Draft. Cite as: Deroy, O., & Auvray, M. (2014). A crossmodal perspective on sensory substitution. In S. Biggs, M. Matthen, & D. Stokes (Eds.), *Perception and Its Modalities* (pp. 327-349), Oxford: Oxford University Press.

A crossmodal perspective on sensory substitution

Ophelia Deroy[1] & Malika Auvray[2]

1. Centre for the Study of the Senses, Institute of Philosophy, University of London, UK

2. Institut Jean Nicod, CNRS UMR 8129, Ecole Normale Supérieure, Paris, France

1. Introduction

Blind people's ability to navigate through unfamiliar environments forces us to wonder how this is achieved. It also leads to philosophical questions, which Descartes already captured:

> No doubt you have had the experience of walking over rough ground without a light, and finding it necessary to use a stick in order to guide yourself. You may then have been able to notice that by means of this stick, you could feel the various objects situated around you, and that you could even tell whether there were trees or rocks, or sand, or water, or grass, or mud, or any other such thing. It is true that this kind of sensation is somewhat confused and obscure in those who do not have long practice with it. But consider it in those born blind, who have made use of it all their lives: with them, you will find it is so perfect and so exact, that one might almost say that they see with their hands, or that their stick is the organ of some sixth sense given to them in place of sight (Descartes, 1637/1985, p. 153).

The development of refined conversion devices designed to compensate for blindness certainly makes Descartes' puzzle all the more relevant and difficult to solve. These substitution systems, which convert visual stimuli into tactile or auditory stimuli, give access to distant objects in a way that was not possible with white sticks. Can these devices offer a genuine substitute for vision or do they compensate its loss through other means? The name chosen for these systems - "sensory substitution devices" (SSDs) - seems to have embraced the most ambitious solution considered by Descartes in his quote, i.e. that technical devices could enable blind people to see with some intact part of their anatomy. At first glance, sensory substitution devices certainly offer a striking functional equivalence with vision: they respond to electromagnetic stimuli, they enable blind people to navigate in new environments, to detect and identify remote objects and to judge their approximate distance. By bypassing the constraints of the mechanical transfer at stake in the white cane, SSDs convey information about the surrounding objects which is *richer* than the information previously accessed through the intact sense. What's more, they can do so at a *distance*, independently of an actual contact between the device and the object. Blind people now have access to the shape of out-of-reach objects, a privilege reserved so far to the sighted. Although training remains necessary and performance uneven, Descartes' first hypothesis seems to have been willingly embraced by most specialists in the field: SSD-users are said to "see with their skin" (White et al., 1970) or with their ears (see Deroy & Auvray, 2012, for a critical review).

Expressions such as "seeing with the skin" show how intuitive the assimilation of SSD-use to vision is. They also reveal an underlying tension: how could blind people ever see? Claiming that they see appears to be an oxymoron: if the definition of seeing involves proper visual processing, it is almost tautological to say that blind people, whose visual system is irreversibly damaged, cannot see. The initial idea that they "almost see" (Descartes, 1637/1985, p. 153) thanks to sensory aids might then have to be taken less literally: is not it more appropriate to say that they have gained or recovered something *like* vision, or close to it? It seems fair to us to say that the intuitive characterization of sensory substitution is more complex than the mere alternative between vision and not vision: What people seem to want to capture is closer to a "quasi-vision".

Two understanding of quasi-vision

Once one agrees to talk about something being quasi-visual, some difficulties arise: To what extent should we grant quasi-vision to trained users of SSDs? And more importantly: should we try to defend the idea that there is something like quasi-vision, and turn it into our best philosophical accounts of sensory substitution?

These, in our view, are the two main questions to be raised if we want to draw philosophical conclusions from SSDs. As SSDs superficially resemble vision, they lead to the intuitive idea that they are, if not visual, at least quasi-visual. This intuition, in turn, seems to have led to two distinct philosophical theories. On a first reading, quasi-vision has been taken to mean "same but less than vision". The claim, shared by what is known as the "deference thesis" (e.g., Hurley and Noë, 2003), is that SSDs give a sense of vision, but to a lesser degree than regular vision. On a second reading, quasi visual only means pseudo-visual (or deceptively visual). Theorists, defending the "dominance thesis", want to stress that SSDs remain in the substituting modality, e.g. touch or audition depending on the device used (e.g., Humphrey, 1992, Block, 2003, Prinz, 2006). What has lead authors to arrive at such strongly opposing views on the basis of what seems to be the same intuition? Each alternative needs to be considered and described in more detail.

