Visual Experiences in the Blind induced by an Auditory Sensory Substitution Device

Jamie Ward[1] and Peter Meijer[2]

[1]Department of Psychology, University of Sussex, Brighton, U.K.

[2]Eindhoven, The Netherlands

Preprint, to appear in *Consciousness and Cognition*

DOI: http://dx.doi.org/10.1016/j.concog.2009.10.006

Running Head: Visual experiences in the blind

Address correspondence to :-

Jamie Ward,

Department of Psychology,

University of Sussex,

Falmer, Brighton,

BN1 9QH, U.K.

Tel. : +44 (0)1273 876598

Fax. : +44 (0)1273 678058

E-mail : jamiew@sussex.ac.uk

Ward, J., & Meijer, P. (2009). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. doi:10.1016/j.concog.2009.10.006

Abstract

In this report, the phenomenology of two blind users of a sensory substitution device – "The vOICe" – that converts visual images to auditory signals is described. The users both report detailed visual phenomenology that developed within months of immersive use and has continued to evolve over a period of years. This visual phenomenology, although triggered through use of The vOICe, is likely to depend not only on online visualization of the auditory signal but also on the users' previous (albeit distant) experience of veridical vision (e.g. knowledge of shapes and visual perspective). Once established, the sensory substitution mapping between the auditory and visual domains is not confined to when the device is worn and, thus, may constitute an example of acquired synaesthesia.

Keywords: sensory substitution, mental imagery, vOICe, visual consciousness, synaesthesia/synesthesia, blind

Ward, J., & Meijer, P. (2009). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. doi:10.1016/j.concog.2009.10.006

1. Introduction

Sensory substitution devices artificially convert information normally delivered to one sense into a representation that is compatible with an alternative sense. For example, by converting visual information into auditory or tactile signals it is possible for the blind and visually impaired to acquire information about the world that is not normally accessible through audition and touch. Sensory substitution devices convert vision to sound or touch, but does the brain reconvert the auditory and tactile stimuli into vision or does it process them solely within the modality within which they are received? There is convincing evidence from both tactile and auditory sensory substitution devices that users of these devices recruit regions of the brain that are normally specialized for vision (e.g. Poirier, De Volder, & Scheiber, 2007). However, this in itself does not necessarily mean that the experience of the user is one of seeing. In the case of tactile devices, there are phenomenological reports that the users of the device claim that the perceptual experience resembles seeing more than touch (Bach-y-Rita, 1972; Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969). However, this interpretation is not universally accepted (Block, 2003). Although there has been much recent research on auditory sensory substitution devices, there have been no detailed phenomenological reports from blind expert users of the device. This paper will provide phenomenological reports of visual experiences in blind users of an auditory sensory substitution device.

The most detailed phenomenological descriptions to date of visual-like experiences induced from a sensory substitution device come from the tactile-visual sensory substitution (TVSS) explorations of Bach-y-Rita and colleagues. The original device used a 20 × 20 array of tactile pins typically mounted on participants' backs (Bach-y-Rita et al., 1969). A later version was based on electrical stimulation of the tongue (Sampaio, Maris, & Bach-y-Rita, 2001). In both instances, participants used a hand-held or head-mounted camera to explore the scene. The visual view of the camera was then converted into an array of pixels, and the brightness levels of the different pixels were mapped to corresponding levels of tactile stimulation by the tactors in a tactile array. Starting with lines and basic shapes, then simple objects (e.g. a cup), blind participants could eventually discriminate different staff: "That is Betty; she is wearing her hair down today and does not have her glasses on; her mouth is open and she is moving her right hand from her left side to the back of her head." (Bach-y-Rita, 1972, p. 6). Although the tactile array was normally placed on the back, within 5–15 h of training the participants reported an externalization of their sensations in front of them, i.e. in front of the camera. This was even maintained if the tactile array was moved to the abdomen. However, externalization only occurred if the participant had full control of the camera during training. If the experimenter increased the zoom level on the lens then the participant experienced a looming sensation, with raised arms and backward movement. They lurched backwards, away from an apparent approaching object, even though the tactile array was on their back (Bach-y-Rita, 1972).

Are blind participants using the TVSS and its later tactile variations seeing? Bach-y-Rita (1972) puts visual phenomenology to one side to give an affirmative answer: "If a subject without functioning eyes can perceive detailed information in space, correctly localize it subjectively, and respond to it in a manner comparable to the response of a normally sighted person, I feel justified in applying the term 'vision'." (Bach-y-Rita, 1972, p. ix). Others, such as Keeley (2002), have given a negative answer noting that acquiring visual information from touch is not seeing, it is touch alone (again, phenomenology was left aside). Blind participants trained to use the tongue-version of TVSS experienced tactile sensations when their occipital ('visual') cortex was stimulated using Transcranial Magnetic Stimulation (TMS; Kupers et al., 2006), but we do not know what they experienced when using the device itself. O'Regan and Noë (2001) in contrast, do place an importance on the phenomenological descriptions of users. They argue that visual experiences occur when our motor system engages with the world in particular ways (see also Hurley & Noe, 2003). Conscious visual experiences feel the way that they do because of the way that visual information changes when we move our eyes, head and body. For example, a circle may look elliptical if we adjust our viewing angle, but a coin felt in the hand will not change shape under haptic exploration. According to O'Regan and Noë (2001) different kinds of sensory experiences reflect different sensorimotor contingencies. In the TVSS, the sensory-motor contingencies resemble vision even though the input is tactile and, hence, the phenomenology is more like vision than touch.

