

COGNITIVE NEUROSCIENCE

A bias-free two-alternative forced choice procedure to examine intersensory illusions applied to the ventriloquist effect by flashes and averted eye-gazes

Jean Vroomen and Jeroen J. Stekelenburg

Department of Cognitive Neuropsychology, Tilburg University, PO Box 90153, 5000 LE, Tilburg, Netherlands

Keywords: human subjects, mismatch negativity, multisensory perception, psychophysics, ventriloquist effect

Abstract

We compared with a new psychophysical method whether flashes and averted eye-gazes of a cartoon face induce a ventriloquist illusion (an illusory shift of the apparent location of a sound by a visual distracter). With standard psychophysical procedures that measure a direct ventriloquist effect and a ventriloquist aftereffect, we found in human subjects that both types of stimuli induced an illusory shift of sound location. These traditional methods, though, are probably contaminated by response strategies. We therefore developed a new two-alternative forced choice procedure that allows measuring the strength of an intersensory illusion in a bias-free way. With this new procedure we found that only flashes, but not averted eye-gazes, induced an illusory shift in sound location. This difference between flashes and eye-gazes was validated in an EEG study in which again only flashes illusorily shifted the apparent location of a sound thereby evoking a mismatch negativity response. These results are important because they highlight that commonly used measures of multisensory illusions are contaminated while there is an easy yet stringent way to measure the strength of an illusion in a bias-free way.

Introduction

Our brains are often multisensory in nature, trying to combine different information sources. One of the most striking manifestations of this is the so-called ventriloquist illusion, referring to the observation that the apparent location of a sound is shifted towards a spatially disparate but otherwise synchronised visual stimulus such as a flash (Howard & Templeton, 1966; Bertelson, 1999). The ventriloquist illusion has been taken as a particularly clear example of conflict reduction between the senses for events that should normally converge, and this has been shown to be consistent with optimal integration of visual and auditory signals (Alais & Burr, 2004). Recent evidence further suggests that not only flashes, but also imagined flashes (Berger & Ehrsson, 2013), arrow cues (Borjon *et al.*, 2011), and averted eye-gazes of a pictured face (Borjon *et al.*, 2011; Collins & Schirillo, 2013) may shift the apparent location of a sound. Possibly, this occurs because all these stimuli induce a shift of attention that re-aligns the interpretation of where a sound is heard, although others have reported that an overt shift of the eyes of an observer actually shifts the apparent location of a sound in the opposite direction (Lewald, 1997).

The perceptual nature of the ventriloquist effect of these different stimuli (whether they cause perceptual fusion or response bias) remains to be examined because many studies suffer from methodological pitfalls (see e.g. Chen & Vroomen, 2013). Here, we first

used traditional methods to examine whether flashes and averted eye-gazes of a cartoon face induced a significant ‘direct ventriloquist’ effect and a ‘ventriloquist aftereffect’. We argue that these traditional methods may be contaminated because with concurrent audiovisual stimulation (direct ventriloquist effect) participants may simply point to the visual stimulus whenever unsure about sound location, while unisensory sound localisation after exposure to audiovisual conflict (the ventriloquist aftereffect) may be contaminated by visual response priming (Bertelson, 1999). We therefore developed a two-alternative forced choice (2AFC) procedure that is devoid of response biases. In this task, participants heard on each trial a static and a left–right alternating sound sequence while visual left–right stimulation (flashes or eye-gazes) was added to both sound sequences. The participants’ task was to decide whether the first or the second sound sequence contained alternating sounds. Weber–Fechner law predicts that performance in the 2AFC task should become more difficult with visual stimulation, to the extent that a visual stimulus captures the apparent location of the sounds because capture decreases the relative difference between static and alternating sequences. A decrease in discrimination performance by visual stimulation thus can serve as a bias-free measure of ventriloquism. To validate this, we further relied on the fact that a mismatch negativity (MMN) response is evoked not only by a real auditory change in sound location (Näätänen, 1992) but also by an illusory change as reported before for flash-induced ventriloquism (Stekelenburg *et al.*, 2004). In Experiment 4, we therefore examined whether eye-gazes would evoke a ventriloquist-MMN in a manner similar to the MMN evoked by flash-induced illusory shifts.

Correspondence: Jean Vroomen, as above.
E-mail: j.vroomen@uvt.nl

Received 10 December 2013, revised 23 January 2014, accepted 24 January 2014

Materials and methods

The study was approved by the ethics committee of Tilburg University, and was conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki. All experiments were undertaken with the understanding and written consent of each subject. In all experiments, auditory stimuli were presented by two broad-band loudspeakers (Harman/Kardon HK206, Woodbury, NY, USA) placed 27° to the left and right of a 19-inch CRT screen (60 Hz refresh). A stereo white-noise burst (30 ms in duration, 5 ms linear fade-in and fade-out, was created and played at ~ 65 dB-A. The apparent horizontal location of the sound was varied by adjusting the sound level between the left and right speaker (the balance), thus providing inter-aural level cues. The balance was adjusted in e-Prime using the pan value function, with pan values varying from 0 (central sound) to +/- 450 in various step sizes (negative pan values for left sounds; pan values in 100th of Decibels). Visual stimuli were presented against a dark background at a distance of ~ 57 cm. The flash consisted of a white circle (1.8° diameter) presented at 10.6° to the left or right of the centre for 33 ms. For the eye-gaze condition, a yellow cartoon face (5.5° diameter) was shown in the centre of the screen with pupils either straight ahead or averted to the left or

right. The flash was synchronised with the sound, while the eyes shifted gaze 300 ms before sound onset (following Borjon *et al.*, 2011). Eyes returned to straight ahead after sound deliverance.

