

Seeing Sounds and Hearing Colors: An Event-related Potential Study of Auditory–Visual Synesthesia

Aviva I. Goller¹, Leun J. Otten², and Jamie Ward¹

Abstract

■ In auditory–visual synesthesia, sounds automatically elicit conscious and reliable visual experiences. It is presently unknown whether this reflects early or late processes in the brain. It is also unknown whether adult audiovisual synesthesia resembles auditory-induced visual illusions that can sometimes occur in the general population or whether it resembles the electrophysiological deflection over occipital sites that has been noted in infancy and has been likened to synesthesia. Electrical brain activity was recorded from adult synesthetes and control participants who were played brief tones and required to monitor for an infrequent auditory target. The synesthetes were instructed to attend either to the auditory or to the visual (i.e., synesthetic) dimension of the tone, whereas

the controls attended to the auditory dimension alone. There were clear differences between synesthetes and controls that emerged early (100 msec after tone onset). These differences tended to lie in deflections of the auditory-evoked potential (e.g., the auditory N1, P2, and N2) rather than the presence of an additional posterior deflection. The differences occurred irrespective of what the synesthetes attended to (although attention had a late effect). The results suggest that differences between synesthetes and others occur early in time, and that synesthesia is qualitatively different from similar effects found in infants and certain auditory-induced visual illusions in adults. In addition, we report two novel cases of synesthesia in which colors elicit sounds, and vice versa. ■

INTRODUCTION

In auditory–visual synesthesia, sounds automatically elicit conscious visual percepts in addition to an auditory percept. For example, a cello may sound like a “dark velvet or reddish-brown tree trunk-like texture” and a flute may be “dry and transparent with pastel colors” (Mills, Boteler, & Larcombe, 2003). For some individuals, the synesthesia is triggered solely by speech (e.g., Nunn et al., 2002; Paulesu et al., 1995; Baron-Cohen, Harrison, Goldstein, & Wyke, 1993), but for others the synesthesia is triggered by all known auditory stimuli (e.g., Thornley Head, 2006; Ward, Huckstep, & Tsakanikos, 2006). This may reflect a qualitative difference between whether synesthesia is linked to linguistic representations (e.g., graphemes) versus perceptual properties of the stimulus such as its pitch (e.g., Simner, Glover, & Mowat, 2006; Frith & Paulesu, 1997). Our study considers the latter, using nonspeech tones. This variety of synesthesia is of particular theoretical interest because of the large literature on audiovisual interactions in the nonsynesthetic brain (e.g., Calvert, Hansen, Iversen, & Brammer, 2001). It raises the possibility that this type of synesthesia reflects an adaptation of normal multisensory processes (Ward et al., 2006). Previous research has demonstrated the authenticity of this type of synesthesia by

showing that the auditory-color associations of these synesthetes are more consistent than controls and by showing that the synesthetic color of a task-irrelevant tone interferes with color naming in a Stroop task (Ward et al., 2006). However, in other respects there are commonalities between the nature of synesthetic experiences and those reported by nonsynesthetes in imagery, matching tasks, or cross-modal interference paradigms (Ward et al., 2006; Marks, 2004). In particular, high-pitched sounds tend to be visually lighter, higher, and smaller than low-pitch sounds in both synesthetic experiences and response biases of nonsynesthetes (Marks, 2004). This suggests common processes between synesthetic perception and audiovisual processing of nonsynesthetes. The present study will use ERPs (small changes in the brain’s electrical activity time locked to an event) to adjudicate between two different theories.

One suggestion is that the newborn infants’ experiences of the world resemble a form of synesthesia (Maurer & Mondloch, 2006; Maurer & Maurer, 1988), in which the senses are yet to be differentiated and in which one sense (e.g., audition) can trigger another (e.g., vision). For example, infants show cross-modal habituation depending on the intensity of light and sounds (Lewkowicz & Turkewitz, 1980). A further claim is that, in some individuals, these early multisensory pathways are retained into adulthood giving rise to developmental synesthesia whereas in most other individuals

¹University of Sussex, ²University College London

they are greatly diminished (Maurer, 1997; Baron-Cohen, 1996). Electrophysiological correlates of infantile auditory–visual “synesthesia” have been reported. These consist of a large negative deflection between 100 and 500 msec over occipital sites (absent by 30 months of age), contrasting with a developmentally more stable potential over temporal sites (Neville, 1995). Although a direct comparison between adult audiovisual synesthetes and normal infants would be impossible to interpret (e.g., due to developmental changes in conductance), one can nevertheless determine whether a qualitatively similar electrophysiological signature is found in adult synesthetes to that previously documented in infants. According to this account, the electrophysiological responses to auditory stimuli should include an early deflection maximal over posterior sites.

Recent studies have shown that there are direct projections from primary auditory cortex (A1) to primary visual cortex (V1) in the mature primate brain although they are primarily found in regions representing peripheral vision (Rockland & Ojima, 2003; Falchier, Clavagnier, Barone, & Kennedy, 2002). Even in nonsynesthetes, direct auditory–visual projections may play a functional role in multisensory processing. It may even give rise to a synesthesia-like illusion in the normal population. Shams, Kamitani, Thompson, and Shimojo (2001) and Shams, Kamitani, and Shimojo (2000) report that if two beeps are played in quick succession and are accompanied by a single flash, then participants often perceive two distinct flashes instead of one: the double-flash illusion. The illusion occurs predominantly in peripheral vision, consistent with the known anatomy of the direct projections. They report that the illusory flash is accompanied by electrical activity over occipital sites (Oz, O1, and O2) less than 110 msec after the onset of the second beep (Shams et al., 2001), and a recent fMRI study shows differences in V1 activity but not in other brain regions (Watkins, Shams, Tanaka, Haynes, & Rees, 2006).

Anatomical studies that have attempted to look for the reverse pathways, from V1 to A1, have not found them (Innocenti, Berbel, & Clarke, 1988). The apparent rarity with which vision evokes sound in synesthesia relative to vice versa (e.g., Simner, Mulvenna, et al., 2006; Day, 2005) could therefore relate to the relative availability of these multisensory pathways. However, visual–auditory synesthesia has been anecdotally noted before. Baron-Cohen, Burt, Smith-Laittan, Harrison, and Bolton (1996) briefly note the case of JR who sees colors when she hears sounds and also hears sounds when she sees colors (a situation that does cause perceptual interference in her day-to-day life). They note that she shows high consistency over time when assessed in both directions although the associations are not necessarily the same as each other in either direction. For instance, a red traffic light may trigger a particular synesthetic sound, but that sound may itself trigger a synesthetic color of, say, blue (rather than red). In this

study, we retest JR using visual- and auditory-evoked ERPs together with one other synesthete, SL, who reports a similar pattern to JR.

There is, however, an alternative to the hypothesis of direct auditory–visual connections. This has been termed the “cross-modal transfer hypothesis” (Ward et al., 2006; Baron-Cohen, 1996). This hypothesis assumes that connections between auditory and visual regions are indirect and are mediated by multisensory audiovisual brain regions. Activation in multisensory neurons may feedback and influence activity in regions traditionally considered to be unisensory (e.g., Driver & Spence, 2000). Although multisensory processes may normally be activated when two senses are stimulated, it is also conceivable that they can sometimes be activated by a unimodal stimulus in some situations (e.g., Giraud, Price, Graham, Truy, & Frackowiak, 2001; Calvert et al., 1997). Synesthesia may be one such example. According to this account, an additional ERP deflection (due to multisensory binding) would follow the normal early auditory deflections, but would precede in time any auditory-evoked visual potential.

To date, there have been very few ERP studies of synesthesia. Two studies have considered grapheme-color synesthesia using visually presented graphemes (Schiltz et al., 1999) or spoken letter names and words (Beeli, Esslen, & Jancke, 2008). In addition, there are two single case studies that have specifically investigated electrophysiological correlates of auditory–visual synesthesia using nonlinguistic stimuli (Rao, Nobre, Alexander, & Cowey, 2007; Rizzo & Eslinger, 1989).