Not everyone accepts that the phenomenological reports from TVSS resemble vision. Block (2003) argues that the tactile input in TVSS may use spatial rather than visual mechanisms, and that the reported phenomenology is consistent with spatial rather than visual experiences. Even Bach-y-Rita (1972) seems to attribute the spatial, rather than strictly visual, nature of TVSS to sensory-motor correspondences: "Self-induced camera motion appears to be analogous to eye movements… external localization of percepts depends critically on such movements… a plausible hypothesis is that a translation of the input that is precisely correlated with self-generated movement is the necessary and sufficient condition for the experienced phenomena to be attributed to a stable outside world." (Bach-y-Rita, 1972, p. 99). It is also to be noted that some of the user reports were from congenital blind participants who described their use of the device in visual terms ('seeing', 'watching', 'field of view') despite having no history of vision (Guarnier.G, 1974)[1]. Claims for visual experiences should therefore be interpreted very cautiously.

Given controversy in the visuo-tactile domain, it is important to look for complementary evidence in other areas. Another sensory substitution device that has recently been increasingly applied in studies is The vOICe system, which converts visual views into an auditory representation termed a 'soundscape' (Meijer, 1992). The basic hardware requirements are simpler than for a tactile system, requiring only a camera (e.g. a webcam), headphones, and a computing device for running the image-to-sound conversion software (e.g. a laptop or a mobile phone). The software subsamples the image into a 176 × 64 array of pixels (PC version). The image is heard column by column, panning from left to right over 1 s (i.e. *x*-axis is mapped to time and stereo position in the auditory domain); with bright pixels sounding loud (i.e. brightness to loudness mapping); and pixels high in the view being higher in pitch (i.e. *y*-axis is mapped to frequency in the auditory domain within the range 500–5000 Hz). Thus, a white '/' would be heard as a single tone rising in pitch, a white 'X' would be heard as two simultaneous sound streams, one ascending in pitch and one descending, and a white '+' would be heard as a constant mid-pitch tone with a brief sound pulse (made up of many different frequencies) mid-way in time. Whereas in TVSS there is a spatial homology between the pixels in the image and the tactors in the array, there is no such spatial homology in The vOICe (visual space is coded spectro-temporally in the auditory domain). As such, reports of spatial phenomenology would

---

[1] To illustrate the point consider the following quote from Guarniero (1974): "I have used the word 'see' for lack of better. The difficulty is not merely one of vocabulary; rather, it is a conceptual one. Very soon after I learned how to scan, the sensations no longer felt as if they were located on my back, and I became less and less aware of the vibrating pins were making contact with my skin." (page 104). There is little to suggest that Guarniero had sensations of light and dark.

be an outcome of a more complex (synaesthesia-like) transformation of the auditory signal rather than a simple remapping from the body surface to extra-personal space.

Several studies have documented how sighted blindfolded participants are able to use The vOICe. Auvray, Hanneton, and O'Regan (2007) trained participants for approximately 15 h and found that they could perform reasonably accurately at locating objects on a table and identifying objects (from a choice of 10), including novel exemplars (e.g. a different plant to the one previously presented). Some participants reported qualitative experiences that were like vision or a non-auditory sense, particularly when trying to locate objects. Proulx, Stoerig, Ludowig and Knoll (2008) found that degree of familiarity with The vOICe improved performance in locating LEDs and objects, although blindfolding during training with The vOICe had no effect. Kim and Zatorre (2008) found that participants could match 'soundscapes' with their corresponding objects irrespective of whether they had explicit knowledge of the algorithm.

In terms of neural responses, one study found an enhanced auditory signal (using MEG) in trained blindfolded participants when they heard vOICe soundscapes, but not other sounds (Pollok, Schnitzler, Stoerig, Mierdorf, & Schnitzler, 2005). A more recent study using fMRI found a region in the lateral-occipital region to be activated in expert users of The vOICe when listening to soundscapes of shapes relative to other sounds (Amedi et al., 2007). This region may be involved in the multi-sensory processing of shape, normally visual or haptic shape, but activation by auditory 'shape' appears possible when participants have learned to interpret a shape-encoding auditory signal from a sensory substitution device.

The study by Amedi et al. (2007) is noteworthy in that the seven expert users that they tested were not all blindfolded sighted people but included two blind participants who had been using The vOICe for some time in their day-to-day environments. One of their participants, PF, is documented further below. PF has previously described her experiences of using The vOICe elsewhere (‘Towards a Science of Consciousness’, Tucson, Arizona, USA; 8th April 2002) and a transcript of relevant parts of the presentation are discussed here, with her permission, in order to explain how her experiences have changed in the intervening years. PF has also taken part in other research, including a recent TMS study in which Transcranial Magnetic Stimulation applied to the occipital lobes temporarily disrupted her ability to use and ‘see’ with The vOICe (Merabet et al., 2009). In addition, a second case is discussed. CC is similar to PF in many respects, but has not used The vOICe as extensively.