Experiment 1 – direct ventriloquist effect

Sixteen first-year students from Tilburg University received course credits for participation. Participants were told to fixate the centre of screen and to report by pressing either of two buttons with the right hand whether the sound came from left or right of central fixation, trying to ignore the flash or eye-gaze (Fig. 1A). Pan values of sounds varied from -450 to +450 in steps of 150, resulting in 28 unique stimuli [7 (pan values) × 2 (flash/gaze) × 2 (left/right)], each presented 15 times in blocks of 80 trials. Flashes and gazes were blocked, while all other conditions were presented randomly. Ten catch trials were included per block (a blink of the fixation) to ensure fixation on the centre.

Experiment 2 – ventriloquist aftereffect

Fifteen new participants were tested in Experiment 2. Multiple short exposure phases to audiovisual conflict were followed by an auditory-only test (see Fig. 1C). During exposure, the central sound was

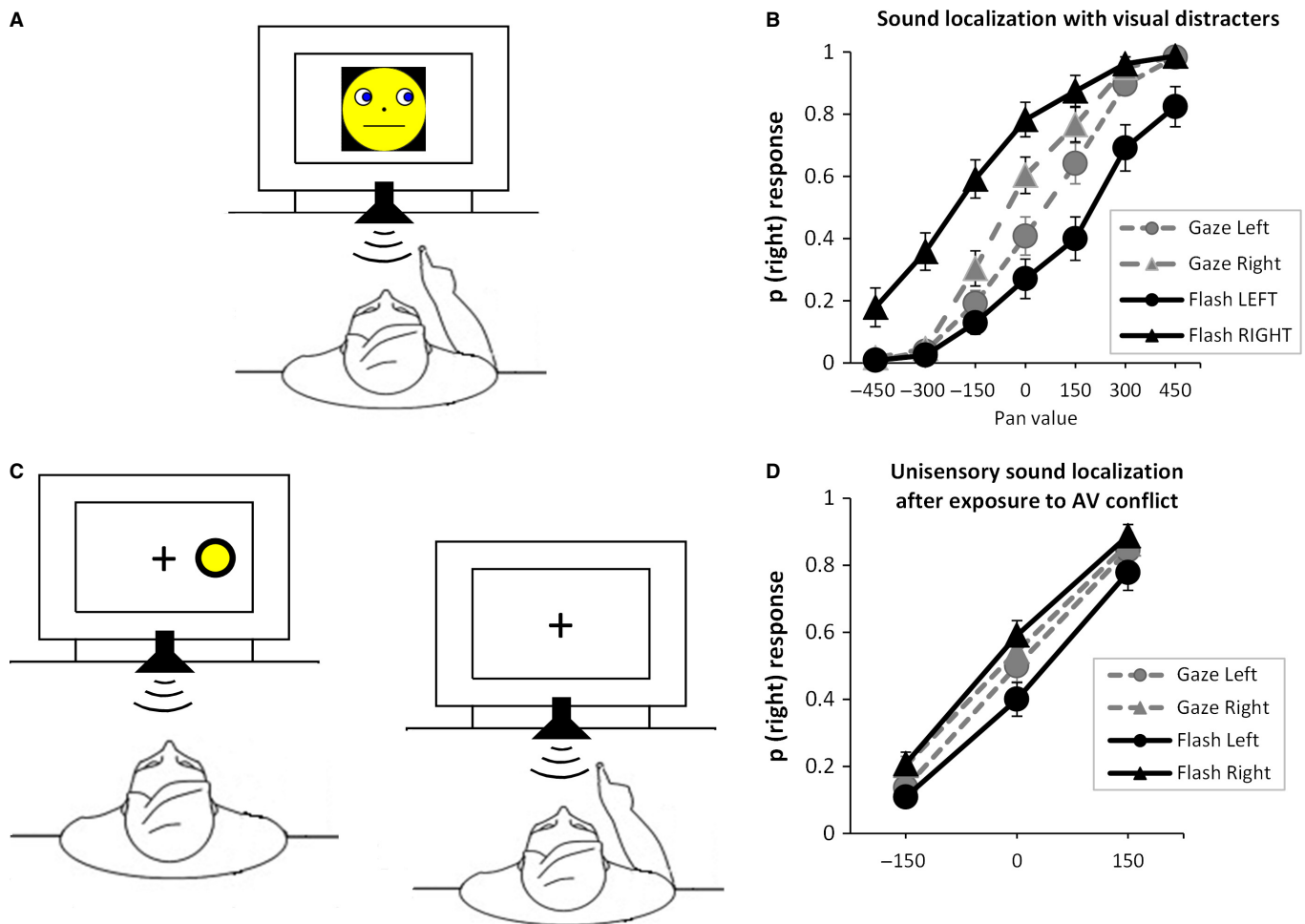


FIG. 1. (A) Observers judged whether a sound came from left or right of central fixation while ignoring a flash or eye-gaze (only eye-gaze depicted). (B) Flashes and eye-gazes both shifted sound location responses in their direction. (C) Following exposure to an audiovisual conflict (a train of eight central sounds with flashes or eye-gazes on the left or right; only a flash on the right is depicted) observers judged whether test sounds came from the left or right of central fixation. (D) Unisensory sound localisation was shifted in the direction of the previously experienced conflict.

presented eight times with a flash (or eye-gaze) on the left (or right). After a 1-s pause, six auditory-only test trials with pan values of -150 , 0 or $+150$ were presented twice in random order. Participants judged with two buttons whether the test sound came from left or right. There were 24 exposure-test blocks with flashes, and 24 with eye-gazes (288 test sounds in total). The direction of the visual stimulation (left or right) and stimulus type (gaze or flash) was blocked, with order counterbalanced over participants. To ensure fixation, an additional six catch trials (a blink of the fixation) were included per visual stimulus. Participants had to press a key as soon as they detected a blink of the fixation.