Schiltz et al. (1999) tested 17 grapheme-color synesthetes who were presented with runs of visual letters and who were required to detect certain target letters (e.g., vowels). They reported an increased positivity at frontal and central scalp sites emerging around 150 msec and maintained until 600 msec, relative to a nonsynesthetic control group. More recently, an ERP study was reported of grapheme-color synesthesia in which spoken letters and words elicit experiences of color (Beeli et al., 2008). As noted above, in this type of “color hearing,” it has been suggested that speech activates graphemic representation of words, which is then linked to experiences of color (e.g., Simner, Glover, et al., 2006; Frith & Paulesu, 1997). The fact that visual graphemes also elicit color for these synesthetes is consistent with this (Beeli et al., 2008). Beeli et al. (2008) found reduced amplitudes and/or increased latencies of the auditory N1, P2, and N2 deflections in synesthetes in response to spoken letters and words. Although few electrode sites were available, source localization implicated additional activity in the synesthetes in OFC and inferior temporal lobe.

Of the two single case studies of auditory–visual synesthesia that have measured ERPs to nonlinguistic sounds, one was an acquired case of synesthesia following blindness, and the other was a case of developmental synes-

thesia more closely related to that investigated here. The case of Rao et al. (2007) suffered destruction of the optic nerves following a car accident and began to report phosphenes from auditory stimuli a year or so later. A comparison of evoked potentials to sounds that did elicit a visual experience versus those that did not revealed modulation of an auditory-evoked N1 deflection (around 100 msec) including over occipital sites. The authors suggest that this reflects auditory activation of visual cortex rather than volume conduction from a distant site because sighted controls showed no comparable activity over the same electrode sites. Rizzo and Eslinger (1989) conducted the only electrophysiological study of developmental auditory–visual synesthesia to date. Auditory clicks were not associated with abnormal potentials at O1, O2, or Oz (other electrodes were not reported), and they conclude that the synesthesia may depend on “more anterior structures with polymodal connections.” The results of this case study were very different to that obtained by Rao et al. (2007), suggesting a possible involvement of different neural pathways in acquired and developmental cases of auditory–visual synesthesia. This will be returned to again in the discussion.

The present study is the first to consider the time course, using ERPs, of synesthesia in a group of people who experience colors in response to nonlinguistic sounds. This type of synesthesia is likely to be qualitatively different to that based on speech sounds and more closely resembles (in terms of phenomenology and candidate mechanisms) synesthesia-like illusions in the normal population (Shams et al., 2001), acquired audiovisual synesthesia (Rao et al., 2007), and possible infantile synesthesia (Neville, 1995). A series of pure tones were played to a group of synesthetes ($n = 10$) and controls ($n = 10$) who were required to detect an infrequent target tone. In different blocks, the synesthetes were additionally instructed to attend either to the pitch of the tone or to the synesthetic color of the tone. Attention modulates synesthesia as measured on tasks such as synesthetic Stroop interference (e.g., Mattingley, Payne, & Rich, 2006). However, inattention does not necessarily eliminate synesthesia (Sagiv, Heer, & Robertson, 2006), and it is unclear whether attention exerts its effects early or late in the induction of synesthesia. Synesthetes and controls were additionally shown unimodal color patches to record visual-evoked potentials. The color patches were similar in hue to some of their auditory-induced visual experiences although only 2 of 10 of the synesthetes reported conscious auditory experiences from seeing color. As such, we were able to explore differences between synesthetes who do and do not experience sound from color (comparing the 2 synesthetes to the remaining 8) and also to assess whether synesthetes have normal electrophysiological responses to visual stimuli (by comparing the 8 synesthetes to the 10 nonsynesthetes). It has been

suggested that colors may implicitly trigger representations of number in synesthetes (Cohen Kadosh, Cohen Kadosh, & Henik, 2007; Cohen Kadosh & Henik, 2006; Cohen Kadosh et al., 2005), and it is conceivable that colors may elicit sounds.

METHODS

Participants

Ten auditory–visual synesthetes (9 women) and 10 healthy controls (6 women) gave written informed consent to take part in this experiment. The procedure was approved by the University College London ethics committee. The mean age of the synesthetes was 39.7 years with a range of 21–68 years, and the mean age of controls was 39.9 years with a range of 20–67 years. All synesthetes reported being right-handed; two controls reported being left-handed. Handedness was considered unlikely to affect performance in auditory detection. Eight of the synesthetes report their sound–vision synesthesia to be unidirectional (i.e., sounds evoke vision but vision does not evoke sound), whereas two claim it to be bidirectional. For example, JR (also studied by Baron-Cohen et al., 1996) reports the following sounds as she moves her gaze around a Kandinsky painting (Composition VIII, 1923): “There is a huge splurge of sound left-hand top—booming but also a bit vulgar! Below it a rather mousy little meee sound which then translates into ohs and ahs and pops at the various circles. The lines are sharp and moving to the right with the sound of steel (like blades scraping against one another). The triangle and boomerang shapes are surprised and pop up laughing with a whooo.”

All participants were free of known neurological illness and reported normal hearing and normal or corrected-to-normal vision. Color vision was assessed with the Ishihara plates and was found to be normal in all participants. Participants were paid at a rate of £7.50/hr and had their travel expenses reimbursed.

The synesthetes were informally asked about the spatial location of their photisms by altering the location of a sound (e.g., a knock) and by altering their posture. If the locations of the photisms are gaze centered, then it would suggest dependence on retinotopically organized regions (e.g., V1), but if the locations depend on posture, it would be more consistent with an involvement of higher visual processes (e.g., Colby & Goldberg, 1999). None reported gaze-centered photisms. For two synesthetes, the photisms appeared to be located in front of them relative to the body trunk. Thus, the photisms would remain “in front” when the eyes and/or head were turned to the side and irrespective of where the sound comes from. For three synesthetes, the photism would be located in line with the location of the sound, irrespective of their own position. For three synesthetes, the photism was reported to be

internal, in their “mind’s eye,” and was unaffected by either their own position or the location of the sound. The remaining two synesthetes reported a combination of the above, namely, that the photism would initially appear to come from the location of the sound but could be shifted in to a different spatial reference frame (relative to their body) when attended.

A modified version of the “test of genuineness” used by Ward et al. (2006) was used on a set of 40 different tones of varying pitch and timbre (10 pure tones, 10 strings, 10 piano, and 10 notes of other timbre). Participants were required to choose the “best” color on two occasions using a standard Windows color palette (controls were encouraged to guess). Synesthetes were significantly more consistent (i.e., smaller differences) in their RGB selections, $t(18) = 2.55$, $p < .05$, relative to the controls reported by Ward et al. (2006) for the same stimuli.

Experimental Design

The experiment was based on an oddball paradigm in which auditory or visual stimuli were presented and participants responded to the presence of an infrequent target that was defined in terms of either pitch or color. This ensured that participants attended to the stimuli, although only the frequent stimuli were analyzed. The control participants were requested to respond to the pitch of the auditory stimuli and the color of the visual stimuli (i.e., two conditions). The auditory-to-visual synesthetes were asked to respond to either the pitch or the synesthetic color of the auditory stimulus or to the color of the visual stimulus (i.e., three conditions). The two synesthetes with bidirectional auditory–visual synesthesia were given a fourth condition, namely, to attend to the synesthetic sound of the colored stimulus. The order of presentation of the conditions was randomized. As such, our main analyses consisted of a between-subject manipulation (presence or absence of synesthesia) and a within-subject manipulation within the group of synesthetes (attend auditory vs. attend visual). This enabled us to determine when differences between synesthetes and controls first emerge (and over which sites) and to what extent these differences reflect whether the synesthete was instructed to attend to the sound or to the synesthetic vision.