Experiences of vision (if that is what they are) induced by touch or sound can be thought of as examples of acquired synaesthesia, albeit acquired artificially through the use of a sensory substitution device (see also Proulx & Stoerig, 2006). Synaesthetic experiences have a number of core properties (e.g. Ward, 2008; Ward & Mattingley, 2006). They are, by definition, percept-like in nature. They are elicited by a stimulus (rather than being spontaneous, like an hallucination), and they are automatic insofar as they occur involuntarily. It is to be noted that in synaesthesia the experienced sensation co-exists with the inducing one rather than replacing it. Thus, for a synaesthete a sound is seen but it is also heard. The same can also be said of induced experiences from sensory substitution devices, although the extent to which it is reported as more like one sense than another may be due to the modality that is attended to.

Synaesthesia has several causes. Developmental synaesthesia occurs throughout the lifespan and has a genetic contribution (Asher et al., 2009). Acquired (or induced) synaesthesia occurs as a result of an environmental trigger in individuals who did not previously report synaesthesia. These triggers include: (1) hallucinogenic drugs such as LSD and ‘magic’ mushrooms (Hartman & Hollister, 1963) and (2) sensory deprivation including both blindness

(Afra, Funke, & Matsuo, 2009) and blindfolding (Merabet et al., 2004). Acquired synaesthetic experiences can occur within hours (induced through hallucinogen use), or in cases of blindness anywhere from as little as a few days to a year (Afra et al., 2009). This suggests that acquired synaesthesia is linked to different mechanisms of plasticity: fast unmasking of existing cross-modal connections and slower reorganisation perhaps associated with changes in synaptic connectivity (Ward, 2008). The question of whether visual (or spatial) experiences induced by auditory or tactile sensory substitution devices constitutes a third type of environmental trigger for acquired synaesthesia is partly a theoretical one (e.g. based on one's definitions) and partly an empirical one. For example, in the interviews below we consider whether long-term use of a sensory substitution device results in visual experiences from sounds even when the device is not worn, arising as a result of long-term plastic changes associated with learning to use the device.

2. Case Descriptions

PF was born in 1956 and became blind at the age of 21 as a result of an industrial accident. She currently has a small amount of light perception in the left eye only, but the right eye was burned out entirely. Before using The vOICe she relied on a cane and guide dog. She came across The vOICe software in 1998 and began using it immersively from 2000. She uses the standard settings of 1 s refresh rate and normal contrast (i.e. bright mapped to loud). She has taken part in several research studies into The vOICe (Amedi et al., 2007; Merabet et al., 2009).

PF began with a non-portable setup for the software, with a flatbed scanner upon which objects were placed:

> “I decided to take and place known objects on my scanner and then maybe I could tell what I was working with. One object that I used was a plastic drinking glass. I would place the glass on the scanner, take the scan and then whilst holding it in my hand, listening to the soundscapes, try to relate that which I was hearing to that which I was touching. This gave me a general understanding of how soundscapes worked and how they related visually.” (presentation by PF, April 2002).

She then had a portable system with a camera mounted on a baseball cap before opting for a more concealed version in which the camera is hidden in the nose bridge of a special pair of ‘spy’ sunglasses and connects to a laptop in a back pack. She describes how moving the camera from on top of a cap to the bridge of her glasses initially resulted in poor sensorimotor coordination:

> “I went down from a webcam on my [base]ball cap to a camera between my eyes and after a while I became clumsy because my mind had adapted so that the soundscapes indicated that they were looking at a door. If I had my ball cap on I’d be looking down so my body language learned to reach up to grab the door knob. When I put my glasses on it became crazy, my body had to learn a new language as to how to interpret the soundscape and reach for the doorknob at my hip. Wherever the camera is placed is very, very important.” (interview with PF, December 2007).

CC was born in 1959. Her blindness is a result of rod dystrophy. She had bad eyesight in childhood that deteriorated significantly as a teenager. She was registered blind around the age of 33 (in 1992). Her visual acuity enables her to count fingers in front of her (in both eyes) and she can notice large objects in strong contrast and movement. She came across The vOICe on the Internet in about 2001. When using it immersively, she says it took around 3 months to adapt. For most of this period she has used the camera of a mobile phone, which is worn around a necklace on her chest. The mobile phone version uses a 64 × 64 pixel image with mono sound (compared to the PC version which is 176 × 64 pixels in stereo). She uses the standard settings of 1 s refresh rate and normal contrast (i.e. bright mapped to loud). CC also has multiple sclerosis and is confined to a wheelchair.

> "I learned the sounds of shapes that Peter [Meijer] had put on his website, and so I knew what simple geometric shapes sounded like. Then I had some more immersive training with a webcam on my head, with the laptop version. It took three months for me to learn enough so that I didn't have to consciously concentrate on it." (interview with CC, November 2007)

3. What Do They Claim to 'See'?

3.1 Perceiving Edges and Contrast

In 2000, PF described her visual experiences as follows:

> "It is sort of looking at the world with dirty glasses on. You can see shapes and placements of the objects about you but not determine the finer details" (posting by PF to The vOICe user group on 7th September 2000).

