and Vroomen, 2007, 2012; Talsma et al., 2007; Santangelo et al., 2008; Stein and Stanford, 2008; for review, see Stein and Meredith, 1993). Also, few studies in animals and humans have already demonstrated at the behavioral level that spatiotemporally congruent audiovisual stimulation can be used to improve accuracy of localization of single sound sources (Strelnikov et al., 2011; Isaiah et al., 2014; Berger et al., 2018; Cai et al., 2018; Kumpik et al., 2019). Even though we failed to find specific effects of audiovisual-congruency training at the behavioral level, the present study extended these previous approaches by showing that electrophysiological correlates of audiospatial attention in the presence of multiple distractor sources were enhanced by this type of training, while no effect was observed for visual-feedback training.

We assume that this result is related to the experience of phenomenal causality of auditory and visual events (i.e., the impression of a common cause) during audiovisual-congruency training, as typically occurs in the ventriloquism effect (Lewald and Guski, 2003). It is important to note that such binding phenomena do not require complex stimuli, with a highly compelling, meaningful association of auditory and visual information. Rather, simple light spots and tone bursts have been shown to be sufficient to induce audiovisual binding if presented in close spatiotemporal proximity (e.g., Thomas, 1941; Bertelson and Radeau, 1981; Radeau and Bertelson, 1987; Lewald et al., 2001; Slutsky and Recanzone, 2001; Lewald and Guski, 2003). In the audiovisual-congruency training used here, light flashes and target animal vocalizations were presented with identical stimulus onset and duration from roughly the same location. Participants may have experienced binding of auditory and visual events, if distractors were successfully suppressed by the occurrence of the “cocktail-party” effect. That is, this type of training may have induced neural processes resulting in more effective distractor suppression and, thus, more accurate representation of auditory targets in the presence of distractor sources. These processes can be described in terms of short-term neural plasticity based on the ventriloquism effect, as has been discussed in the context of the normally occurring continuous calibration of auditory space by visual experience or its counterpart, the ventriloquism after-effect, in which repetitive or trial-wise presentation of synchronized auditory and visual stimuli with consistent spatial disparity shifts the representation of auditory space relative to the visual space (Recanzone, 1998; Lewald, 2002b). In an EEG source-imaging study focused on the neural basis of the ventriloquism aftereffect, Park and Kayser (2020) recently reported that prolonged exposure to consistent auditory-visual discrepancies recruits, in addition to sensory (occipital and temporal) cortices and multisensory parietal areas, prefrontal regions, including inferior frontal, middle frontal, and superior frontal gyri about 240 ms after stimulus onset. This finding could, potentially, be related to the enhancement of DLPFC activity found here

after audiovisual-congruency training at the time of the N2. Taken together, spatiotemporal congruency of auditory and visual stimuli during training appears to be a key feature enhancing neural processes of auditory selective spatial attention. Our results suggested that this training-induced short-term plasticity occurs particularly in the DLPFC region at the time of the N2 component of the ERP.

Training-Induced Effects on N1 Amplitude and P2 Latency

Only minor and rather non-specific post-intervention changes were observed for N1 and P2 components, and no consistent effects at all for the P1 component and the LPC. The N1 amplitude was generally reduced in the post-training-blocks compared with 1-h-post-blocks, independently of the training condition. In terms of learning, changes in N1 amplitude are often referred to as early automatic stimulus processing (Näätänen and Picton, 1987; Lange, 2013) depending on attentional phenomena (Hillyard et al., 1973; Eimer, 2014; for review, see Luck et al., 2000; McEvoy et al., 2001). Increased N1 amplitudes have been associated with attention catching properties of auditory stimuli and task difficulty (Näätänen et al., 2011). On the other hand, decreased N1 amplitudes could indicate less attentional effort after training due to improved early processing of the stimuli, task familiarity, or the participants' impression that the task was less demanding after already having performed it (Tallus et al., 2015). Furthermore, the N1 typically decreases with repeated stimulation (Näätänen and Picton, 1987), but also recovers with time. This might be reflected by the result that N1 amplitudes were comparably strong in the pre-training and 1-h-post-blocks, but reduced in the post-block. Neural refractoriness as well as adaptation has been suggested as underlying mechanisms of this amplitude attenuation in N1 (Budd et al., 1998; Rosburg et al., 2010). These factors could also be relevant for our findings, given that the participants performed the test blocks three times per session.

P2 latencies were specifically shortened in the younger group 1 h after audiovisual-congruency training, but not for the other conditions, the older group, and the post-training block. On the one hand, this result suggests an accelerated occurrence of P2-related processes after this type of training. On the other hand, it seems difficult to interpret since one may generally expect stronger effects of training on ERPs immediately after the interventions, rather than delayed by 1 h. As the analyses also revealed a general shortening of P2 latencies in the 1-h-post-blocks, compared with post-blocks, it seems likely that an unspecific effect occurred that was due to the repeated execution of the task, independently of the type of intervention. To which extent such unspecific effects have superimposed potential effects specific to the audiovisual-congruency training remains unclear.

If one compares the N1 and P2 waves in the two age groups (Figure 2), it seems as if older adults have overall larger N1 amplitude and reduced P2 amplitude. This could be due to a longer processing negativity overlapping the N1 and P2 wave, which might explain why the N1 wave appears broader in older

adults and why older adults show an overall more negative amplitude between 80 and 220 ms.