Experiment 3 – 2AFC task

Fifteen new students were recruited. Pan values varied between -400 and $+400$ in steps of 50 . Successive sounds also alternated in pitch (by low-pass filtering them at either 9 or 10 kHz) to ensure that non-spatial difference between sounds could not serve as a cue for left–right alternation. Each trial consisted of two sequences of four pitch-alternating sounds (interstimulus interval, 900 ms; 1900 ms between sequences). Each sequence was indexed by a ‘1’ or ‘2’. One of the two sequences was static (pan value 0), the other sequence contained sounds whose pan values alternated at eight different separations with pan values from $-50/+50$ to $-400/+400$ in steps of ± 50 . The static and alternating sounds were all accompanied by flashes or eye-gazes, the alternating sounds in a spatially congruent fashion (i.e., left sounds were combined with flashes or eye-gazes on the left, and right sounds were combined with flashes or eye-gazes on the right). The position of the static sequence (first or second), the pitch value of the first sound (lower- or higher-pitched),

and the pan value of the first alternating sound (left or right) varied randomly from trial to trial, yielding 64 unique trials in a block. Each block also contained eight catch trials (a blink of fixation) to ensure that participants watched the screen. The whole test consisted of six blocks: two blocks with flashes, two blocks with eye-gazes and two blocks without visual distracters (but only a fixation) serving as baseline. These six blocks were presented in an ABCCBA order, and were counterbalanced across participants. Participants fixated the screen during sound presentation and judged which of the two sequences contained the alternating sounds with two designated keys with the right hand. Upon detection of a catch trial they had to press a key with the left hand.

Experiment 4 – EEG study

Nineteen new students (10 female, nine male) participated. For purposes of subtraction of the EEG signal, there were five conditions in total: an auditory-only (A), two audiovisual (AV-flash; AV-gaze), and two visual-only (V-flash, V-gaze) conditions. Each condition comprised 765 (85%) standards and 135 deviants (15%) administered across three blocks for a total of 4500 stimuli (interstimulus interval, 1150 ms). Trial order was randomised with the restriction that at least two standards preceded each deviant. In the A-condition, the standard sound was presented slightly to the left (pan value -50) and the deviant sound slightly to the right (pan value $+50$). Auditory stimulation in the AV conditions was identical to A, but standard sounds were now combined with a flash or eye-gaze on the left, and deviant sounds with a flash or gaze on the right. The V conditions were identical to the AV counterpart, but without sound. Blocks alternated in counterbalanced order across

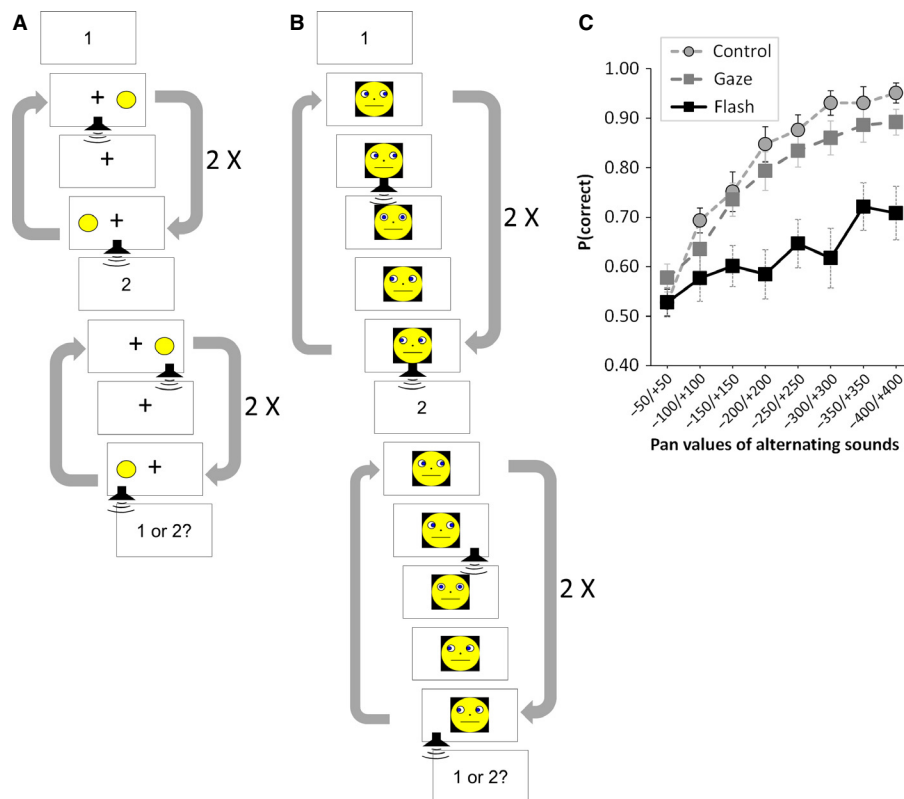


FIG. 2. (A) Participants heard two sequences of four sounds and decided whether the first or the second sequence contained left–right alternating sounds. Flashes were presented as distracter. (B) Same for gazes. (C) The proportion of correct responses as a function of the separation of the alternating sounds for each condition.

participants. Participant had to detect by key-press an occasional colour change at central fixation (in 3.6% of the standard trials).