In 2002, PF described her visual experiences in terms of a non-detailed greyscale sketch. Her description was as follows:

> "I cannot tell fine little tiny details. Rather my vision is based upon black and white and all the little gradients in between. The best way I try to describe this to people is: take a large black sheet of paper; now take a magical piece of white chalk and sketch me here on this stage in a line drawing; now make me three dimensional and you've just about represented how my sight looks." (presentation by PF, April 2002).

In the context of a discussion about whether The vOICe gives 'just sound' relative to the prospects of neuroprostheses, PF is again clear that she has experiences of light:

> "Just sound?… No, it is by far more, it is sight! There IS true light perception generated by The vOICe. When I am not wearing The vOICe the light I perceive from a small slit in my left eye is a grey fog. When wearing The vOICe the image is light with all the little greys and blacks… The light generated is very white and clear then it erodes down the scale of color to the dark black. I don't really see a difference in this light as compared to the "light phosphenes" they are talking about." (email to Meijer from PF, 29th August 2002).

When asked about her level of detail 5 years later, in the context of recognizing faces, she again reports lacking fine detail:

> "Faces to me are still kind of blurry. I can tell eye sockets, brows. I can tell if they have long hair, curly hair; heavy set face. If it's a heavy set woman put up against a

> male I couldn't tell you which is which. It just isn't down to that point yet." (interview with PF, December 2007).

For PF, what she sees is not solely driven by The vOICe but is also guided by her knowledge of and interactions with the world. There are several references to this in the interviews that suggest that her experience with The vOICe has provided some basic auditory-induced visual experiences that have become strongly augmented by other senses (notably touch), by her knowledge of her current environment, and by her previous experience of seeing.

> "With the program I first experience what I call the 'sound smear image'. I detect something that's there, "ah there's something there" an edge, and I say to myself I bet I can find what that is. But the soundscape tells me 'there is something there'. Once I go up to that particular soundscape and touch it, I can equate myself with what I am hearing. This is where I get what I call the line drawings. I have an image of what a tree looks like. I have an image of what a car looks like. So it's like a line drawing – windows, stairs, doors – all these have line drawings in the beginning. It's just like trying to look at a sketch of them. " (presentation by PF, April 2002).

CC, like PF, also claims not to get enough visual detail to identify a person's sex or age with The vOICe saying "I could by their clothes, but you really don't get that much detail… I trained it down the arch [an arch at her place of work]. When people walk through the arch, and when cars were coming through, I could tell the size of the car and what people were roughly wearing – whether they had a long coat on, or shorts. You can tell stuff like that but you can't tell a lot." (interview with CC, November 2007).

When listening to a soundscape created from a static photograph of a monstrance, she describes her phenomenology as:

> "a bright band with, on either side, groups of small vertical objects. The last two on the right side of the picture are at slightly different levels… I have no idea of the scale of this picture… Zooming in, the circle does not have an even texture but has small differences in its colour or light. It certainly does not have an even edge." (email by CC to The vOICe user group, 30 January 2006).

When asked how the visual experiences from The vOICe interact with her real residual vision, CC describes it as normally being co-extensive with her residual vision (interview with CC, November 2007):

CC: You can still see blurred things ahead of you [from residual normal vision] but what will occupy your attention is the image of something else. If I looked ahead of me and pointed the camera to the side of me, what would occupy my attention, and what my brain would process the most, is the image [of the thing] at the side of me - which is a bit bizarre, so you have to point it ahead of you.

JW: If you look ahead at a dark object you would see the object through your eyes but you would also see something through the soundscape. But would both be in the same space?

CC: Yes it would. It would be in the same space.

JW: Would you think that the soundscape is helping you to see the true object or is it creating a separate image? Or both?

CC: You are seeing with two different sets of eyes, because you see the blurry sort of dark version with my eyes and a very crisp clear, but not as entirely detailed as other people see it, but very detailed version with the camera of the phone.

3.2 Perceiving Depth

Both CC and PF report being able to perceive depth. However, both of them report that this ability occurred after having ‘flat’ visual experiences of edges and shading. CC recalls one incident in which she emailed Peter Meijer (developer of The vOICe) to say "there appears to be a car driving up the side of a building!". CC, of course, realized that there was not really a car driving up a building, but that the road surface was perceived in an inappropriate vertical rather than horizontal perspective (i.e. a distant vanishing point). However, over time such distortions became less apparent. She recalls the car incident as occurring several months after starting to use The vOICe.

PF also described being able to ‘see’ things with The vOICe but without those things having any sense of depth. She reports her sense of depth perception arising in a kind of Eureka moment:

> “I’ve developed it where one day I was washing dishes and without thinking I grabbed the towel, washed my hands, and looked down into the sink to make sure that the water had got out and I realised “Oh! I can see down. I can see depth.” And I stepped back from that sink and looked down again to make sure that I wasn’t fooling myself , and I walked slowly through my house looking into the rooms and it was like incredible. I could see into the room. That flat drawing now has depth to it. I can sense it. I can’t tell you what is across the other end of the room but I can tell you that it is across from me.” (presentation by PF, April 2002).

This particular event can be dated to late 2000 (a posting by PF to The vOICe user group on 1st November 2000 described it as occurring "the other day"). She had been describing other visual phenomenology for at least 2 months prior to this (e.g. a posting by PF to The vOICe user group on 7th September 2000 quoted above).