Stimulus Materials

In the auditory stimulus conditions, the frequent auditory stimulus consisted of five tones within a limited pitch range. Slightly different tones were used to avoid habituation or adaptation. The pitches of the tones were 262, 277, 294, 311, and 330 Hz (each corresponds to a semitone difference on the Western musical scale). The infrequent target stimulus was much higher in pitch (1000 Hz) and was distorted with an auditory filter to

make it more distinctive. The stimuli were presented for 200 msec.

To select the colors for the visual stimulus conditions, the sounds were played to the synesthetes in advance of the EEG session. They used a standard Windows color picker to choose the color that best represented their experience to the tone. The sounds were not labeled and were each presented twice to ensure reliability. The order was randomized. For each sound and for each synesthete, the two RGB values were averaged, and this was used as the representative color for that sound for that person. The colors were chosen in this way to match them, approximately, to the visual reports of their synesthesia and also to test the hypothesis that synesthesia may be bidirectional (i.e., vision to auditory) even in synesthetes who do not report auditory phenomenology with vision. The color for the infrequent (higher pitch) sound differed from synesthete to synesthete but tended to be lighter in color and was always easily discriminable from the others. Participants were shown the infrequent and the frequent colors before the visual trials and informed of which one was the target. In the conditions in which visual stimuli were presented, each synesthete was presented with their own synesthetic colors, and an age-matched control was shown the same colors.

Procedure

In a dark cubicle, a fixation cross appeared on a black screen for the duration of the experiment to minimize eye movements. The screen was at a distance of 90 cm from the participant.

In the auditory conditions, participants were presented with the tones binaurally through headphones at intervals between 1.75 and 2.25 sec. The interval was varied to reduce ERP components associated with preparation and expectation. Participants were instructed to press a button in response to the infrequent target. The infrequent target was defined in terms of pitch for controls and in terms of either pitch or color for the synesthetes. Speed and accuracy of their responses was recorded. The participants were first presented with a practice block containing 10 trials, followed by two 3-min-long blocks each consisting of 80 trials, 10 of which were infrequent targets.

The visual stimulus conditions followed the same procedure as the auditory ones, except that colored squares were presented in place of tones and the instruction was to respond to the infrequent target color. The colors appeared as 2×2 -cm squares in the center of an otherwise black screen (a visual angle of 1.3 degrees). We did not predict a difference between synesthetes and controls in these conditions (except in the two synesthetes who report visual-to-auditory synesthesia) given that for most synesthetes and all controls, colors do not elicit sounds. However, it is also conceivable that colors could

implicitly activate auditory processes in synesthetes as similar effects have been noted in grapheme-color synesthesia (Cohen Kadosh et al., 2007). Finally, the two synesthetes who reported experiencing sounds from color were presented with colors and were asked to attend to (and respond to) the sound associated with the infrequent color.

EEG Acquisition and ERP Analysis

The EEG was continuously recorded from 31 silver/silver chloride electrodes. Twenty-nine of these were mounted in an elastic cap placed according to an equidistant montage (montage 10; www.easycap.de), and the remaining electrodes were situated on the right and left mastoid processes. The recordings were made with reference to the midfrontal electrode, Fz, and were subsequently rereferenced off-line to the algebraic average of the right and left mastoids. Blinks and other vertical eye movements were monitored by recording bipolar EOGs from a pair of electrodes placed above and below the right eye. Horizontal eye movements were measured by recording EOGs from electrodes on the outer right and the left canthi. Impedances for all electrodes were always kept below 10 k Ω . The EEG and the EOG signals were amplified with a bandwidth of 0.3–35 Hz (3-dB roll-off) and sampled at a rate of 250 Hz. The EOG activity was visually monitored during data collection to ensure that the participants' eyes were kept open and that they did not blink at regular intervals.

ERPs were averaged off-line according to condition and electrode site for 1024 msec epochs, including a 100-msec prestimulus baseline. All ERP waveforms were based on a minimum of 105 artifact-free trials (75% of trials). Artifacts were both automatically detected and monitored based on visual inspection at an individual level without knowledge of condition. Specifically, trials were removed if they were contaminated at any electrode site by artifacts such as horizontal or nonblink vertical eye movements, A/D saturation, or EEG drifts larger than 50 μ V. Blink artifacts were minimized by estimating and correcting their contribution to the ERP waveform using a standard linear regression technique (Rugg, Mark, Gilchrist, & Roberts, 1997). Incorrect trials were excluded from both the ERP and the behavioral analyses.

All analyses focused on the ERPs elicited by the frequent stimuli. The analyses comprised two primary comparisons. First, ERPs elicited by tones when attention had to be directed to auditory information were compared across synesthetes and controls. Second, in synesthetes only, ERPs for tones were contrasted depending on whether attention was paid to the auditory stimuli or the visual experience they evoked. Additional analyses were directed at the two bidirectional synesthetes. The ERPs elicited by visual stimuli in these individuals were compared with those elicited in the eight

unidirectional synesthetes and 10 controls. Auditory ERPs were also compared across unidirectional and bidirectional synesthetes.

RESULTS

Response Times

Participants correctly identified the infrequent stimulus on 99.8% of trials. RTs for individual trials were removed as outliers if they were more than three standard deviations above the mean RT for each subject in each condition. The synesthetes and controls did not differ significantly from each other for the response time for detecting the auditory targets (synesthetes: mean = 442 msec, $SD = 79$; controls: mean = 419 msec, $SD = 71$), $t(18) = .52$, *ns*, or the visual targets (synesthetes: mean = 497 msec, $SD = 89$; controls: mean = 460 msec, $SD = 74$), $t(18) = .78$, *ns*. The synesthetes responded just as fast when asked to attend to the synesthetic color of a tone (mean = 467 msec, $SD = 98$) as when asked to attend to the pitch of the tones (mean = 442 msec, $SD = 79$), $t(9) = 1.41$, *ns*.

ERP Results

For the analyses of auditory stimuli, the 10 synesthetes were treated as a single group as all reported auditory-to-visual synesthesia. It will be shown later that the two synesthetes who additionally report visual-to-auditory synesthesia do not differ in their response to auditory stimuli relative to the other eight synesthetes. For the analyses of visual stimuli, the synesthetes are divided according to their reported experiences.

Auditory Stimuli: Differences between Synesthetes and Controls

At debriefing, the synesthetes did report color experiences to the auditory stimuli. The group average ERP waveforms elicited by the auditory stimuli when attending to the pitch of tones in synesthetes and controls at all electrode sites are presented in Figure 1. Both synesthetes and controls showed identifiable N1, P2, and N2 peaks that are characteristic of auditory-evoked potentials (AEP; Picton, 1990). For both groups, the N1 was maximal over midline frontal sites, the P2 over midline central sites, and the N2 over frontal sites (see Figure 2). A small positive deflection after 200 msec corresponds to the offset of the sound. The waveforms of the controls were generally more negative going than those of synesthetes, particularly at frontocentral electrode sites. Synesthetes elicited a smaller N1 and N2 than controls, and a negative slow wave around 400–800 msec apparent in the controls was virtually absent in the synesthetes.