EEG recording and analysis

The EEG was sampled at 512 Hz from 64 locations using active Ag–AgCl electrodes (BioSemi, Amsterdam, The Netherlands) mounted in an elastic cap, and two mastoid electrodes. Electrodes were placed according to the extended International 10–20 system. Two additional electrodes served as reference (Common Mode Sense active electrode) and ground (Driven Right Leg passive electrode). EEG was referenced offline to an average of left and right mastoids, and bandpass-filtered (0.1–30 Hz, 24 dB/octave). The 50-Hz interference was removed by a 5-Hz notch filter. Raw data were segmented into epochs of 800 ms, including a 100-ms prestimulus baseline. Event-related potentials (ERPs) were time-locked to the sound onset in the A and AV conditions, and to the corresponding time stamp in the V conditions. After electro-oculogram correction (Gratton *et al.*, 1983), epochs with an amplitude change exceeding $\pm 120 \mu\text{V}$ at any EEG channel were rejected.

ERPs of standards and deviants were averaged per condition and then subtracted, resulting in a deviant – standard difference wave. The difference waves of the AV conditions may be composed of overlapping components pertaining to the illusory change in the sound, as well as the change in visual and auditory information. To subtract activity by visual and auditory change, A and V were subtracted from AV. These (AVflash – Vflash – A) and (AVgaze – Vgaze – A) difference waves thus represent brain activity evoked by the illusory change in sound location without contribution of the auditory and visual components (Stekelenburg *et al.*, 2004). A

second analysis involved the spatiotemporal dynamics of the difference waves by conducting point-by-point running two-tailed *t*-tests in a 1–700 ms window. Using a procedure to minimise type I errors (Guthrie & Buchwald, 1991), difference waves were considered significantly different from the pre-stimulus baseline level when at least 12 consecutive points (24 ms) were significantly more negative than zero.

Results

Performance on catch trials was high in each experiment (> 94% correct detection) and was not analysed any further.

Experiment 1 – direct ventriloquist effect

Figure 1B plots the group-averaged proportion of ‘right’ responses when participants had to locate a sound while ignoring a visual distracter. Flashes and eye-gazes both shifted localisation responses, though flashes were more potent. A 2 (Distracter type: Flash/Gaze) \times 2 (Direction of distracter: Left/Right) \times 7 (Pan value of the sound) ANOVA of the individual proportions of left responses confirmed that the interaction between Distracter type and Direction was significant, $F_{1,15} = 26.72$, $P < 0.001$, indicating that flashes induced a greater shift (34.1%) than eye-gazes (7.2%). Separate *t*-test on the direct ventriloquist effect of flashes and eye-gazes (i.e., the difference between visual stimulation on the right vs. left, pooled over sound locations; all tests Bonferroni-corrected), though, showed that both visual stimuli shifted sound localisation responses in their direction: Flashes, $t_{15} = 6.87$, $P < 0.001$; Eye-gazes, $t_{15} = 3.99$; $P < 0.002$.