Limitations

Unlike the clear-cut electrophysiological result, task performance was found to be unspecifically improved after all training interventions and independently of age. This outcome was probably due to the multiple repetition of the task, as is often found in research on learning and memory (Ebbinghaus, 1885). That we failed to find specific training effects could, possibly, be a result of ceiling effects (Brungart and Simpson, 2007) after repeated training, as were found previously in a similar task (Hananberg et al., 2019). However, given mean rates of correct responses of about 84% in the younger and 70% in the older group, ceiling effects should not play a major role here. Alternatively, the effect of the audiovisual-congruency training observed for the N2 amplitude could be confined to specific sub-processes required to solve a “cocktail-party” speech localization task. Assuming that the N2 especially reflects cognitive control processes mainly related to the inhibition of task-relevant information (as argued above), the audiovisual-congruency training might have enhanced specific cognitive control processes (reflected by the increase in N2 amplitude), rather than speech-in-noise localization in general.

This study left open the question of whether more extended and more long-lasting improvements in task performance in “cocktail-party” scenarios could be achieved by training in both, younger and older adults. Motivation has been shown to have a significant impact on task performance (Green and Seitz, 2015), and future studies offering more attractive training paradigms, such as game-based training, should be considered here. Recent work on this topic has shown that various age groups profit from action video game training, showing enhanced performances in task switching abilities after playing for 3 weeks (Basak et al., 2008; Strobach et al., 2012; Wang et al., 2016). Recently, Schuchert and Lewald (2020), using a similar “cocktail-party” task as used here, demonstrated that both audio action game training and video non-action game training improved auditory selective spatial attention in younger adults. The present results thus suggest that a bimodal (audio-visual synchronous) game training may also be promising in this respect. It has, however, to be noted that other training approaches, such as music training have also been shown to have beneficial effects on auditory processing in acoustically complex situations, especially in older adults. For example, Zendel et al. (2019) recently found for a group of older adults that learning to play the piano improved understanding words presented in loud background noise after 6 months. Also, a 10-week participation in a choir was shown to improve speech-in-noise perception, pitch discrimination ability, and the strength of the neural representation of speech fundamental frequency, suggesting that short-term choir participation may be considered as an effective intervention for mitigating age-related hearing losses (Dubinsky et al., 2019). Future studies will show which of these quite different training approaches is best suited to improve hearing performance of older people in everyday situations requiring auditory selective spatial attention.

CONCLUSION

In summary, the present study showed that short-term audiovisual-congruency training, but not visual-feedback training and a control condition, enhanced the N2 component in a multiple-speaker target localization task. The increase in N2 was associated with an increase of electrical activity in DLPFC and may indicate enhancement of neural processes of auditory selective spatial attention. Both effects were observed in younger, but not older, participants. Further experiments are necessary in order to examine whether more intensive, longer lasting and more realistic audio-visual training settings are suitable to obtain improvements also in behavioral measures.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Committee of the Leibniz Research Centre for Working Environment and Human Factors, Dortmund. The participants provided their written informed consent to participate in this study.

REFERENCES

- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., and Grady, C. L. (2001). 'What' and 'where' in the human auditory system. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12301–12306.
- Alain, C., Campeanu, S., and Tremblay, K. (2010). Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *J. Cogn. Neurosci.* 22, 392–403. doi: 10.1162/jocn.2009.21279
- Alain, C., Snyder, J. S., He, Y., and Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cereb. Cortex* 17, 1074–1084. doi: 10.1093/cercor/bhl018
- Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262. doi: 10.1016/j.cub.2004.01.029
- Anderer, P., Semlitsch, H. V., and Saletu, B. (1996). Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. *Electroencephalogr. Clin. Neurophysiol.* 99, 458–472. doi: 10.1016/s0013-4694(96)96518-9
- Arnott, S. R., Binns, M. A., Grady, C. L., and Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *NeuroImage* 22, 401–408. doi: 10.1016/j.neuroimage.2004.01.014
- Artchakov, D., Tikhonravov, D., Vuontela, V., Linnankoski, I., Korvenoja, A., and Carlson, S. (2007). Processing of auditory and visual location information in the monkey prefrontal cortex. *Exp. Brain Res.* 180, 469–479. doi: 10.1007/s00221-007-0873-8
- Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765–777. doi: 10.1037/a0013494
- Berger, C. C., Gonzalez-Franco, M., Tajadura-Jiménez, A., Florencio, D., and Zhang, Z. (2018). Generic HRTFs may be good enough in virtual reality.

AUTHOR CONTRIBUTIONS

CH, SG, and JL contributed to conception and design of the study. CH collected the data and wrote the first draft of the manuscript. CH and JL wrote sections of the manuscript. JL supervised the project and takes final responsibility for this article. All authors contributed to the analyses, manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by the German Federal Ministry of Education and Research in the framework of the TRAIN-STIM project (Grant number 01GQ1424E).