3.3 Perceiving Movement

The vOICe software normally converts one visual image per second into a soundscape and is not well suited for detecting fast moving objects. It would perhaps not be expected to give rise to visual experiences of smooth movement (1 Hz is too slow for apparent motion, for example). However, both PF and CC no longer report any subjective experience of jerkiness nor is their experience a series of snapshots. As PF describes it:

> "It depends on the speed of the movement. If I am just walking, it will see the images as if you would be walking. I don't see the second break anymore, it's like totally gone. It's like using a flip-book, you don't see the different breaks between the images. My mind no longer does that, it just sees the whole image. It just redraws the difference of the image, there is no time lag. Recently, I tried using a bicycle with The vOICe and I was surprised that, although I could see where I was going, The vOICe did not track or update as fast as the wheels were moving under me. It was quite an experience to be on a bike after 30 years." (interview with PF, December 2007).

For CC, she reports that her experiences of smooth movement have come about gradually (interview with CC, November 2007):

CC: Well, at first when I did it, it was like a very jerky movie, especially with the phone version. You can alter the speed of its processing with the computer version. You can alter the speed of it. I found in some situations it was good to speed it up and in others to slow it down.

JW: Is it on a constant speed at the moment?

CC: I do because I have the mobile phone version that you can't alter really. I don't tend to use the computer version much.

JW: With the mobile phone version you experience it as smooth movement not jerky?

CC: I do now but at first I didn't.

3.4 Perceiving Color

The vOICe does not encode color information in the soundscape. As such, any subjective experiences of color must be driven either by previous knowledge of the color of objects, or synaesthetically related to the auditory signal. The only person who reports experiences of color is PF, but these experiences have emerged very slowly over time. When describing her experiences in April 2002 they were described as "black and white and all the little gradients in between". However, when interviewed in December 2007 she said that colors had started to emerge during late 2006.

PF: Now it has developed in to what I perceive as color.

JW: Really? Before you had said that it is not colored?

PF: Yes, that's true. But before my brain wasn't seeing the finer detail. Over time my brain seems to have developed, and pulled out everything it can from the soundscape and then used my memory to color everything.

JW: Aside from The vOICe, if you think of a strawberry you could still think of it in your mind's eye as being colored?

PF: Yes, red color with yellow seeds all around it and a green stalk. I can see it instantly.

JW: But if you look at someone's sweater or pants you wouldn't necessarily know the color? It could be blue or red.

PF: My brain would probably take a guess at that time. It would be greyish black. Something I know such as grass, tree bark, leaves, my mind just colors it in.

JW: How long ago was it when you started having the colors?

PF: Gradual, gradual but it is strongest now. Within the past year, year and a half, after my depth perception developed.

3.5 Role of Prior Visual Knowledge

PF's description of perceiving colors for known objects points to the involvement of prior visual knowledge in her phenomenology. Another clear example of this, is the level of detail with which she reports being able to 'see' objects and the distinction that she draws between objects she was very familiar with before becoming blind (e.g. trees) and objects she only became familiar with as a blind person (e.g. computers):

> "I took the program down to see my Christmas tree. I just wanted to see the pretty lights, that's all, but to my surprise I was able to see the branches swirling and twirling around. And then one thing caught my eye. That was that by looking at the branches, I could see the points of the needles and before I even could stick out my finger and touch them I could feel that it would be sharp as if it was saying "touch me and I will prick you". … as I've gone along later in life, life has changed and I have things in

> my house that I have no memories other than what I have learned from my hands what they look like. Things like a microwave. Things like my computer. Unfortunately even though I love a computer (it has given me so much) when I look at a computer all I again see is the line box drawing. It's not filled in solid like the tree. I cannot explain why. I just know that it is. I don't have a memory for what computers truly look like." (presentation by PF, April 2002).

3.6 Interim Summary

Both PF and CC report a rich range of visual experiences as a result of using The vOICe, and these experiences have evolved over time. But are they really seeing? The TMS study of PF carried out by Merabet et al. (2009) demonstrates that PF uses her 'visual' cortex for identifying vOICe soundscapes. This does not prove that she is seeing the soundscapes (although it is consistent with it). One counter suggestion is that the experiences are spatial but not visual. This is at odds with the reported phenomenology of both CC and PF in which light and dark is perceived (and color in PF) in addition to spatial characteristics such as shape. This conclusion hinges on the reliability of their reports. However, it is to be noted that both CC and PF have some natural light perception (i.e. they understand what light is) and both state that The vOICe enables experiences beyond the range of their limited residual vision. Another possibility is that the experiences reflect visual imagery rather than perception. It is certainly the case that their experiences are augmented by visual knowledge acquired before blindness. But this is equally true of normal vision in which prior visual knowledge plays an essential role. It is to be noted that their visual experiences are reported to occur involuntarily and, in this regard at least, they resemble perceiving more than imagining.