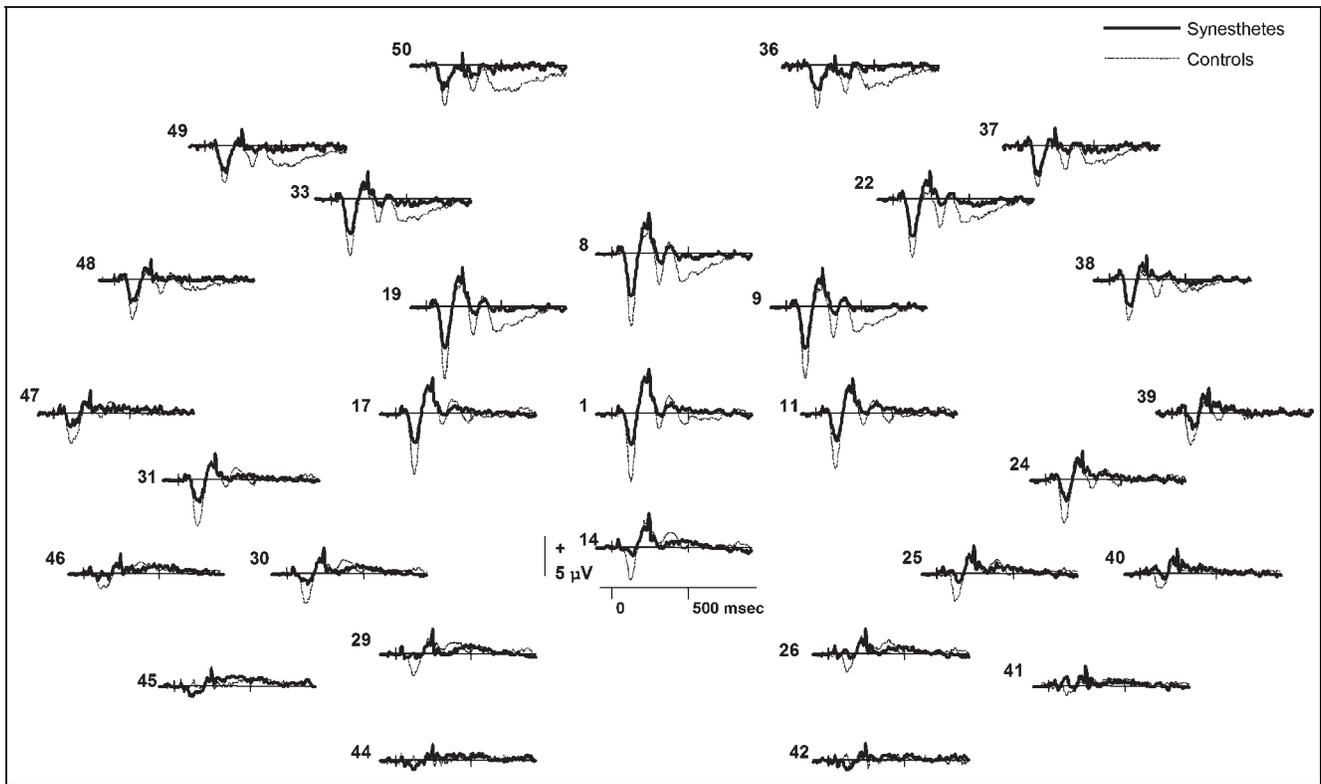


Figure 1. Group-averaged ERP waveforms at all 29 recorded electrode sites for synesthetes and controls when attending to the pitch of an auditory stimulus (thick line = controls, thin line = synesthetes). Positive values are plotted upward.

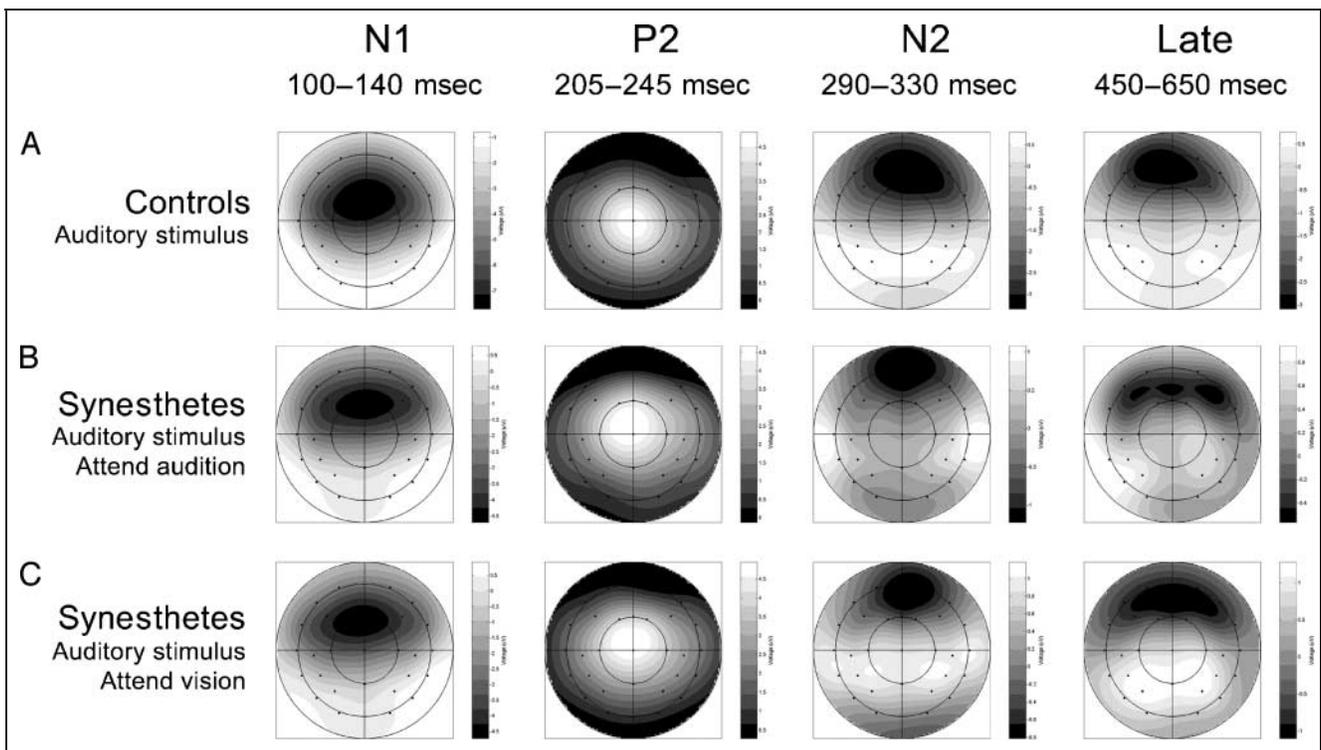


Figure 2. Voltage spline maps showing the distribution of the N1, P2, N2, and late negative deflection elicited by unimodal auditory stimuli across the scalp. The maps are range scaled. Warm colors represent positive values; cold colors represent negative values. Maps are shown separately for controls and synesthetes, when synesthetes attend to either the auditory stimulus or the visual illusion.

The analyses of the ERP data focused on the three known AEP deflections, N1, P2, and N2. The auditory P1 was not prominent and so was not considered. These deflections were quantified by centering 40-msec-wide latency regions on the deflections' maximum peaks based on visual inspection of the grand average (100–140, 205–245, and 290–330 msec, respectively) and measuring the mean amplitudes with relation to the mean of the 100-msec prestimulus baseline. A late latency region of 450–650 msec was additionally chosen for analysis to capture the visible differences on the later negative slow wave. The spline maps illustrating the scalp distributions of the deflections for synesthetes and controls are shown in Figure 2. The analyses were performed across all 29 electrode sites to consider changes in amplitude as well as scalp distribution.

At each latency region, 2×29 ANOVAs were conducted to assess the between-subjects effect of group (synesthetes vs. controls), the within-subjects effect of electrode site (29 levels), and the interaction between them. All ANOVAs used the Greenhouse–Geisser corrections for sphericity violations (Keselman & Rogan, 1980). Significant main effects of electrode site were found at all latency regions ($p < .001$). The ANOVA on the N1 latency region found a main effect of group, $F(1,18) = 6.46$, $p = .02$, but no interaction between group and electrode site, $F(1.97,35.42) = 1.35$, *ns*. This indicates that the amplitude of the N1 deflection was significantly larger in controls than in synesthetes and that this did not vary as a function of scalp location. The ANOVAs on the P2 and the N2 latency regions demonstrated no group main effects, $F(1,18) = .17$, *ns* and $F(1,18) = 2.55$, *ns*, respectively, or Group \times Site interactions, $F(2.03,36.48) = .57$, *ns* and $F(1.73,31.10) = 3.21$, $p = .06$, respectively. The final ANOVA on the 450- to 650-msec latency region revealed both a group main effect, $F(1,18) = 4.91$, $p = .04$, and an interaction between group and electrode site, $F(2.29,41.18) = 4.95$, $p = .009$. The interaction did not remain significant, $F(2.66,47.85) = .49$, *ns*, after the data were normalized using the max/min scaling procedure of McCarthy and Wood (1985). This suggests that the interaction was due to amplitude differences between synesthetes and controls, with synesthetes showing a virtual absence of this late negative deflection.