ACKNOWLEDGMENTS

The authors wish to thank Jonas Heyermann and Stefan Weber for help in running the experiments, and Tobias Blanke and Peter Dillmann for preparing software and parts of the electronic equipment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2021.645702/full#supplementary-material>

- Improving source localization through cross-modal plasticity. *Front. Neurosci.* 12:21. doi: 10.3389/fnins.2018.00021
- Bernal, B., and Altman, N. (2010). The connectivity of the superior longitudinal fasciculus: a tractography DTI study. *Magn. Reson. Imaging* 28, 217–225. doi: 10.1016/j.mri.2009.07.008
- Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.* 29, 578–584. doi: 10.3758/bf03207374
- Bertoli, S., Smurzynski, J., and Probst, R. (2005). Effects of age, age-related hearing loss, and contralateral cafeteria noise on the discrimination of small frequency changes: psychoacoustic and electrophysiological measures. *J. Assoc. Res. Otolaryngol.* 6, 207–222. doi: 10.1007/s10162-005-5029-6
- Besle, J., Fort, A., Delpuech, C., and Giard, M.-H. (2004). Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur. J. Neurosci.* 20, 2225–2234. doi: 10.1111/j.1460-9568.2004.03670.x
- Bharadwaj, H. M., Lee, A. K. C., and Shinn-Cunningham, B. G. (2014). Measuring auditory selective attention using frequency tagging. *Front. Integr. Neurosci.* 8:6. doi: 10.3389/fnint.2014.00006
- Bolognini, N., Rasi, F., Coccia, M., and Ladavas, E. (2005). Visual search improvement in hemianopic patients after audio-visual stimulation. *Brain* 128, 2830–2842. doi: 10.1093/brain/awh656
- Bosen, A. K., Fleming, J. T., Allen, P. D., O'Neill, W. E., and Paige, G. D. (2018). Multiple time scales of the ventriloquism aftereffect. *PLoS One* 13:e0200930. doi: 10.1371/journal.pone.0200930
- Braga, R. M., Fu, R. Z., Seemungal, B. M., Wise, R. J. S., and Leech, R. (2016). Eye movements during auditory attention predict individual differences in dorsal attention network activity. *Front. Hum. Neurosci.* 10:164. doi: 10.3389/fnhum.2016.00164