4. Acquired Synaesthesia for non-vOICe Stimuli

What is particularly interesting in the present context is that both PF and CC claim to be able to ‘see sounds’ also when not using The vOICe. Their brains have internalized The vOICe rules for mapping between hearing and vision and the rules are applied both when the device is worn and when it is not. Neither experienced this prior to using The vOICe. CC describes this as:

“Monochrome artificially induced synaesthesia, only in certain frequencies of sound. A small price to pay for very detailed vision, but the consultant's music next door sets me off as well (Bach Mass in B Minor)… The thing I experience is not in color, is in my mind's eye, and can be very distracting. The shapes are consistent and can be reproduced by the same sound. It is not triggered by all sounds but by vOICe-like sounds (the program, not people's voices). It is almost as if you had a computer with two monitors running simultaneously different pictures, one was a very grey blurred version of the real world, and the other was a pure grey background with a big semicircular light grey arc on it, and sometimes you switched your attention between both. The arc picture was triggered by the sound of a police car going by my office.” (email to Ward, November 2007).

PF also gives visual descriptions, all monochrome, to a number of simple sounds (e.g. sine wave tones, noises) including some sounds that fall below the frequency range of The vOICe (500 to 5000 Hz). Some representative examples are given below (from March 2008):

“high small flat white flashing disk” (pure tone of 2000 Hz, 500 ms duration).

“white circle on left circle is rotating. Reminds me of looking down or into a well or cave.” (pure tone of 150 Hz, 500 ms duration).

"clear hollow moving squares middle of face movement left to right." (noise between 250 and 300 Hz, 500 ms duration).

PF also believes that her synaesthetic experiences to these sounds are stable over time. In developmental synaesthesia, consistency over time is a hallmark feature,

> "Today I wanted to determine if I would have the same results from listening to the files… I opened the wav files one by one and then compared them to my descriptions I sent you. I am shocked [that] the descriptions for the files and what I see are almost the same. A silver disk in one file. The lines going down the left side of the hall. Each matching my previous description. How does this happen? I thought my mood, time, environment factors, etc., would change the images, color, motion effects. It does not." (email from PF, 26th March 2008).

For PF, but not CC, tactile experiences (e.g. of manipulating objects) also trigger visual experiences. However, there are important differences between her tactile induced vision and auditory induced vision. The sound of a car horn does not trigger a visual image of a car (or a visual image of a car horn) but the touch of a pencil triggers a visual image of a pencil. Whilst the latter could be considered as imagery rather than synaesthesia, it is also possible that both are true instances of synaesthesia that reflect different types of mapping between perceptual systems. In the case of touch, the visuo-tactile mapping may be based upon stored multi-modal representations of known object shapes, whereas the audio-visual mapping uses the learned features of The vOICe (pitch to height, loudness to brightness, etc.). The following is from the interview in December 2007:

JW: If you are not wearing The vOICe and you hear certain sounds could those also trigger vision? I am thinking of artificial noises such as a SHHH or a truck reversing or certain other sounds that aren't related to The vOICe?

PF: Yes, it does. Absolutely. Because my mind automatically records it as a visual sound. It has to be in a certain vOICe frequency. I understand that now. But you can't use a high car horn and it become a vision of a car. But if I hear a car horn, I see it in my mind through the 'vOICe sight'. I don't think of it like I use to 'see sight'. The vOICe sight, I call it The vOICe sight.

JW: But you would have The vOICe sight without using The vOICe for certain sounds?

PF: Yes.

JW: If you were to touch things you wouldn't get any visual experiences through that?

PF: Yes, I can! If I pick up a pencil, I feel pencil, I see pencil.

JW: Even if you touch it? If you touch an object with your hands you have an experience of seeing it?

PF: Yes! Touch is vision.

JW: If you weren't using The vOICe and you were to touch an object you would see that object?

PF: I'd see an image in my head, yes.

JW: You would feel this as occurring automatically rather than something you are deliberately creating?

PF: Yes, absolutely.

5. General Discussion

The aim of this paper was to present detailed descriptions of the experiences of two blind expert users of a sensory substitution system that converts visual images into auditory signals. There have been previous descriptions of blind users of a tactile sensory substitution system, but no known descriptions exist from users who have used such a device over a period of years in their day-to-day environment. As such, this is a unique and important record. Both blind users had normal or near-normal vision in their childhood and teenage years, and both users had experienced a prolonged length of time as being registered blind prior to using The vOICe. As such, it is inconceivable that their visual phenomenology from sounds arose as a spontaneous consequence to becoming blind per se even though this does occur in some cases of acquired synaesthesia (e.g. Rao, Nobre, Alexander, & Cowey, 2007). Moreover, although there was substantial time for cross-modal plasticity to occur after the onset of blindness (i.e. for the visual cortex to learn to process auditory and tactile stimuli) this cannot explain the emergence of visual phenomenology which occurred only after use of The vOICe. The crucial factor is probably that the sounds conveyed by The vOICe change in predictable ways when the wearer of the device moves his/her body and that it is possible to correlate certain aspects of the soundscapes with objects that can be touched and located via touch in external space. This is in broad agreement with sensorimotor accounts of visual consciousness (Hurley & Noe, 2003; O'Regan & Noe, 2001). There are certain properties of vOICe soundscapes that cannot be known via sensorimotor contingencies, such as the fact that bright objects are loud. However, both participants have some residual light perception that could facilitate learning of this rule and it is also a general multi-sensory association that they are likely to have acquired prior to becoming blind (Marks, 1974, 1982).