To look more closely for evidence of a visual potential evoked by auditory information in synesthetes relative to controls, we constructed difference waves at 50-msec intervals over the four most posterior scalp locations (sites 26, 29, 42, and 44 of montage M10). These liberal analyses only revealed significant differences in two intervals at one or two sites [site 29 between 250–300 and 350–400 msec, $t(18) = 2.30$ and 3.26 , respectively, both $p < .034$; site 26 between 350 and 400 msec, $t(18) = 2.18$, $p < .05$]. Importantly, in all instances, there was a reduced amplitude in synesthetes. This is qualitatively different to reports of auditory-evoked visual potentials

found in multisensory illusions (Shams et al., 2001), in acquired synesthesia (Rao et al., 2007), and in early infancy (Neville, 1995). Thus, there is no evidence in the present data that tones evoke a visual potential in synesthetes.

Auditory Stimuli: Effects of Attention on Synesthesia

Further analyses compared the conditions in which synesthetes attended to either the auditory percept or their visual synesthetic experience when they were presented with a unimodal auditory stimulus. The group-averaged ERP waveforms for these conditions are presented in Figure 3. The first thing to note is that there are few visible differences between these conditions. To determine any effects of attention, we carried a 2 (attend auditory vs. visual) \times 29 (electrode sites) ANOVA out at each latency region previously used in the analysis of the AEP. No difference was found for attention at any of the latency regions, and only the late latency region, 450–650 msec, demonstrated a significant interaction between electrode site and attention, $F(3.01,27.12) = 5.08$, $p = .006$. This reflects increased positivity at more posterior sites accompanied by increased negativity at more anterior sites (i.e., attention increases the amplitude of this late deflection). After scaling the data, this interaction was no longer significant, $F(2.32,20.88) = 1.04$, *ns*. This suggests that there is an effect of attention, but the effect is to modulate a deflection rather than introduce a new deflection. All latency regions showed an electrode site main effect ($p < .05$), except for the N2 region, $F(1.96,17.65) = 2.22$, *ns*.

These findings suggest that attention toward or away from their synesthesia has a late modulatory influence on the ERP to auditory stimuli, but differences between synesthetes and controls emerge far earlier (from 100 msec). These differences are thus unlikely to be due to between-group differences in attention. The fact that some attention-related differences were observed implies that synesthetes were complying with instructions. A visual inspection of the data from those synesthetes who reported colors in their mind's eye ($n = 3$) versus elsewhere ($n = 7$) did not reveal extra deflections, although we lacked the power to assess this statistically.

Visual Stimuli

The group-averaged waveforms for visual-evoked potentials are shown in Figure 4. As with the auditory potentials, the analysis of the visual-evoked potentials concentrated on two known deflections, the visual P1 and the N1, together with a later positive deflection that was particularly prominent in the synesthetes. In both groups, the P1 was maximal at occipital sites, the N1 at left temporal sites, and the late positive deflection at central/parietal sites. These deflections were quantified

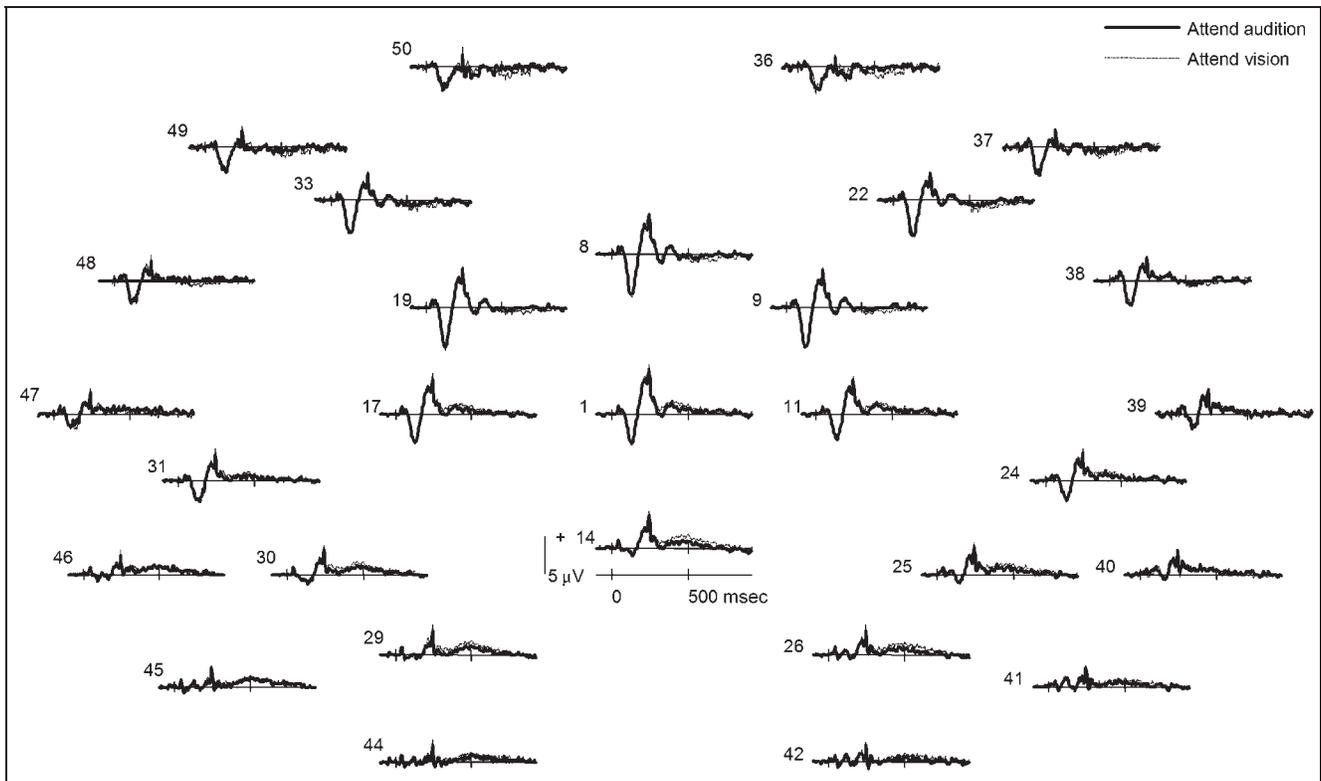


Figure 3. Group-averaged ERP waveforms for synesthetes when presented with an auditory stimulus and attending to either the pitch (thick line) or the synesthetic color (thin line).