The deference theory, according to which SSDs provide their users with a lower form of vision, is perhaps best expressed in Heil's comment that "a person making intelligent use of a TVSS (Tactile Vision Sensory Substitution) may be said to be seeing (though perhaps only dimly) features of his environment" (Heil, 1983, p. 16). Heil implies here that SSD-use differs from traditional, non-augmented sight in that it provides a diminished form of vision: In other words, this mean that SSD-users have a less exact or less powerful sense of vision. Although the claim is not always clear, it seems to concern the capacity itself, not just the performance: SSD-users are not compared to say, short-sighted individuals who are granted the same kind of sensory capacity or competence as others, but with a less good performance. They are considered as having a diminished competence altogether). The perceptual acuity of blind persons perceiving information through a visual-to-tactile SSD has been quantified using a standard Ophthalmological test, and their "visual" acuity averaged 40/860 (Sampaio, Maris & Bach-y-Rita, 2001); which would be enough for anyone refusing to even consider these users as seeing. The idea is then not to put them on the same scale as sighted people, but consider another parallel scale to measure their performance.

One thing that can be used to distinguish the scale used for traditional vision that of SSD-use is the kind of objects at stake: Trained users can indeed negotiate their way in simplified environments, with widely separated, very high-contrast objects (Collins, 1985; Jansson, 1983), but they cannot navigate successfully in real-world situations, such as environments cluttered with small and low-contrast objects. They are better at recognizing simplified drawings than ordinary objects, not to mention whole scenes (Kim & Zatorre, 2008). This issue discussed here is known as a bandwidth problem when visual information is converted into touch (see Loomis, Klatzky, & Giudice, 2012, for a review). However it is sometimes forgotten when it comes to discussing the status of SSDs. In addition, the bandwidth problem has not prevented optimists from claiming that SSD-users would be in the future on an equal footing with sighted people. They should qualify any straightforward assimilation of SSD-use to classic or full-blown vision, and at least encourage the defenders of the deference thesis to follow Heil's moderate formulation when he notes that "it is, to an extent, misleading to describe a TVSS-user as seeing in an unqualified sense" (Heil, 1983, p. 145)

Even once it is nuanced and takes into account the bandwidth problem, the thesis remains hard to pin down. What is the "qualified" sense in which blind SSD-users are supposed to see? In this instance some theorists seem prone to defend the idea that visual perception comes in degrees; in other words, that it is not an all-or-none state, but a scale of skills which progressively develop as mastery improves (Hurley & Noë, 2003; Noë, 2004; O'Regan, 2011). According to these theorists, individuals are not just more or less good at seeing; they also have more primitive or complex seeing skills. An appropriate analogy here might be to that of improving in Karate and updating the colour of the Karate belt with each level of skill: Going from a white belt to a black one does not just correspond to getting better at the same moves, it also involves expanding your move repertoire and being capable of new moves. In this way, it is possible to say that SSD-users, like holders of a karate white belt, generally have poorer seeing (or karate) skills than sighted individuals, while participating nonetheless in the same general practice. The analogy comes from the way we individuate actions, by reference to their goals and independently of variations in means. We count performing a simple or complex karate move as the same kind of practice, but we also count cutting with a knife or cutting with an ax as the same kind of action; as a consequence, the case should be the same for seeing with one's eyes or through a SSD.

Such interpretation of SSDs comes with compromises, because it abandons the classical definition of vision as a single skill and as a non-graded term. Is there a less compromising option? According to a second, more conservative interpretation, the superficial similarity between SSD-use and vision must not be taken seriously and SSD-use is only deceptively visual. Our first intuitions simply suggest that using a SSD can be *mistaken for* vision; that it looks like vision when it is not. "Seeing with the skin" or with the hand, as Descartes suggested, is considered to be only a confused way of speaking. As Leon wrote, saying that TVSS enables sight is not persuasive:

> "It is not more persuasive than the suggestion that we would hear sounds and various properties by means of the eyes, simply because we observe an optical transformation of an aural input by using, say, an oscilloscope" (Leon, 1988, p. 252).

This second interpretation is more often favored by supporters of an all-or-none approach to sensory perception, especially those who think that acquiring specific qualitative states is constitutive of vision. Using a SSD or a white cane is allegedly not like having a visual experience: it does not give characteristically visual sensations, like colors, the feeling of empty spaces or other modes of phenomenal presence of objects (see Auvray & Farina, forthcoming, for discussion).

There are different ways to extend the interpretation that SSDs are only deceptively visual, yet the most common one is to claim that SSD-users still exert their intact sense of touch or audition, albeit in a more expert way. Audition or touch is recruited to achieve new tasks, usually performed through vision, but they do not become visual for this reason.