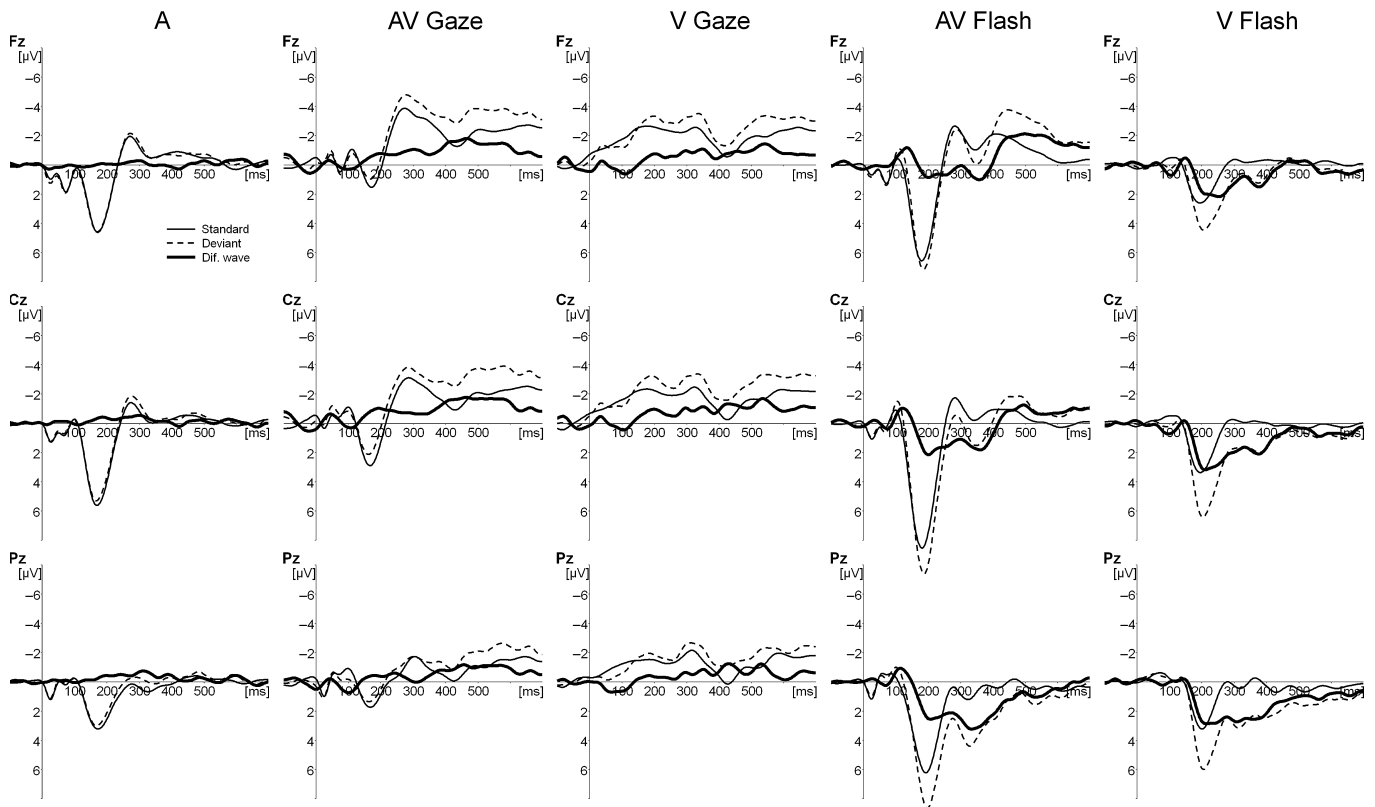


FIG. 3. Grand-averaged ERPs at electrodes Fz, Cz and Pz for the standard, deviant and the corresponding difference waves (deviant – standard) for the AV and V eye-gaze and flash, and A-only conditions.

Experiment 2 – ventriloquist aftereffect

Figure 1D shows the group-averaged proportion of ‘right’ responses to unimodal sounds after participants were exposed to audiovisual conflict. Both flashes and eye-gazes induced a ventriloquist aftereffect (i.e., more ‘right’ responses after exposure to central sounds with concurrent visual stimulation on the right than left), though the effect was greater for flashes than eye-gazes. A 2 (Visual stimulus: Flash/Gaze) \times 2 (Direction: Left/Right) \times 3 (Pan value) ANOVA on the individual proportion of ‘right’ responses confirmed that there was a main effect of Pan value ($F_{2,13} = 149.13$, $P < 0.001$, and Direction, $F_{1,14} = 15.57$, $P < 0.001$). The interaction between Visual stimulus and Direction was significant ($F_{1,14} = 8.80$, $P < 0.01$), indicating that flashes induced greater aftereffects than eye-gazes (13.3% vs. 4.2%, respectively). The second-order interaction was also significant ($F_{2,13} = 5.82$, $P < 0.05$) because the aftereffect was greatest for the most ambiguous central sound that participants were originally exposed to. Separate t -tests on the aftereffects of flashes and eye-gazes (i.e., the difference in sound localisation after exposure to visual stimulation on the right vs. left, pooled over the three sound locations; all tests Bonferroni-corrected) showed that flashes induced an aftereffect > 0 ($t_{1,14} = 3.99$, $P < 0.002$), while the aftereffect of eye-gazes was marginally significant ($t_{14} = 2.31$, $P = 0.074$).

Experiment 3 – 2AFC task

Figure 2C displays the group-averaged percentage of correct responses in the 2AFC task. When no distracter was present (the

auditory-only baseline), performance increased from near chance-level (50%) for sounds that alternated close to the centre (pan values ± 50) to near ceiling for the most widely separated sounds (pan values ± 400). Concurrent flashes drastically hampered performance, and even the most widely separated alternating sounds were still difficult to discriminate from static sounds (71% correct). Importantly, the effect of eye-gaze was now only minimal and did not differ statistically from baseline. A 3 (Visual stimulus: no distracter, flash, eye-gaze) \times 8 (Pan values of alternating sound) repeated-measures ANOVA on the individual proportion of correct responses showed that there was a main effect of Pan value ($F_{7,8} = 26.70$, $P < 0.001$, Visual stimulus, $F_{2,13} = 17.60$, $P < 0.001$), with no interaction ($F_{14,1} = 3.08$, $P = 0.42$). Average performance pooled across the pan values was 81.3% at baseline, 77.7% in the eye-gaze condition and 62.3% in the flash condition. Separate t -test showed that, compared to baseline, the 19.0% drop in performance with flashes was significant ($P < 0.001$) whereas the 3.7% drop with eye-gazes was not ($P = 0.19$). The 15.4% difference between flashes and eye-gaze was also significant ($P < 0.001$).

Experiment 4 – EEG study

Figure 3 shows the grand-averaged ERPs for the standard, the deviant and the corresponding difference waves (deviant – standard) for the AV, V and A conditions. Figure 4 plots the AV – V – A difference waves for eye-gazes and flashes. Figures 3 and 5A (representing point-wise t -test) show that in the A condition there was a small MMN at ~ 300 ms with a central-parietal distribution. For the flash