- Braga, R. M., Hellyer, P. J., Wise, R. J. S., and Leech, R. (2017). Auditory and visual connectivity gradients in frontoparietal cortex: frontoparietal audiovisual gradients. *Hum. Brain Mapp.* 38, 255–270. doi: 10.1002/hbm.23358
- Brainard, M. S., and Knudsen, E. I. (1993). Experience-dependent plasticity in the inferior colliculus: a site for visual calibration of the neural representation of auditory space in the barn owl. *J. Neurosci.* 13, 4589–4608. doi: 10.1523/jneurosci.13-11-04589.1993
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attent. Percept. Psychophys.* 77, 1465–1487. doi: 10.3758/s13414-015-0882-9
- Brungart, D. S., and Simpson, B. D. (2007). Cocktail party listening in a dynamic multitalker environment. *Percept. Psychophys.* 69, 79–91. doi: 10.3758/bf03194455
- Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., and Michie, P. T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness. *Int. J. Psychophysiol.* 31, 51–68. doi: 10.1016/s0167-8760(98)00040-3
- Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.* 2, 759–766. doi: 10.1038/11239
- Cai, Y., Chen, G., Zhong, X., Yu, G., Mo, H., Jiang, J., et al. (2018). Influence of audiovisual training on horizontal sound localization and its related ERP response. *Front. Hum. Neurosci.* 12:423. doi: 10.3389/fnhum.2018.00423
- Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J. Exp. Psychol.* 84, 141–147. doi: 10.1037/h0028925
- Canon, L. K. (1971). Directed attention and maladaptive ‘adaptation’ to displacement of the visual field. *J. Exp. Psychol.* 88, 403–408. doi: 10.1037/h0030878
- Chavis, D. A., and Pandya, D. N. (1976). Further observations on corticofrontal connections in the rhesus monkey. *Brain Res.* 117, 369–386. doi: 10.1016/0006-8993(76)90089-5
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229
- Covey, T. J., Shucard, J. L., Benedict, R. H., Weinstock-Guttman, B., and Shucard, D. W. (2018). Improved cognitive performance and event-related potential changes following working memory training in patients with multiple sclerosis. *Mult. Scler. J. Exp. Transl. Clin.* 4:2055217317747626.
- Covey, T. J., Shucard, J. L., and Shucard, D. W. (2019). Working memory training and perceptual discrimination training impact overlapping and distinct neurocognitive processes: evidence from event-related potentials and transfer of training gains. *Cognition* 182, 50–72. doi: 10.1016/j.cognition.2018.08.012
- Daffner, K. R., Alperin, B. R., Mott, K. K., Tusch, E. S., and Holcomb, P. J. (2015). Age-related differences in early novelty processing: using PCA to parse the overlapping anterior P2 and N2 components. *Biol. Psychol.* 105, 83–94. doi: 10.1016/j.biopsycho.2015.01.002
- Dennis, T. A., and Chen, C. C. (2007). Neurophysiological mechanisms in the emotional modulation of attention: the interplay between threat sensitivity and attentional control. *Biol. Psychol.* 76, 1–10. doi: 10.1016/j.biopsycho.2007.05.001
- Doricchi, F., Thiebaut de Schotten, M., Tomaiuolo, F., and Bartolomeo, P. (2008). White matter (dis)connections and gray matter (dys)functions in visual neglect: gaining insights into the brain networks of spatial awareness. *Cortex* 44, 983–995. doi: 10.1016/j.cortex.2008.03.006
- Doricchi, F., and Tomaiuolo, F. (2003). The anatomy of neglect without hemianopia: a key role for parietal–frontal disconnection? *Neuroreport* 14, 2239–2243. doi: 10.1097/00001756-200312020-00021
- Dubinsky, E., Wood, E. A., Nespole, G., and Russo, F. A. (2019). Short-term choir singing supports speech-in-noise perception and neural pitch strength in older adults with age-related hearing loss. *Front. Neurosci.* 13:1153. doi: 10.3389/fnhum.2019.01153
- Ebbinghaus, H. (1885). *Über das Gedächtnis: Untersuchungen zur Experimentellen Psychologie*. Berlin: Duncker & Humblot.
- Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends Cogn. Sci.* 18, 526–535. doi: 10.1016/j.tics.2014.05.005
- Eldar, S., and Bar-Haim, Y. (2010). Neural plasticity in response to attention training in anxiety. *Psychol. Med.* 40, 667–677. doi: 10.1017/S0033291709990766
- Falkenstein, M., Hoormann, J., and Hohnsbein, J. (2002). Inhibition-related ERP components: variation with modality, age, and time-on-task. *J. Psychophysiol.* 16, 167–175. doi: 10.1027//0269-8803.16.3.167
- Feierabend, M., Karnath, H.-O., and Lewald, J. (2019). Auditory space perception in the blind: horizontal sound localization in acoustically simple and complex situations. *Perception* 48, 1039–1057. doi: 10.1111/0301006619872062
- Folstein, J. R., and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170.
- Fuster, J. M., Bodner, M., and Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature* 405, 347–351. doi: 10.1038/35012613
- Gajewski, P. D., and Falkenstein, M. (2012). Training-induced improvement of response selection and error detection in aging assessed by task switching: effects of cognitive, physical, and relaxation training. *Front. Hum. Neurosci.* 6:130. doi: 10.3389/fnhum.2012.00130
- Gamble, M. L., and Luck, S. J. (2011). N2ac: an ERP component associated with the focusing of attention within an auditory scene: the N2ac component. *Psychophysiology* 48, 1057–1068. doi: 10.1111/j.1469-8986.2010.01172.x
- Gazzaley, A., and D’Esposito, M. (2007). Top-down modulation and normal aging. *Ann. N. Y. Acad. Sci.* 1097, 67–83. doi: 10.1196/annals.1379.010
- Getzmann, S., Falkenstein, M., and Wascher, E. (2015a). ERP correlates of auditory goal-directed behavior of younger and older adults in a dynamic speech perception task. *Behav. Brain Res.* 278, 435–445. doi: 10.1016/j.bbr.2014.10.026
- Getzmann, S., Hananberg, C., Lewald, J., Falkenstein, M., and Wascher, E. (2015b). Effects of age on electrophysiological correlates of speech processing in a dynamic ‘cocktail-party’ situation. *Front. Neurosci.* 9:341. doi: 10.3389/fnhum.2015.00341
- Getzmann, S., Wascher, E., and Falkenstein, M. (2015c). What does successful speech-in-noise perception in aging depend on? Electrophysiological correlates of high and low performance in older adults. *Neuropsychologia* 70, 43–57. doi: 10.1016/j.neuropsychologia.2015.02.009
- Getzmann, S., Lewald, J., and Falkenstein, M. (2014). Using auditory pre-information to solve the cocktail-party problem: electrophysiological evidence for age-specific differences. *Front. Neurosci.* 8:413. doi: 10.3389/fnhum.2014.00413
- Grasso, P. A., Ládavas, E., and Bertini, C. (2016). Compensatory recovery after multisensory stimulation in hemianopic patients: behavioral and neurophysiological components. *Front. Syst. Neurosci.* 10:45. doi: 10.3389/fnsys.2016.00045
- Gratton, G., Coles, M. G., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–484. doi: 10.1016/0013-4694(83)90135-9
- Green, C. S., and Seitz, A. R. (2015). The impacts of video games on cognition (and how the government can guide the industry). *Policy Insights Behav. Brain Sci.* 2, 101–110. doi: 10.1177/2372732215601121
- Griffiths, T. D., and Warren, J. D. (2002). The planum temporale as a computational hub. *Trends Neurosci.* 25, 348–353. doi: 10.1016/s0166-2236(02)02191-4
- Hananberg, C., Getzmann, S., and Lewald, J. (2019). Transcranial direct current stimulation of posterior temporal cortex modulates electrophysiological correlates of auditory selective spatial attention in posterior parietal cortex. *Neuropsychologia* 131, 160–170. doi: 10.1016/j.neuropsychologia.2019.05.023
- Hasher, L., and Zacks, R. T. (1988). “Working memory, comprehension, and aging: a review and a new view,” in *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 22, ed. G. H. Bower (Cambridge, MA: Academic Press), 193–225. doi: 10.1016/S0079-7421(08)60041-9
- Held, R. (1955). Shifts in binaural localization after prolonged exposures to atypical combinations of stimuli. *Am. J. Psychol.* 68, 526–548. doi: 10.2307/1418782
- Helmholtz, H. V. (1867). *Handbuch der Physiologischen Optik*. Bonn: Voss.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science* 182, 177–180. doi: 10.1126/science.182.4108.177
- Howard, I. P., and Templeton, W. B. (1966). *Human Spatial Orientation*. Hoboken, NJ: Wiley.