The experiences reported here are intended to be taken as 'proof of principle'. There are many things that we do not understand. Is blindness a necessary condition for such experiences or could sighted individuals, trained extensively, acquire them? Would congenitally blind users acquire an altered sense of visual-like space as a result of using the device, and would that include other visual qualities such as lightness? Are the particular mappings used by The vOICe special in some way, or could any consistent mapping between vision and audition lead to these kind of experiences? The latter is predicted by sensori-motor accounts provided the user has control of the camera.

It is noteworthy that, for both users, their phenomenology has not been stable but has developed over time. Initially, both report their visual experiences induced from The vOICe to be flat, jerky, lacking in detail and monochrome. The next major transitions are smooth movement perception and the perception of depth[2]. Both involve going beyond the information explicitly given in the soundscape, as interpreting depth in a photograph based on perspective, shading and occlusion. It is not readily apparent how movement of external objects can be perceived smoothly given the 1 s left-to-right scanning time. Perhaps the user-initiated movements of the device are used to achieve perceptual stability, and reduce the feeling of jerkiness. What is particularly striking about PF is that she now reports some of her soundscapes to be colored, some 5–6 years after wearing The vOICe on a daily basis.

One could certainly debate how much of the visual phenomenology is really attributable to The vOICe, and how much is attributable to visual imagery, visual memories, knowledge of

---

[2] It is worth noting that similar transitions are reported after removal of dense congenital cataracts in which adult patients see for the first time (Fine et al., 2003; Gregory & Wallace, 1963; Kurson, 2007).

Ward, J., & Meijer, P. (2009). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. doi:10.1016/j.concog.2009.10.006

the environment (e.g. knowing which things are where), and information gained from touch. All of these are clearly important. However, if the visual phenomenology of The vOICe was indeed largely dependent on top-down mechanisms and only loosely constrained and guided by the sensory input this would not be uninteresting given that many believe that this is an apt description of normal (non-vOICe) visual consciousness. What is genuinely interesting and novel is that the sensory constraints are coming from audition and not vision, even though the phenomenology is visual, and that they have been environmentally learned via interactions through a man-made device.

One example of how expertise in The vOICe can give rise to visual phenomenology in the absence of any obvious top-down or sensory-motor constraints, is that both PF and CC report 'seeing sounds' when the device is not worn. Neither participant reported having synaesthesia prior to using the device, or prior to becoming blind. Neither of them has more 'standard' types of synaesthesia unrelated to the device such as colored graphemes. Moreover, their synaesthesia involving sounds has somewhat different characteristics to equivalent developmental cases (e.g. Ward, Huckstep, & Tsakanikos, 2006). For vOICe users, but not developmental cases, most of the experiences are monochrome grey and they tend to be elicited only by certain sounds, i.e. those that resemble The vOICe. For example, the sound of human speech and the sound of most instrumental music do not elicit visual experiences in either CC or PF. The plastic changes in the brains of PF and CC that enabled them to learn to interpret The vOICe have long-term stability and these neural pathways are not disconnected when the camera that they wear gets disconnected. Acquired synaesthesia may well be an inevitable consequence of long-term adaptation to a sensory substitution device, and is unlikely to be limited to The vOICe. Acquired synaesthesia has been documented before as a spontaneous symptom of blindness (e.g. Rao et al., 2007) reflecting, presumably, cross-modal plasticity and reorganisation to compensate for sensory loss. However, the acquired synaesthesia in PF and CC has not arisen spontaneously but has come about through the sustained use of a technology.

Other types of synaesthesia (both acquired and developmental) do not generally fit well with sensorimotor accounts and others have even argued that synaesthesia provides a challenge to these models (Gray, 2003; Gray et al., 2002). For example, some normally sighted synaesthetes have visual experiences (plus auditory ones) whenever they listen to sounds (Goller, Otten, & Ward, 2009). In this example, hearing does not have the sensori-motor functions of seeing so it should not have the phenomenology of seeing (e.g. Gray, 2003). The same logic does not apply to the vOICe in which the soundscapes encode veridical visual information and parallel vision-like sensori-motor contingencies. The existing data therefore suggests that the question of why some sounds (or touch) can elicit visual experiences has multiple answers: either through genetic differences creating 'rewired' brains (in developmental synaesthesia); through spontaneous cross-modal plasticity (in most acquired synaesthesia); or through experience-dependent learning of sensorimotor contingencies (in sensory substitution) augmented by non-learned multi-sensory associations such as high pitch being associated with bright stimuli and spatial height; or a combination of the former. The mechanisms that let soundscapes give visual experiences in the blind may be further aided by reduced thresholds for having spontaneous visual experiences, as evidenced by the high incidence of Charles Bonnet Syndrome hallucinations after loss of sight, such that it could in part be a process of binding (to visually meaningful sound) rather than eliciting (from sound).

References

Ward, J., & Meijer, P. (2009). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. doi:10.1016/j.concog.2009.10.006

Afra, M., Funke, M., & Matsuo, F. (2009). Acquired auditory-visual synesthesia: A window to early cross-modal sensory interactions. *Psychology Research and Behavior Management, 2*, 31-37.