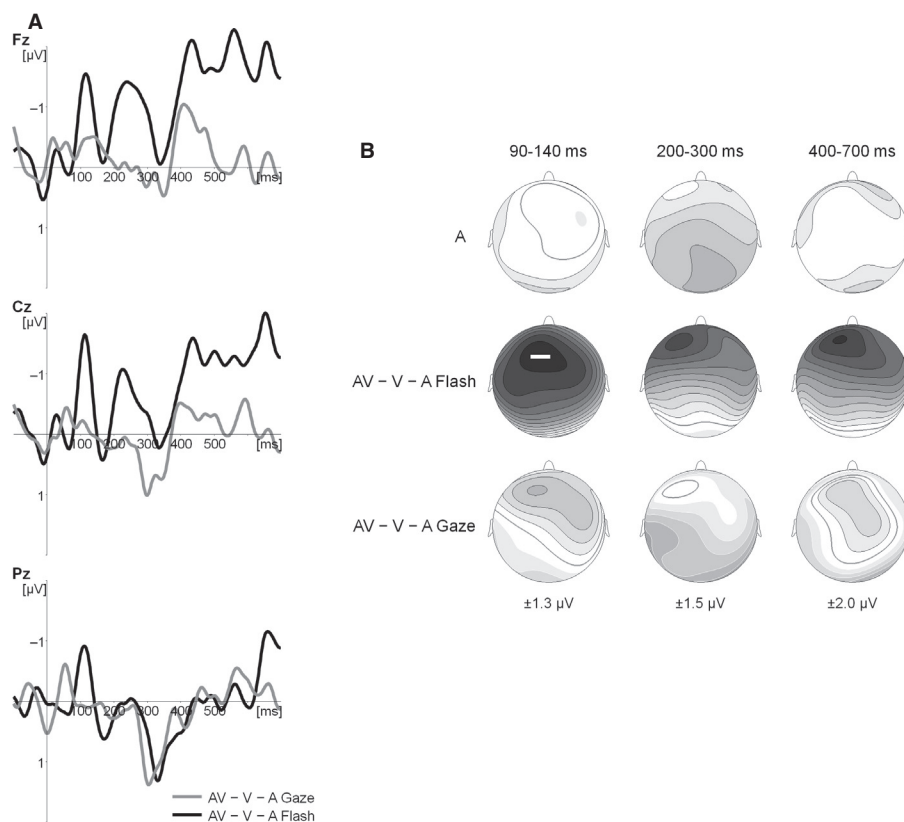


FIG. 4. (A) Grand-averaged ERPs at electrodes Fz, Cz and Pz for the (AV – V – A) difference waves for the eye-gaze and flash conditions. (B) Scalp topographies of mean activity in 90–140, 200–300 and 400–700 ms windows for the A and (AV – V – A) difference waves. The range of the voltage maps in μV is displayed below.

condition, the MMN consisted of an early (at ~ 100 ms) and late (200–300 ms) negative deflection with frontal topographies (Figs 4 and 5C). A similar distinction between an early and late phase of the auditory location MMN was reported by Deouell *et al.* (2006). This flash-induced ventriloquist-MMN was followed by a late negative frontal wave in a window of 400–700 ms which probably reflects the reorienting negativity (RON), indicating a reorienting of attention back towards centre (Schroger & Wolff, 1998). Importantly, for the eye-gaze condition, the AV – V – A difference wave hardly deviated from baseline (Figs 4 and 5B). To test for differences between eye-gazes and flashes, we calculated mean activity in three temporal windows based on running *t*-tests (see Fig. 5D), i.e., 90–140, 200–300 and 400–700 ms. A repeated-measures ANOVA for

the period 90–140 ms with Electrode (AF3, AFz, AF4, F3, Fz, F4, FC3, FCz, FC4) and Condition (eye-gaze vs. flash) as factors indicated that there was no main effect or interaction (*P*-values > 0.19) at that time. For the 200–300 ms and 400–700 ms windows, though, there were main effects of Condition ($F_{1,18} = 5.86, P < 0.05, F_{1,18} = 11.43, P < 0.05$) respectively, and no main effects of Electrodes and no interactions (*P*-values > 0.38), indicating that flashes were more effective at eliciting an MMN-like response than were eye-gazes.

We also tested whether brain activity differed from prestimulus baseline. Separate ANOVAS on the AV – V – A difference wave per temporal window for flashes and eye-gazes with Electrode were conducted. For flashes, mean activity in the 90–140, 200–300 and