- Hwang, J., and Romanski, L. M. (2015). Prefrontal neuronal responses during audiovisual mnemonic processing. *J. Neurosci.* 35, 960–971. doi: 10.1523/jneurosci.1328-14.2015
- Hyde, P. S., and Knudsen, E. I. (2000). Topographic projection from the optic tectum to the auditory space map in the inferior colliculus of the barn owl. *J. Comp. Neurol.* 421, 146–160. doi: 10.1002/(sici)1096-9861(20000529)421:2<146::aid-cne2>3.0.co;2-5
- Hyde, P. S., and Knudsen, E. I. (2002). The optic tectum controls visually guided adaptive plasticity in the owl's auditory space map. *Nature* 415, 73–76. doi: 10.1038/415073a
- Isaiah, A., Vongpaisal, T., King, A. J., and Hartley, D. E. H. (2014). Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J. Neurosci.* 34, 11119–11130. doi: 10.1523/jneurosci.4767-13.2014
- Jack, C. E., and Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the 'ventriloquism' effect. *Percept. Mot. Skills* 37, 967–979. doi: 10.2466/pms.1973.37.3.967
- Jackson, C. V. (1953). Visual factors in auditory localization. *Q. J. Exp. Psychol.* 5, 52–65. doi: 10.1080/17470215308416626
- Kalil, R., and Freedman, S. J. (1967). Compensation for auditory re-arrangement in the absence of observer movement. *Percept. Mot. Skills* 24, 475–478. doi: 10.2466/pms.1967.24.2.475
- Kikuchi-Yorioka, Y., and Sawaguchi, T. (2000). Parallel visuospatial and audiopsatial working memory processes in the monkey dorsolateral prefrontal cortex. *Nat. Neurosci.* 3, 1075–1076. doi: 10.1038/80581
- Klemm, O. (1910). Lokalisation von Sinneseindrücken bei disparaten Nebenreizen. *Psychol. Stud.* 5, 73–162.
- Knudsen, E. I. (1999). Mechanisms of experience-dependent plasticity in the auditory localization pathway of the barn owl. *J. Comp. Physiol. A* 185, 305–321. doi: 10.1007/s003590050391
- Knudsen, E. I., and Knudsen, P. F. (1985). Vision guides the adjustment of auditory localization in young barn owls. *Science* 230, 545–548. doi: 10.1126/science.4048948
- Knudsen, E. I., and Knudsen, P. F. (1989). Vision calibrates sound localization in developing barn owls. *J. Neurosci.* 9, 3306–3313. doi: 10.1523/jneurosci.09-09-03306.1989
- Kong, L., Michalka, S. W., Rosen, M. L., Sheremata, S. L., Swisher, J. D., Shinn-Cunningham, B. G., et al. (2014). Auditory spatial attention representations in the human cerebral cortex. *Cereb. Cortex* 24, 773–784. doi: 10.1093/cercor/bhs359
- Krumbholz, K., Schönwiesner, M., Rübsem, R., Zilles, K., Fink, G. R., and von Cramon, D. Y. (2005). Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *Eur. J. Neurosci.* 21, 230–238. doi: 10.1111/j.1460-9568.2004.03836.x
- Kuk, F., Keenan, D. M., Lau, C., Crose, B., and Schumacher, J. (2014). Evaluation of a localization training program for hearing impaired listeners. *Ear Hear.* 35, 652–666. doi: 10.1097/aud.0000000000000067
- Kumpik, D. P., Campbell, C., Schnupp, J. W. H., and King, A. J. (2019). Re-weighting of sound localization cues by audiovisual training. *Front. Neurosci.* 13:1164. doi: 10.3389/fnins.2019.01164
- Lange, K. (2013). The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Front. Hum. Neurosci.* 7:263. doi: 10.3389/fnhum.2013.00263
- Lee, A. K. C., Larson, E., and Maddox, R. K. (2012). Mapping cortical dynamics using simultaneous MEG/EEG and anatomically-constrained minimum-norm estimates: an auditory attention example. *J. Vis. Exp.* 68:e4262.
- Lee, A. K. C., Larson, E., Maddox, R. K., and Shinn-Cunningham, B. G. (2014). Using neuroimaging to understand the cortical mechanisms of auditory selective attention. *Hear. Res.* 307, 111–120. doi: 10.1016/j.heares.2013.06.010
- Lewald, J. (2002a). Opposing effects of head position on sound localization in blind and sighted human subjects. *Eur. J. Neurosci.* 15, 1219–1224. doi: 10.1046/j.1460-9568.2002.01949.x
- Lewald, J. (2002b). Rapid adaptation to auditory-visual spatial disparity. *Learn. Mem.* 9, 268–278. doi: 10.1101/lm.51402
- Lewald, J. (2002c). Vertical sound localization in blind humans. *Neuropsychologia* 40, 1868–1872. doi: 10.1016/s0028-3932(02)00071-4