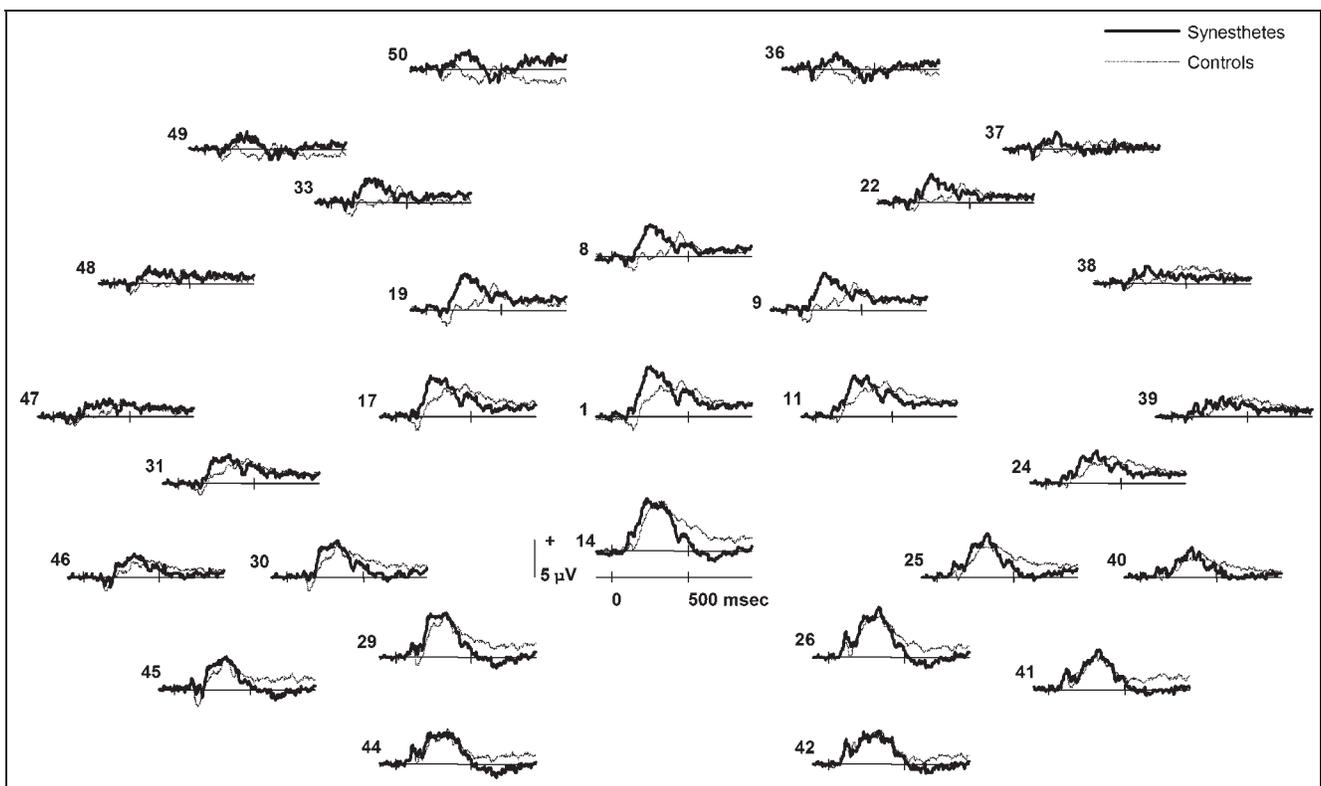


Figure 4. Group-averaged ERP waveforms for synesthetes ($n = 8$) experiencing auditory–visual synesthesia (but not visual–auditory synesthesia) and controls when attending to the color of a visual stimulus (thick line = controls, thin line = synesthetes).

by centering 40-msec-wide latency regions on the maximum peaks based on visual inspection of the group averages (P1 = 80–120 msec; N1 = 120–160 msec; late positive = 230–270 msec). These latency regions are in line with those reported elsewhere (Luck, 2005). The mean amplitudes were measured in relation to the mean of the 100-msec prestimulus baseline. Given that two of our synesthetes report conscious auditory experiences from vision, these synesthetes were considered separately (see next section). Thus, the analyses compared 10 controls and 8 synesthetes using a 2×29 ANOVA contrasting group and electrode site. In general, the differences between synesthetes and controls were less pronounced than for auditory stimuli, consistent with their phenomenological reports. Nonetheless, some differences were found. The synesthetes showed a reduced visual N1 deflection, which manifested itself as a main effect of group, $F(1,16) = 8.62$, $p = .01$, but the effect of group did not interact with site. The late positive deflection showed a Group \times Site interaction that was of borderline significance, $F(2.93,46.82) = 2.79$, $p = .052$. There were no significant group effects for the other deflections and no interactions between group and site. The main effect of site was significant for all three deflections. In general, the results suggest that the visual perception of color in synesthesia is atypical even if color does not evoke overt synesthetic perceptions. This could either reflect implicit bidirectionality between colors and sounds based on their lifelong association or lack of inhibition between areas (Cohen Kadosh et al., 2005) or more fundamental differences in color processing. It is important for future studies to contrast colors that are associated with sounds relative to colors that are not to distinguish between these interpretations.

Visual–Auditory Synesthetes

For the two synesthetes who report auditory experiences to color, the analyses focused on the same three deflections as described above. However, these deflections were analyzed over a limited selection of electrodes to minimize a type II error. The electrodes chosen were those in which the visual-evoked potential was noted to be maximal in the grand average of the controls (electrodes 42, 19, and 14 for the visual P1, N1, and late positive deflection, respectively). The mean amplitudes for the two synesthetes were compared with the other synesthetes ($n = 8$), the controls ($n = 10$), and the combined group ($n = 18$) using the modified t test reported by Crawford and Howell (1998). This procedure compares a single score to a group mean, taking into account the size of the group (the degrees of freedom is $N - 1$, where N is the size of the comparison group). The data are summarized in Figure 5. One of the synesthetes, JR, had a significantly accentuated P1 deflection (compared with the combined group

and synesthete group, respectively), $t(17) = 3.33$, $p < .001$ and $t(7) = 3.14$, $p < .05$. In contrast, the other synesthete, SL, had a significantly more negative-going late deflection at 230–270 msec compared with the synesthete group, $t(7) = 2.45$, $p < .05$. This could be due to a longer latency of the earlier N1 deflection. The same pattern was found when these two synesthetes were explicitly instructed to attend to the auditory (i.e., synesthetic) component of the color (JR had a visual P1 of $+7.59 \mu\text{V}$ and SL had a late “positive” deflection of $-2.83 \mu\text{V}$). Thus, the basic finding is unlikely to be due to the focus of attention per se. It is to be noted that these two synesthetes did not stand out as anomalous on the auditory tasks. We reanalyzed the data for the four auditory deflections (N1, P2, N2, and late negative deflection) over their maximal electrode sites comparing JR and SL to the other synesthetes for both auditory presentation conditions (attend visual and attend auditory). No significant differences were found. As such, these two synesthetes do not appear to be globally different. The differences that they manifest are limited to the condition in which they report a different perceptual experience to the other synesthetes. This provides the first evidence for the authenticity of this type of synesthesia. We shall consider potential reasons for the discrepancy between the two vision-to-auditory synesthetes in the Discussion section.