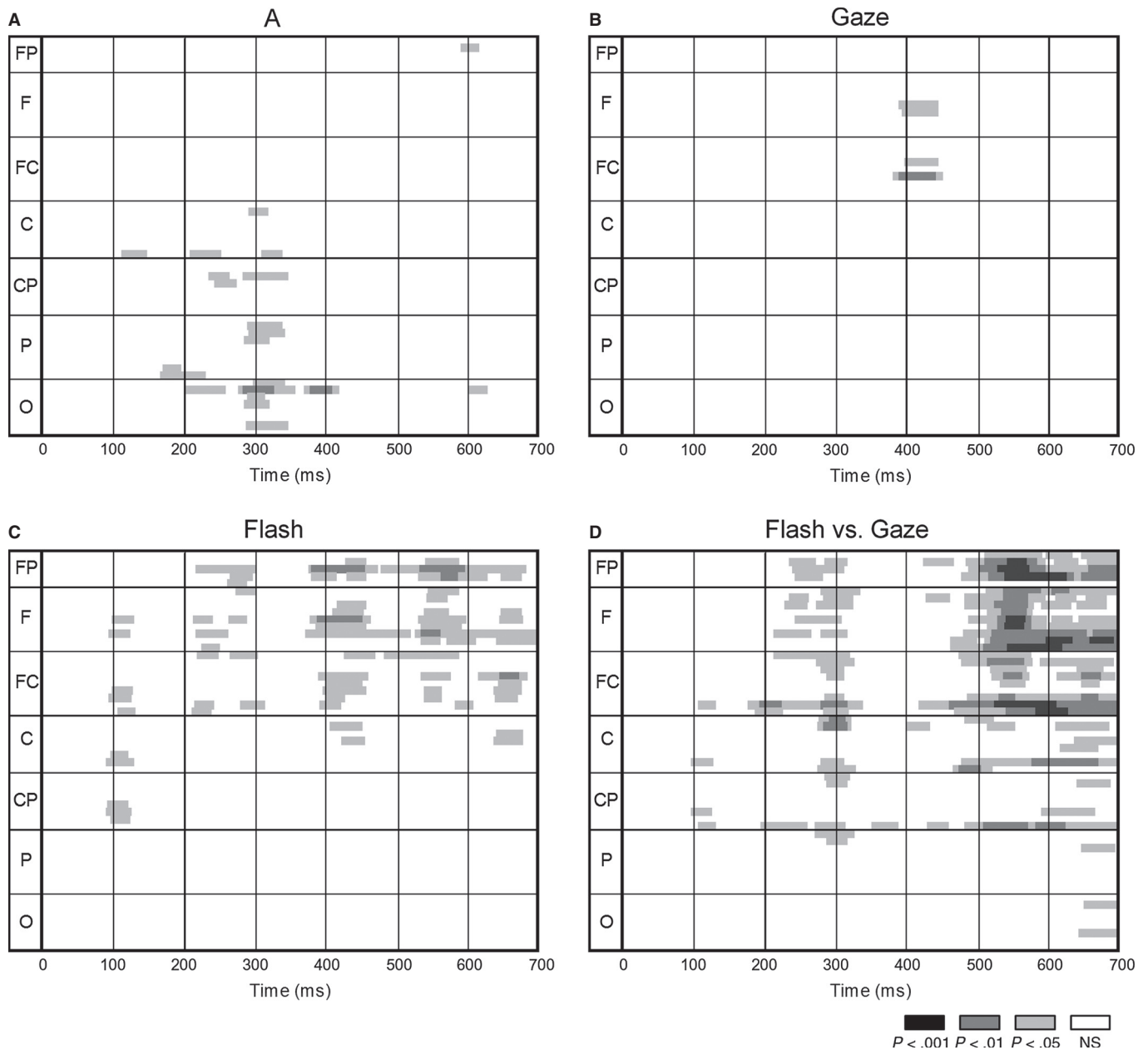


FIG. 5. Point-wise running *t*-tests in a 1–700 ms window at frontopolar (FP), frontal (F), frontocentral (FC), central (C), centroparietal (CP), parietal (P) and occipital (O) regions. The plot is highlighted only if at least 12 consecutive points were significant. (A) Running *t*-tests of the A difference wave tested against prestimulus baseline. (B and C) The AV – V – A difference waves of eye-gazes and flashes tested against prestimulus baseline. (D) The AV – V – A difference wave of flashes was more negative than eye-gazes.

400–700 ms windows were all significantly more negative than prestimulus baseline ($F_{1,18} = 4.49$, $P < 0.05$, $F_{1,18} = 4.93$, $P < 0.05$, $F_{1,18} = 6.87$, $P < 0.05$), but for eye-gazes these effects were all nonsignificant (all F -values < 1 ; see Fig. 5B and C). These analyses thus confirm that flashes induced a ventriloquist-MMN and RON while eye-gazes were ineffective.

Discussion

In the literature on multisensory perception, the most widely used method for examining multisensory integration is to ask participants to selectively attend to a target modality while ignoring concurrent stimulation in a distracter modality. With this procedure, we found that flashes and eye-gazes both shifted sound localisation responses. Unfortunately, of all procedures available, this is also the one that is most probably contaminated by response biases because participants may, whenever unsure about the target modality, rely on the distracter modality thus blurring the distinction between perceptual fusion and response strategy. An alternative procedure relies on aftereffects where targets are presented unisensorily after exposure to intersensory conflict. With this procedure, we found that flashes induced strong aftereffects indicative of ventriloquism whereas eye-gazes induced only a small and nonsignificant effect. This suggests that aftereffects provide a clearer picture on true ventriloquism. However, in practice it appears that the interpretation of aftereffects is also often complicated because it may entail different mechanisms. That is, besides issues with the build-up (the amount of exposure), decay (time between exposure and test) and the optimal stimulus for inducing aftereffects (see e.g. Pages & Groh, 2013), it cannot be excluded that the exposure stimuli themselves promote 'priming' so that after repeated exposure to, say, a ventriloquised sound on the right, participants are more willing to respond 'right' at test.

To avoid these difficulties, we developed a 2AFC procedure as a more stringent test of perceptual fusion, a test that relies on one of the most established laws in psychology – Weber–Fechner. The crux of this procedure is that participants heard on each trial a static and a left–right alternating sound sequence while visual left–right stimulation (flashes or eye-gazes) was added to both sequences. The participants' task was to decide which of the two sequences contained the alternating sound sequence, first or second? Performance was expected to become more difficult to the extent that the concurrent visual stimulation captures the apparent location of the sounds because capture decreases the relative difference between static and alternating sound sequences. With this procedure, we found that flashes, but not eye-gazes, were able to hamper performance, presumably because only flashes truly shifted the apparent location of sounds. This interpretation was validated in an EEG study in which we again found that only flashes evoked a ventriloquist-MMN, which is considered to be the brain's signature of a perceptual shift in sound location.

Others, including the first author, have argued that a psychophysical staircase may also eliminate any possible influence of voluntary post-perceptual decisions on the ventriloquism effect (Bertelson & Aschersleben, 1998; Vroomen *et al.*, 2001). In this procedure, the presentation of a visual stimulus at fixation has been found to lead to earlier uncertainty, in the form of response reversals, about the location of sounds presented at locations further from fixation compared to when no visual stimulus is presented at fixation. Our 2AFC procedure bears resemblance to this, but a crucial difference is that it does not exclusively rely on the early reversals. Our new procedure is therefore much more reliable because each trial contributes

to the response distribution rather than just the first few reversals. Note also that especially early reversals might be caused by distraction by visual stimulation (because participants may not yet be familiar with task requirements) or, in the case where a visual stimulus is imagined (Berger & Ehrsson, 2013), there may be interference by the dual task of imagining a flash while a sound is presented.