- Lewald, J. (2013). Exceptional ability of blind humans to hear sound motion: implications for the emergence of auditory space. *Neuropsychologia* 51, 181–186. doi: 10.1016/j.neuropsychologia.2012.11.017
- Lewald, J. (2016). Modulation of human auditory spatial scene analysis by transcranial direct current stimulation. *Neuropsychologia* 84, 282–293. doi: 10.1016/j.neuropsychologia.2016.01.030
- Lewald, J. (2019). Bihemispheric anodal transcranial direct-current stimulation over temporal cortex enhances auditory selective spatial attention. *Exp. Brain Res.* 237, 1539–1549. doi: 10.1007/s00221-019-05525-y
- Lewald, J., Ehrenstein, W. H., and Guski, R. (2001). Spatio-temporal constraints for auditory-visual integration. *Behav. Brain Res.* 121, 69–79. doi: 10.1016/s0166-4328(00)00386-7
- Lewald, J., Foltys, H., and Töpper, R. (2002). Role of the posterior parietal cortex in spatial hearing. *J. Neurosci.* 22:RC207.
- Lewald, J., and Getzmann, S. (2015). Electrophysiological correlates of cocktail-party listening. *Behav. Brain Res.* 292, 157–166. doi: 10.1016/j.bbr.2015.06.025
- Lewald, J., and Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Brain Res. Cogn. Brain Res.* 16, 468–478. doi: 10.1016/s0926-6410(03)00074-0
- Lewald, J., Hananberg, C., and Getzmann, S. (2016). Brain correlates of the orientation of auditory spatial attention onto speaker location in a 'cocktail-party' situation. *Psychophysiology* 53, 1484–1495. doi: 10.1111/psyp.12692
- Lewald, J., and Hausmann, M. (2013). Effects of sex and age on auditory spatial scene analysis. *Hear. Res.* 299, 46–52. doi: 10.1016/j.heares.2013.02.005
- Lewald, J., Kentridge, R. W., Peters, S., Tegenthoff, M., Heywood, C. A., and Hausmann, M. (2013). Auditory-visual localization in hemianopia. *Neuropsychology* 27, 573–582. doi: 10.1037/a0033451
- Lewald, J., Peters, S., Tegenthoff, M., and Hausmann, M. (2009a). Dissociation of auditory and visual straight ahead in hemianopia. *Brain Res.* 1287, 111–117. doi: 10.1016/j.brainres.2009.06.085
- Lewald, J., Peters, S., Tegenthoff, M., and Hausmann, M. (2009b). Distortion of auditory space in hemianopia. *Eur. J. Neurosci.* 30, 1401–1411. doi: 10.1111/j.1460-9568.2009.06905.x
- Lewald, J., Riederer, K. A. J., Lentz, T., and Meister, I. G. (2008). Processing of sound location in human cortex. *Eur. J. Neurosci.* 27, 1261–1270. doi: 10.1111/j.1460-9568.2008.06094.x
- Lewald, J., Schlüter, M.-C., and Getzmann, S. (2018). Cortical processing of location changes in a "cocktail-party" situation: spatial oddball effects on electrophysiological correlates of auditory selective attention. *Hear. Res.* 365, 49–61. doi: 10.1016/j.heares.2018.04.009
- Lewald, J., Tegenthoff, M., Peters, S., and Hausmann, M. (2012). Passive auditory stimulation improves vision in hemianopia. *PLoS One* 7:e31603. doi: 10.1371/journal.pone.0031603
- Lewald, J., Wienemann, M., and Borojerdi, B. (2004). Shift in sound localization induced by rTMS of the posterior parietal lobe. *Neuropsychologia* 42, 1598–1607. doi: 10.1016/j.neuropsychologia.2004.04.012
- Lovelace, C. T., Stein, B. E., and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res. Cogn. Brain Res.* 17, 447–453. doi: 10.1016/s0926-6410(03)00160-5
- Luck, S. J., Woodman, G. F., and Vogel, E. K. (2000). Event-related potential studies of attention. *Trends Cog. Sci.* 4, 432–440.
- Macaluso, E. (2010). Orienting of spatial attention and the interplay between the senses. *Cortex* 46, 282–297. doi: 10.1016/j.cortex.2009.05.010
- Majdak, P., Walder, T., and Laback, B. (2013). Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions. *J. Acoust. Soc. Am.* 134, 2148–2159. doi: 10.1121/1.4816543
- Makris, N., Kennedy, D. N., McInerney, S., Sorensen, A. G., Wang, R., Caviness, V. S. Jr., et al. (2005). Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb. Cortex* 15, 854–869. doi: 10.1093/cercor/bhh186
- Marcell, M. M., Borella, D., Greene, M., Kerr, E., and Rogers, S. (2000). Confrontation naming of environmental sounds. *J. Clin. Exp. Neuropsychol.* 22, 830–864. doi: 10.1076/jcen.22.6.830.949
- Martin, B. A., Tremblay, K. L., and Korczak, P. (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear Hear.* 29, 285–313. doi: 10.1097/aud.0b013e3181662c0e