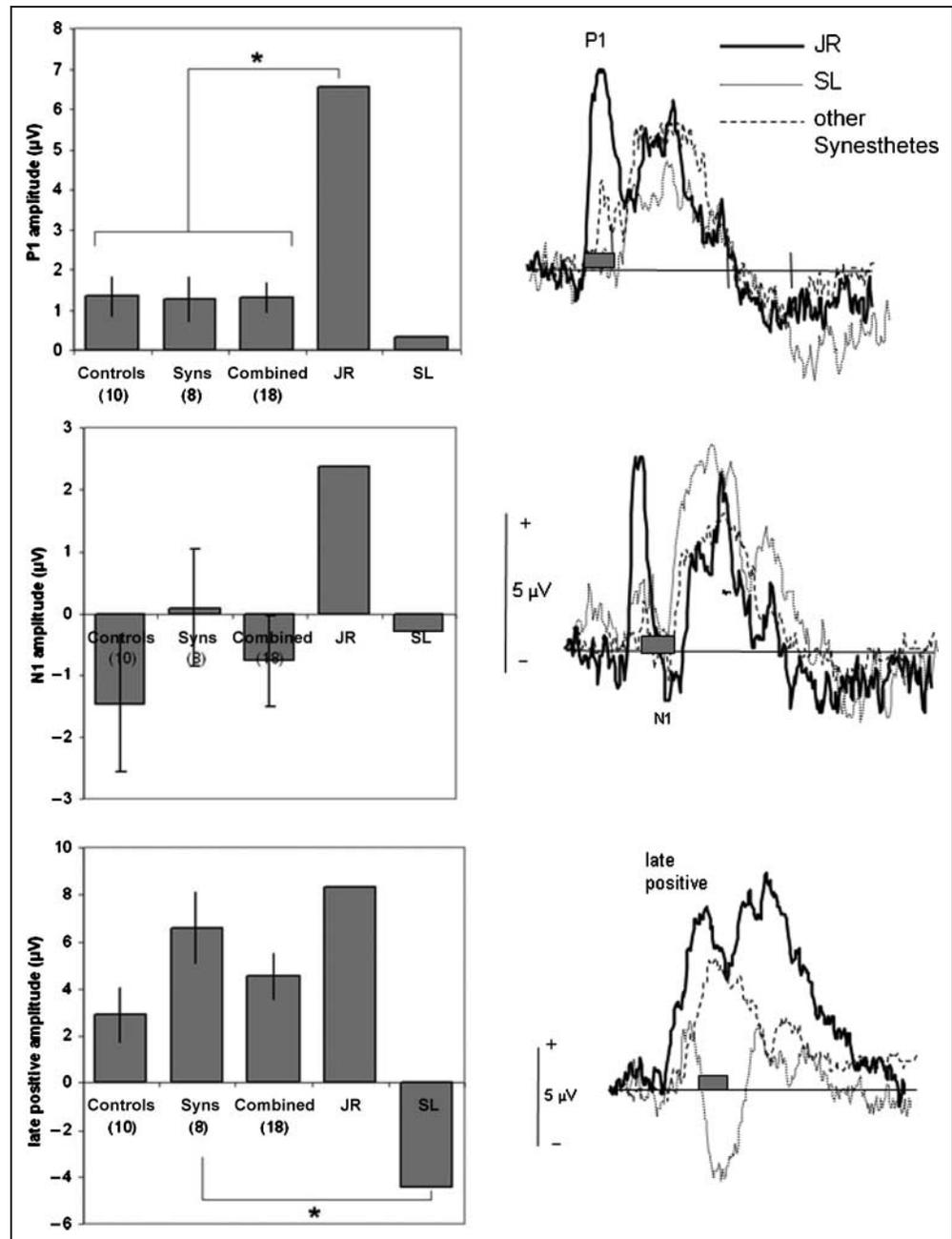
DISCUSSION

In summary, our results demonstrate significant electrophysiological differences between synesthetes and controls presented with unimodal auditory stimuli, in which the auditory tone is reliably associated with a visual experience in the synesthetes but not the controls. These differences reflect modulations of deflections of the AEP (i.e., N1) together with a greatly attenuated late slow negative deflection in the synesthetes. These differences are found irrespective of whether the synesthetes were instructed to attend to the pitch of the tone or the color of the tone (i.e., attention directed away or toward the synesthetic experience), although there were some late differences between these conditions. There was no evidence of an auditory-evoked visual potential over occipital electrode sites.

Our results differ from those showing greater electrophysiological responses over occipital sites in response to sounds in infants (Neville, 1995), in the “double-flash” illusion (Shams et al., 2001), and in a case of acquired synesthesia following blindness (Rao et al., 2007). As such, we suggest that a different mechanism is in operation in developmental synesthesia to that which is tapped in these studies. That is, we suggest that developmental synesthesia does not reflect long-range projections between early auditory and early visual areas (e.g., A1 to V1).

We also did not find evidence that auditory stimuli elicited a distinct multisensory ERP deflection in synesthetes.

Figure 5. Group-averaged ERP waveforms for three electrode sites when presented with unimodal visual stimuli (colors). The graphs on the left show average amplitudes over a 40-msec time window relative to a 100-msec prestimulus baseline for 2 synesthetes who report conscious experiences to tones (SL and JR; $*p < .05$), for 8 synesthetes who do not but who show the reverse type of synesthesia (sounds trigger colors), and for 10 control participants who lack synesthesia. Error bars show 1 SEM.



The data thus do not strongly support the idea of cross-model transfer in this type of synesthesia. Nonetheless, our results are more consistent with this account. A remarkable feature of our results is that no differences in scalp distribution were observed for synesthetes relative to controls. If multisensory AV regions are spatially close to those normally involved in auditory perception, this would explain the lack of a Group \times Site interaction in addition to the significant group effects that were observed. For instance, there could be anomalous cross-activation between adjacent regions of auditory cortex and regions in the superior temporal gyri/sulcus that are implicated in audiovisual perception (e.g., Calvert, 2001). It is also to be noted that tradition-

ally defined unimodal auditory areas can sometimes respond to nonauditory events. Brosch, Selezneva, and Scheich (2005) report that neurons in the monkey primary auditory cortex and posterior belt areas respond to the presence of visual stimuli when the visual stimulus is predictive of a subsequent auditory event (but not in other circumstances). These neurons also respond to unimodal auditory stimuli. That is, they have audiovisual response properties rather than being strictly auditory or visual. It is conceivable that neurons such as these, which lie in or around the cortical auditory pathways, contribute to audiovisual synesthesia. Specifically, synesthetes may contain far more neurons with audiovisual response properties than unimodal auditory re-

sponses. Although direct evidence is lacking, the present research is consistent with this view given our findings of early modulations of the AEP. Our results are also compatible with the idea that there is disinhibition or unmasking of visual neurons within predominantly auditory regions (Cohen Kadosh & Walsh, 2006).

Our research does not disprove the notion that infants may have some form of synesthesia, but it does raise questions about whether this type of synesthesia is directly comparable to that found in adults even when the synesthesia is for relatively simple pairings (pure tones and color). Our results are entirely consistent with a single case study of developmental synesthesia that also recorded auditory-evoked ERPs (Rizzo & Eslinger, 1989). They are also broadly consistent with recent results reported by Beeli et al. (2008) showing reduced amplitudes and/or longer latencies in the auditory N1 and P2 deflections for colors induced for spoken graphemes and words. The question of why synesthetes should have reduced deflections when they are reporting an “extra” experience remains to be fully explored, but it is by no means a general feature of synesthesia (Cohen Kadosh et al., 2007; Schiltz et al., 1999). It is possible that the reduced negativity actually reflects an additional presence of a positive-going deflection. It is also possible that early auditory processing is attenuated as a result of their habitual experiences of accompanying vision.

A novel aspect of the present study concerns the processing of colored visual stimuli in synesthesia. For most of the synesthetes studied here, colors do not induce any synesthetic experiences. Nevertheless, there were significant early (120–160 msec) differences in the visual-evoked potential between synesthetes and controls. This suggests that the presence of synesthesia has repercussions for visual perception even when the visual stimuli do not themselves elicit synesthesia (although there could still be implicit bidirectionality; for example, see Cohen Kadosh et al., 2005). This is consistent with a recent study showing that grapheme-color synesthetes have better perceptual discrimination of color and better memory for colors than nonsynesthetic controls (Yaro & Ward, 2007), and synesthetes who experience touch have enhanced spatial tactile discrimination even when the task does not involve their synesthesia (Banissy & Ward, 2008). These results suggest that there are more fundamental differences in perceptual processing in people with synesthesia that reveal themselves both behaviorally and electrophysiologically rather than synesthesia being solely the presence of some additional sensory attribute.

We also report the first empirical evidence to support the existence of vision-to-auditory synesthesia, although this pattern has been briefly noted before for one of the participants that we studied (Baron-Cohen et al., 1996). The two synesthetes who report this symptom showed abnormal visual-evoked potentials (relative to other syn-

esthetes who do not report this experience) but showed normal AEPs relative to other synesthetes. This suggests that they are not outliers on all measures. The two synesthetes did, however, differ from each other. JR showed a greater amplitude of visual P1 (80–120 msec) and SL showed a negative-going later deflection (230–270 msec) that tended to be positive in other participants. Individual differences in the timing of synesthetic experiences have been postulated in grapheme-color synesthesia (Ward, Salih, Li, & Sagiv, 2007; Dixon, Smilek, & Merikle, 2004) with one suggestion being that externalized color percepts occur earlier than those reported in “the mind’s eye.” This could be one explanation of the difference that we find (JR experiences external photisms for sound-induced vision but SL does not). For the present purposes, it is sufficient to note that this type of synesthesia can occur. It poses a potential challenge to present theories (Hubbard & Ramachandran, 2005) that fail to account for why it is more common in one direction (auditory to visual) than the other (visual to auditory).