Amedi, A., Stern, W., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P., & Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience, 10*, 687-689.

Asher, J. E., Lamb, J. A., Brocklebank, D., Cazier, J. B., Maestrini, E., Addis, L., Sen, M., Baron-Cohen, S., & Monaco, A. P. (2009). A Whole-Genome Scan and Fine-Mapping Linkage Study of Auditory-Visual Synesthesia Reveals Evidence of Linkage to Chromosomes 2q24, 5q33, 6p12, and 12p12. *American Journal of Human Genetics, 84*(2), 279-285.

Auvray, M., Hanneton, S., & O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: Localisation and object recognition with 'The vOICe'. *Perception, 36*, 416-430.

Bach-y-Rita, P. (1972). *Brain Mechanisms in Sensory Substitution*. New York: Academic Press.

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature, 221*, 963.

Block, N. (2003). Tactile sensation via spatial perception. *Trends in Cognitive Neurosciences, 7*, 285-286.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., & MacLeod, D. I. A. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience, 6*(9), 915-916.

Goller, A. I., Otten, L. J., & Ward, J. (2009). Seeing sounds and hearing colors: An event-related potential study of auditory-visual synesthesia. *Journal of Cognitive Neuroscience, 21*, 1869-1881.

Gray, J. (2003). How are qualia coupled to functions? *Trends in Cognitive Sciences, 7*, 192-194.

Gray, J. A., Chopping, S., Nunn, J., Parslow, D., Gregory, L., Williams, S., Brammer, M. J., & Baron-Cohen, S. (2002). Implications of synaesthesia for functionalism. *Journal of Consciousness Studies, 9*, 5-31.

Gregory, R. L., & Wallace, J. G. (1963). *Recovery from early blindness: A case study*. London: Expperimental Psychology Society Monograph 2.

Guarnier.G. (1974). EXPERIENCE OF TACTILE VISION. *Perception, 3*(1), 101-104.

Hartman, A. M., & Hollister, L. E. (1963). Effect of mescaline, lysergic acid diethylamide and psilocybin on color perception. *Psychopharmacolgia, 4*, 441-451.

Hurley, S., & Noe, A. (2003). Neural plasticity and consciousness. *Biology and Philosophy, 18*, 131-168.

Keeley, B. L. (2002). Making sense of the senses: Individuating modalities in humans and other animals. *Journal of Philosophy, 99*, 5-28.

Kim, J.-K., & Zatorre, R. J. (2008). Generalized learning of visual-to-auditory substitution in sighted individuals. *Brain Research, 242*, 263-275.

Kupers, R., Fumal, A., de Noordhout, A. M., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proceedings of the National Academy of Science, USA, 103*, 13256-13260.

Ward, J., & Meijer, P. (2009). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. doi:10.1016/j.concog.2009.10.006

Kurson, R. (2007). *Crashing Through: A True Story of Risk, Adventure, and the Man Who Dared to See*. New York: Random House.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *American Journal of Psychology, 87*, 173-188.

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 177-193.

Meijer, P. B. L. (1992). An Experimental System for Auditory Image Representations. *IEEE Transactions on Biomedical Engineering, 39*, 112-121.

Merabet, L. B., Battelli, L., Obretenova, S., Maguire, S., Meijer, P., & Pascual-Leone, A. (2009). Functional recruitment of visual cortex for sound encoded object identification in the blind. *NeuroReport, 20*, 132-138.

Merabet, L. B., Maguire, D., Warde, A., Alterescu, K., Stickgold, R., & Pascual-Leone, A. (2004). Visual hallucinations during prolonged blindfolding in sighted subjects. *Journal of Neuro-opthalmology, 24*, 109-113.

O'Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*, 939–1031.

Poirier, C., De Volder, A. G., & Scheiber, C. (2007). What neuroimaging tells us about sensory substitution. *Neuroscience and Biobehavioral Reviews, 31*, 1064-1070.

Pollok, B., Schnitzler, I., Stoerig, P., Mierdorf, T., & Schnitzler, A. (2005). Image-to-sound conversion: Experience-induced plasticity in auditory cortex of blindfolded adults. *Experimental Brain Research, 167*, 287-291.

Proulx, M. J., & Stoerig, P. (2006). Seeing sounds and tingling tongues: Qualia in synesthesia and sensory substitution. *Anthropology and Philosophy, 7*, 135-151.

Ward, J., & Meijer, P. (2009). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. doi:10.1016/j.concog.2009.10.006

Proulx, M. J., Stoerig, P., Ludowig, E., & Knoll, I. (2008). Seeing "where" through the ears: Effects of learning by doing and long-term sensory deprivation on localization based on image to sound conversion. *PLoS One, 3*, e1840.

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.

Sampaio, E., Maris, S., & Bach-y-Rita, P. (2001). Brain plasticity: 'visual' acuity of blind persons via the tongue. *Brain Research, 908*, 204-207.

Ward, J. (2008). *The Frog who Croaked Blue: Synesthesia and the Mixing of the Senses*. London: Routledge.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex, 42*, 264-280.

Ward, J., & Mattingley, J. B. (2006). Synaesthesia: An overview of contemporary findings and controversies. *Cortex, 42*, 129-136.