- Mazzoni, P., Bracewell, R. M., Barash, S., and Andersen, R. A. (1996). Spatially tuned auditory responses in area LIP of macaques performing delayed memory saccades to acoustic targets. *J. Neurophysiol.* 75, 1233–1241. doi: 10.1152/jn.1996.75.3.1233
- McEvoy, L. K., Pellouchoud, E., Smith, M. E., and Gevins, A. (2001). Neurophysiological signals of working memory in normal aging. *Brain Res. Cogn. Brain Res.* 11, 363–376. doi: 10.1016/s0926-6410(01)00009-x
- Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6, 414–417. doi: 10.1016/0166-2236(83)90190-x
- Näätänen, R., Kujala, T., and Winkler, I. (2011). Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology* 48, 4–22. doi: 10.1111/j.1469-8986.2010.01114.x
- Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x
- Nakai, T., Kato, C., and Matsuo, K. (2005). An fMRI study to investigate auditory attention: a model of the cocktail party phenomenon. *Magn. Reson. Med. Sci.* 4, 75–82. doi: 10.2463/mrms.4.75
- Nakajima, R., Kinoshita, M., Shinohara, H., and Nakada, M. (2020). The superior longitudinal fascicle: reconsidering the fronto-parietal neural network based on anatomy and function. *Brain Imaging Behav.* 14, 2817–2830. doi: 10.1007/s11682-019-00187-4
- Olfers, K. J. F., and Band, G. P. H. (2018). Game-based training of flexibility and attention improves task-switch performance: near and far transfer of cognitive training in an EEG study. *Psychol. Res.* 82, 186–202. doi: 10.1007/s00426-017-0933-z
- Pages, D. S., and Groh, J. M. (2013). Looking at the ventriloquist: visual outcome of eye movements calibrates sound localization. *PLoS One* 8:e72562. doi: 10.1371/journal.pone.0072562
- Park, H., and Kayser, C. (2020). The neurophysiological basis of short- and long-term ventriloquism aftereffects. *bioRxiv* [Preprint]. doi: 10.1101/2020.06.16.154161
- Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24(Suppl. D), 5–12.
- Passamonti, C., Bertini, C., and Lådavas, E. (2009). Audio-visual stimulation improves oculomotor patterns in patients with hemianopia. *Neuropsychologia* 47, 546–555. doi: 10.1016/j.neuropsychologia.2008.10.008
- Pichora-Fuller, M. K., Alain, C., and Schneider, B. A. (2017). “Older adults at the cocktail party,” in *The Auditory System at the Cocktail Party*. Springer Handbook of Auditory Research, Vol. 60, eds J. C. Middlebrooks, J. Z. Simon, A. N. Popper, and R. R. Fay (Berlin: Springer), 227–259. doi: 10.1007/978-3-319-51662-2_9
- Plakke, B., and Romanski, L. M. (2016). Neural circuits in auditory and audiovisual memory. *Brain Res.* 1640, 278–288. doi: 10.1016/j.brainres.2015.11.042
- Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain Cogn.* 56, 5–13. doi: 10.1016/j.bandc.2004.03.006
- Pugh, K. R., Offywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Byrd, D., Skudlarski, P., et al. (1996). Auditory selective attention: an fMRI investigation. *NeuroImage* 4(3 Pt 1), 159–173.
- Radeau, M. (1994). Auditory-visual spatial interaction and modularity. *Curr. Psychol. Cogn.* 13, 3–51.
- Radeau, M., and Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. *Percept. Psychophys.* 22, 137–146. doi: 10.3758/bf03198746
- Radeau, M., and Bertelson, P. (1978). Cognitive factors and adaptation to auditory-visual discordance. *Percept. Psychophys.* 23, 341–343. doi: 10.3758/bf03199719
- Radeau, M., and Bertelson, P. (1987). Auditory-visual interaction and the timing of inputs: thomas (1941) revisited. *Psychol. Res.* 49, 17–22. doi: 10.1007/bf00309198
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114. doi: 10.1126/science.7701330
- Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc. Natl. Acad. Sci. U.S.A.* 95, 869–875. doi: 10.1073/pnas.95.3.869
- Romanski, L. M., Bates, J. F., and Goldman-Rakic, P. S. (1999a). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 403, 141–157. doi: 10.1002/(sici)1096-9861(19990111)403:2<141::aid-cne1>3.0.co;2-v
- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999b). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136. doi: 10.1038/16056
- Rosburg, T., Zimmerer, K., and Huonker, R. (2010). Short-term habituation of auditory evoked potential and neuromagnetic field components in dependence of the interstimulus interval. *Exp. Brain Res.* 205, 559–570. doi: 10.1007/s00221-010-2391-3
- Rueda, M., Pozuelos, J., and Cómbita, L. (2015). Cognitive neuroscience of attention. From brain mechanisms to individual differences in efficiency. *AIMS Neurosci.* 2, 183–202. doi: 10.3934/Neuroscience.2015.4.183
- Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiol. Aging* 30, 507–514. doi: 10.1016/j.neurobiolaging.2008.09.023
- Santangelo, V., Van der Lubbe, R. H. J., Olivetti Belardinelli, M., and Postma, A. (2008). Multisensory integration affects ERP components elicited by exogenous cues. *Exp. Brain Res.* 185, 269–277. doi: 10.1007/s00221-007-1151-5
- Schuchert, J. B., and Lewald, J. (2020). Training with audio and video games improves audiospatial performance in a “cocktail-party” task: a controlled intervention study in young adults. *bioRxiv* [Preprint]. doi: 10.1101/2020.11.17.386300
- Shams, L., and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends Cogn. Sci.* 12, 411–417. doi: 10.1016/j.tics.2008.07.006
- Slotnick, S. D., and Moo, L. R. (2006). Prefrontal cortex hemispheric specialization for categorical and coordinate visual spatial memory. *Neuropsychologia* 44, 1560–1568. doi: 10.1016/j.neuropsychologia.2006.01.018
- Slutsky, D. A., and Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport* 12, 7–10. doi: 10.1097/00001756-200101220-00009
- Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press.
- Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266. doi: 10.1038/nrn2331
- Stekelenburg, J. J., and Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J. Cogn. Neurosci.* 19, 1964–1973. doi: 10.1162/jocn.2007.19.12.1964
- Stekelenburg, J. J., and Vroomen, J. (2012). Electrophysiological correlates of predictive coding of auditory location in the perception of natural audiovisual events. *Front. Integr. Neurosci.* 6:26. doi: 10.3389/fnint.2012.00026
- Stratton, G. M. (1896). Some preliminary experiments on vision without inversion of the retinal image. *Psychol. Rev.* 3, 611–617. doi: 10.1037/h0072918
- Stratton, G. M. (1897). Vision without inversion of the retinal image. *Psychol. Rev.* 4, 341–360. doi: 10.1037/h0075482
- Strelnikov, K., Rosito, M., and Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One* 6:e18344. doi: 10.1371/journal.pone.0018344
- Strobach, T., Liepelt, R., Schubert, T., and Kiesel, A. (2012). Task switching: effects of practice on switch and mixing costs. *Psychol. Res.* 76, 74–83. doi: 10.1007/s00426-011-0323-x
- Suchan, J., Umarova, R., Schnell, S., Himmelbach, M., Weiller, C., Karnath, H.-O., et al. (2014). Fiber pathways connecting cortical areas relevant for spatial orienting and exploration: fiber pathways relevant for spatial orienting. *Hum. Brain Mapp.* 35, 1031–1043. doi: 10.1002/hbm.22232
- Tallus, J., Soveri, A., Hämäläinen, H., Tuomainen, J., and Laine, M. (2015). Effects of auditory attention training with the dichotic listening task: behavioural and neurophysiological evidence. *PLoS One* 10:e0139318. doi: 10.1371/journal.pone.0139318
- Talsma, D., Doty, T. J., and Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb. Cortex* 17, 679–690. doi: 10.1093/cercor/bhk016