In conclusion, the present findings suggest that adult forms of auditory–visual synesthesia with a developmental origin utilizes pathways based on cross-modal transfer in and around the normal auditory processes rather than direct audiovisual pathways.

Acknowledgments

All research was conducted at the Institute of Cognitive Neuroscience, University College London. Stimulus presentation was programmed with the Cogent2000 software of the physics group of the Wellcome Trust Centre for Neuroimaging. Both L. J. O. and J. W. are supported by the Wellcome Trust.

Reprint requests should be sent to Jamie Ward, Department of Psychology, University of Sussex, Falmer, Brighton, BN1 9QH, UK, or via e-mail: jamiew@sussex.ac.uk.

REFERENCES

- Banissy, M., Walsh, V., & Ward, J. (2008). “Mirror-touch” synaesthesia is associated with enhanced tactile discrimination. *Journal of Cognitive Neuroscience Annual Meeting Program* (Supplement). Abstract presented at 15th annual meeting of Cognitive Neuroscience Society, p. 52.
- Baron-Cohen, S. (1996). Is there a normal phase of synaesthesia in development? *Psyche*, 2.
- Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: Prevalence and familiarity. *Perception*, 25, 1073–1079.
- Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. (1993). Colored speech perception: Is synaesthesia what happens when modularity breaks down? *Perception*, 22, 419–426.
- Beeli, G., Esslen, M., & Jancke, L. (2008). Time course of neural activity correlated with colored-hearing synesthesia. *Cerebral Cortex*, 18, 375–381.
- Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *Journal of Neuroscience*, 25, 6797–6806.

- Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, *11*, 1110–1123.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*, 593–596.
- Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, *14*, 427–438.
- Cohen Kadosh, R., Cohen Kadosh, K., & Henik, A. (2007). The neuronal correlates of bi-directional synaesthesia: A combined ERP and fMRI study. *Journal of Cognitive Neuroscience*, *19*, 2050–2059.
- Cohen Kadosh, R., & Henik, A. (2006). When a line is a number: Color yields magnitude information in a digit-color synaesthete. *Neuroscience*, *137*, 3–5.
- Cohen Kadosh, R., Sagiv, N., Linden, D. E. J., Robertson, L. C., Elinger, G., & Henik, A. (2005). Do colors influence numerical cognition in synaesthesia? *Journal of Cognitive Neuroscience*, *17*, 1766–1773.
- Cohen Kadosh, R., & Walsh, V. (2006). Rewired or cross-wired brains? *Current Biology*, *16*, R962–R963.
- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, *22*, 319–349.
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *Clinical Neuropsychologist*, *11*, 482–486.
- Day, S. (2005). Some demographic and socio-cultural aspects of synesthesia. In L. C. Robertson & N. Sagiv (Eds.), *Synesthesia: Perspectives from cognitive neuroscience*. Oxford: Oxford University Press.
- Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector vs. associator synaesthetes. *Cognitive, Affective, and Behavioral Neuroscience*, *4*, 335–343.
- Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*, *10*, R731–R735.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, *22*, 5749–5759.
- Frith, C. D., & Paulesu, E. (1997). The physiological basis of synaesthesia. In S. Baron-Cohen & J. E. Harrison (Eds.), *Synaesthesia: Classic and contemporary readings*. Oxford: Blackwell.
- Giraud, A. L., Price, C. J., Graham, J. M., Truy, E., & Frackowiak, R. S. J. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*, 657–663.
- Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron*, *48*, 509–520.
- Innocenti, G. M., Berbel, P., & Clarke, S. (1988). Development of projections from auditory to visual areas in the cat. *Journal of Comparative Neurology*, *272*, 242–259.
- Keselman, H. J., & Rogan, J. C. (1980). Repeated measures *F* tests and psychophysiological research: Controlling the number of false positives. *Psychophysiology*, *17*, 499–503.
- Lewkowicz, D., & Turkewitz, G. (1980). Cross-modal equivalence in infancy: Auditory–visual intensity matching. *Developmental Psychology*, *16*, 597–607.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes*. Cambridge, MA: MIT Press.
- Mattingley, J. B., Payne, J., & Rich, A. N. (2006). Attentional load attenuates synaesthetic priming effects in grapheme-color synaesthesia. *Cortex*, *42*, 213–221.
- Maurer, D. (1997). Neonatal synaesthesia: Implications for the processing of speech and faces. In S. Baron-Cohen & J. E. Harrison (Eds.), *Synaesthesia: Classic and contemporary readings*. Oxford: Blackwell.
- Maurer, D., & Maurer, C. (1988). *The world of the newborn*. New York: Basic Books.
- Maurer, D., & Mondloch, C. J. (2006). The infant as synesthete? *Attention and performance XXI* (pp. 449–471).
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208.
- Mills, C. B., Boteler, E. H., & Larcombe, G. K. (2003). “Seeing things in my head”: A synesthete's images for music and notes. *Perception*, *32*, 1359–1376.
- Neville, H. J. (1995). Developmental specificity in neurocognitive development in humans. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.
- Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., et al. (2002). Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience*, *5*, 371–375.
- Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D. G., Goldstein, L., Heather, J., et al. (1995). The physiology of colored hearing: A PET activation study of color-word synaesthesia. *Brain*, *118*, 661–676.
- Picton, T. W. (1990). Auditory evoked potentials. In D. D. Daly & T. A. Pedley (Eds.), *Current practice of clinical electroencephalography* (2nd ed.). New York: Raven Press.
- Rao, A. L., Nobre, A. C., Alexander, I., & Cowey, A. (2007). Auditory evoked visual awareness following sudden ocular blindness: An EEG and TMS investigation. *Experimental Brain Research*, *176*, 288–298.
- Rizzo, M., & Eslinger, P. J. (1989). Colored hearing synesthesia: An investigation of neural factors. *Neurology*, *39*, 781–784.
- Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, *50*, 19–26.
- Rugg, M. D., Mark, R. E., Gilchrist, J., & Roberts, R. C. (1997). ERP repetition effects in indirect and direct tasks: Effects of age and interitem lag. *Psychophysiology*, *34*, 572–586.
- Sagiv, N., Heer, J., & Robertson, L. (2006). Does binding of synesthetic color to the evoking grapheme require attention? *Cortex*, *42*, 232–242.
- Schiltz, K., Trocha, K., Wieringa, B. M., Emrich, H. M., Johannes, S., & Munte, T. F. (1999). Neurophysiological aspects of synesthetic experience. *Journal of Neuropsychiatry and Clinical Neuroscience*, *11*, 58–65.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, *408*, 788.
- Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *NeuroReport*, *12*, 3849–3852.
- Simner, J., Glover, L., & Mowat, A. (2006). Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex*, *42*, 281–289.
- Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., et al. (2006). Synaesthesia: The prevalence

- of atypical cross-modal experiences. *Perception*, *35*, 1024–1033.
- Thornley Head, P. D. (2006). Synaesthesia: Pitch-color isomorphism in RGB-space? *Cortex*, *42*, 164–174.
- Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-color synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, *42*, 264–280.
- Ward, J., Salih, S., Li, R., & Sagiv, N. (2007). Varieties of grapheme-color synaesthesia: A new theory of phenomenological and behavioural differences. *Consciousness and Cognition*.
- Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage*, *31*, 1247–1256.
- Yaro, C., & Ward, J. (2007). Searching for Shereshevskii: What is superior about the memory of synaesthetes? *Quarterly Journal of Experimental Psychology*, *60*, 682–696.

Copyright of *Journal of Cognitive Neuroscience* is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.