Note also that our 2AFC procedure can be applied to other intersensory illusions to distinguish response bias from perceptual fusion. For example, it is well known that lip-read information can change the way a sound is heard (e.g., an auditory/b/combined with lip-read/g/is often 'heard' as/d/, (McGurk & MacDonald, 1976), while letters may have similar effects on speech identification (van Atteveldt *et al.*, 2004). To examine whether letters and lip-read information either bias or fuse with auditory speech, one can envisage a task in which listeners discriminate between a train of 'constant' or 'alternating' speech sounds while trying to ignore letters or lip-read distracters. The amount of interference by these distracters could well serve as a bias-free measure of perceptual fusion.

It should be stressed that we fully acknowledge that the deeper reason why flashes differ from eye-gazes remains to be examined. There are numerous low-level stimulus differences such as saliency, energy, and audiovisual timing that may explain why flashes were more potent in capturing the apparent location of a sound than eye-gazes (for review, see Chen & Vroomen, 2013). Also, for synchronised flash–sound pairings one might argue that the brain assumes that a single event has occurred (de Gelder & Bertelson, 2003), whereas for eye-gazes this single-event explanation seems unlikely because an eye-gaze by itself is different from an event that produced a sound on which the eyes were subsequently directed. Ideally, one might want to vary these stimulus dimensions in a systematic way to examine their impact on ventriloquism. In the end, it may even turn out to be the case that if flashes and eye-gazes are equated on all relevant dimensions, there is no fundamental difference in their impact on sounds. For the purpose of the current study, though, this is all irrelevant because we only wanted to demonstrate that our 2AFC procedure is a better and more stringent method of examining intersensory illusions than is commonly used in this field.

Acknowledgements

The authors declare no competing financial interests. We would like to thank Rolf Ulrich for insightful comments.

Abbreviations

2AFC, two-alternative forced choice; ERP, event-related potential; MMN, mismatch negativity.

References

- Alais, D. & Burr, D. (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.*, **14**, 257–262.
- van Atteveldt, N., Formisano, E., Goebel, R. & Blomert, L. (2004) Integration of letters and speech sounds in the human brain. *Neuron*, **43**, 271–282.
- Berger, C.C. & Ehrsson, H.H. (2013) Mental Imagery Changes Multisensory Perception. *Curr. Biol.*, **23**, 1367–1372.
- Bertelson, P. (1999) Ventriloquism: a case of crossmodal perceptual grouping. In Aschersleben, G., Bachmann, T. & Musseler, J. (Eds), *Cognitive Contributions to the Perception of Spatial and Temporal Events*. Elsevier, North-Holland, pp. 347–363.
- Bertelson, P. & Aschersleben, G. (1998) Automatic visual bias of perceived auditory location. *Psychon. B. Rev.*, **5**, 482–489.
- Borjon, J.L., Shepherd, S.V., Todorov, A. & Ghazanfar, A.A. (2011) Eye-gaze and arrow cues influence elementary sound perception. *P. Roy. Soc. B-Biol. Sci.*, **278**, 1997–2004.

- Chen, L. & Vroomen, J. (2013) Intersensory binding across space and time: a tutorial review. *Atten. Percept. Psycho.*, **75**, 790–811.
- Collins, L. & Schirillo, J. (2013) Attention to endogenous and exogenous cues affects auditory localization. *Exp. Brain Res.*, **231**, 13–18.
- Deouell, L.Y., Parnes, A., Pickard, N. & Knight, R.T. (2006) Spatial location is accurately tracked by human auditory sensory memory: evidence from the mismatch negativity. *Eur. J. Neurosci.*, **24**, 1488–1494.
- de Gelder, B. & Bertelson, P. (2003) Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.*, **7**, 460–467.
- Gratton, G., Coles, M.G. & Donchin, E. (1983) A new method for off-line removal of ocular artifact. *Electroen. Clin. Neuro.*, **55**, 468–484.
- Guthrie, D. & Buchwald, J.S. (1991) Significance testing of difference potentials. *Psychophysiology*, **28**, 240–244.
- Howard, I.P. & Templeton, W.B. (1966) *Human Spatial Orientation*. John Wiley and Sons, Oxford, England.
- Lewald, J. (1997) Eye-position effects in directional hearing. *Behav. Brain Res.*, **87**, 35–48.
- McGurk, H. & MacDonald, J. (1976) Hearing lips and seeing voices. *Nature*, **264**, 746–748.
- Näätänen, R. (1992) *Attention and Brain Function*. Psychology Press, Hillsdale.
- Pages, D.S. & Groh, J.M. (2013) Looking at the ventriloquist: visual outcome of eye movements calibrates sound localization. *PLoS One*, **8**, e72562.
- Schroger, E. & Wolff, C. (1998) Attentional orienting and reorienting is indicated by human event-related brain potentials. *NeuroReport*, **9**, 3355–3358.
- Stekelenburg, J.J., Vroomen, J. & de Gelder, B. (2004) Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neurosci. Lett.*, **357**, 163–166.
- Vroomen, J., Bertelson, P. & de Gelder, B. (2001) The ventriloquist effect does not depend on the direction of automatic visual attention. *Percept. Psychophys.*, **63**, 651–659.