- Tark, K.-J., and Curtis, C. E. (2013). Deciding where to look based on visual, auditory, and semantic information. *Brain Res.* 1525, 26–38. doi: 10.1016/j.brainres.2013.06.002
- Thiebaut de Schotten, M., Kinkingnehun, S., Delmaire, C., Lehericy, S., Duffau, H., Thivard, L., et al. (2008). Visualization of disconnection syndromes in humans. *Cortex* 44, 1097–1103. doi: 10.1016/j.cortex.2008.02.003
- Thiebaut de Schotten, M., Urbanski, M., Duffau, H., Volle, E., Lévy, R., Dubois, B., et al. (2005). Direct evidence for a parietal-frontal pathway subserving spatial awareness in humans. *Science* 309, 2226–2228. doi: 10.1126/science.1116251
- Thomas, G. J. (1941). Experimental study of the influence of vision on sound localization. *J. Exp. Psychol.* 28, 163–177. doi: 10.1037/h0055183
- Wang, P., Liu, H.-H., Zhu, X.-T., Meng, T., Li, H.-J., and Zuo, X.-N. (2016). Action video game training for healthy adults: a meta-analytic study. *Front. Psychol.* 7:907. doi: 10.3389/fpsyg.2016.00907
- Wascher, E., and Getzmann, S. (2014). Rapid mental fatigue amplifies age-related attentional deficits. *J. Psychophysiol.* 28, 215–224. doi: 10.1027/0269-8803/a000127
- Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C. M., Cohen, L. G., et al. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. *J. Neurosci.* 20, 2664–2672. doi: 10.1523/jneurosci.20-07-02664.2000
- Yang, W., Guo, A., Li, Y., Qiu, J., Li, S., Yin, S., et al. (2018). Audio-visual spatiotemporal perceptual training enhances the P300 component in healthy older adults. *Front. Psychol.* 9:2537. doi: 10.3389/fpsyg.2018.02537
- Yordanova, J., Falkenstein, M., Hohnsbein, J., and Kolev, V. (2004). Parallel systems of error processing in the brain. *NeuroImage* 22, 590–602. doi: 10.1016/j.neuroimage.2004.01.040
- Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., and Tam, C. (2006). Perceptual recalibration in human sound localization: learning to remediate front-back reversals. *J. Acoust. Soc. Am.* 120, 343–359. doi: 10.1121/1.2208429
- Zatorre, R. J., Bouffard, M., Ahad, P., and Belin, P. (2002). Where is 'where' in the human auditory cortex? *Nat. Neurosci.* 5, 905–909. doi: 10.1038/nn904
- Zendel, B. R., de Boysson, C., Mellah, S., Démonet, J. F., and Belleville, S. (2016). The impact of attentional training on event-related potentials in older adults. *Neurobiol. Aging* 47, 10–22. doi: 10.1016/j.neurobiolaging.2016.06.023
- Zendel, B. R., West, G. L., Belleville, S., and Peretz, I. (2019). Musical training improves the ability to understand speech-in-noise in older adults. *Neurobiol. Aging* 81, 102–115. doi: 10.1016/j.neurobiolaging.2019.05.015
- Zheng, W., and Knudsen, E. I. (2001). Gabaergic inhibition antagonizes adaptive adjustment of the owl's auditory space map during the initial phase of plasticity. *J. Neurosci.* 21, 4356–4365. doi: 10.1523/jneurosci.21-12-04356.2001
- Zündorf, I. C., Karnath, H.-O., and Lewald, J. (2014). The effect of brain lesions on sound localization in complex acoustic environments. *Brain* 137, 1410–1418. doi: 10.1093/brain/awu044
- Zündorf, I. C., Lewald, J., and Karnath, H.-O. (2013). Neural correlates of sound localization in complex acoustic environments. *PLoS One* 8:e64259. doi: 10.1371/journal.pone.0064259
- Zündorf, I. C., Lewald, J., and Karnath, H.-O. (2016). Testing the dual-pathway model for auditory processing in human cortex. *NeuroImage* 124, 672–681. doi: 10.1016/j.neuroimage.2015.09.026
- Zwiers, M. P., Van Opstal, A. J., and Cruysberg, J. R. (2001a). A spatial hearing deficit in early-blind humans. *J. Neurosci.* 21, RC142: 1–5.
- Zwiers, M. P., Van Opstal, A. J., and Cruysberg, J. R. (2001b). Two-dimensional sound-localization behavior of early-blind humans. *Exp. Brain Res.* 140, 206–222. doi: 10.1007/s002210100800

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hananberg, Schlüter, Getzmann and Lewald. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.