A philosophical dispute

The initial intuitions about the similarity between SSD-use and vision turn out to be rather confuse. This partly explains why they have led to such opposite philosophical interpretations. The representatives of the enactive tradition (e.g., Hurley & Noë, 2003; O'Regan, 2011; Noë, 2004), along with remote sympathizers of the Gibsonian view of perception as a pick-up of information (such as Heil, 1983) claim that visually-impaired users of SSDs gain a primitive sense of sight through learning. This can be maintained, according to them, even if what is obtained in terms of compensation falls short of being perfectly similar to vision as we commonly understand it. By contrast, the pseudo-visual interpretation attracts most of the representationalists and those who put emphasis on the qualitative characters of experience. People like Block (2003) or Leon (1988), for instance, claim that the intact modality, where sensations are obtained, remains dominant in sensory substitution. It is clear that the main source of disagreement between these two interpretations comes down to the definition given to seeing. Is it sufficient to have an access to the same kind of information or to respond to the same kind of stimuli for an ability to be defined as being visual? Is it possible to see without using one's eyes, or without enjoying any of the classic visual phenomenology of sighted people? These questions converge toward the more general problem of the individuation of the senses which has recently benefited from a resurgence of interest (see Macpherson, 2011, for an overview). However, this renewed interest has not spread to the field of sensory substitution, or at least it has not led to an interest in updating their philosophical and empirical understanding. On the contrary, it has rather led to the idea that SSDs are "not paradigmatic" (Gray, 2011) enough to constitute proper tests to assess the distinction between the senses. As a result, they have shifted out of the scope of recent discussions. At best, following Nudds (2004), they seem to encourage the negative conclusion that SSDs mostly raise terminology problems. As suggested by Gray (2011, p. 255), "the presence of disagreement can itself be regarded as providing some support for the anti-realist view that there is no mind-independent fact of the matter about the case" (e.g., whether TVSS-users see or not).

The aim of our present chapter is to challenge this conclusion. SSDs, we claim, can contribute to the debates about the nature of the senses. To do so, the first thing to achieve is to bring the recent work about SSDs into the philosophical arena: Thinking of SSDs as more advanced white canes as they were in the 1970's, has been responsible, we believe, for the imperfect grasp of their quasi-visual status. Recent empirical studies, detailed in section 2, show how complex and deeply integrated the use of SSDs is. This makes it harder, but also more crucial, to determine whether or not they allow a form of genuine vision. The next step is to find better ways to address this difficulty. SSDs certainly provide good reasons to reject the dominant method of distinguishing the senses (e.g., Grice, 1962), that is through an ecumenical combination of personal and sub-personal criteria (sections 3 and 4); they also challenge single-criterion distinctions between the senses which have been attempted more recently (section 5). This is, however, not a reason to think of SSDs as undecided cases, or to embrace a form of anti-realism about the senses. A way out of the dilemma comes with a third way of understanding SSDs which does not try to reduce their use to any of the existing modalities. SSD-use leads to some novel ability, which does not however constitute a new sense properly speaking but a form of artificial crossmodal rewiring resting on pre-existing sensory capacities. The idea that SSDs are not to be aligned with canonical senses is detailed in section 6.

2. The integration of SSDs

SSDs were first developed in the late sixties to improve blind people's ability to navigate in their environment and to identify objects (e.g., Bach-y-Rita et al., 1969). The principle they rest on differs from the mechanical transfer at play in white canes: they start with a series of sensors, most often a camera, which respond to certain electromagnetic signals. These stimuli usually accessed through vision are converted into sensory cues that are detectable through an intact modality. Devices such as the TVSS (Tactile Vision Sensory Substitution), the TDU (Tongue Display Unit), or the Optacon (Optical to Tactile Converter) convert these light signals into tactile stimuli, that is electromechanical or vibratory stimulation of different intensities and spatial distribution applied to the skin of a part of the body surface. In this case, the conversion from visual to tactile information is analogous. Other types of devices like the vOICe (Meijer, 1992) or the Vibe (Hanneton et al., 2010) capture the shapes of objects and their locations through a camera and convert this information into auditory signals. The conversion code translates several dimensions of the original signal into auditory dimensions, such as visual brightness into auditory loudness.

At first, and as was noted for instance by Leon (1988) and Ross (2001), SSD-use seems to be a form of indirect perception, analogous to the case where one indirectly sees the temperature of the stove through directly seeing its redness:

> "It may seem that use of a TVSS is a kind of vision without color. But while use of TVSS detects properties usually detected by vision - namely, spatial properties at a distance - and while its use can provide such information without reflective inference after a period of training, still its use is not a kind of vision because it is not a kind of direct perception. Rather, tactual properties are used as a basis to infer spatial properties" (Ross, 2001, p. 501).

The same interpretation could be applied to users of the vOICe *indirectly* accessing information about shapes through *directly* hearing certain sounds. However, this analogy proves to be inadequate. Many things have been encompassed under the heading of indirect perception. For the present purposes, and in line with what Ross means by the term, indirect perception it is defined as a two-steps form of perception. It is supposed to be distinct from an inferential process where a first perceptual stage leads to a judgment. If one first *sees* the redness of the stove and then infers that the stove is hot, this does not qualify as indirect perception. One has to (indirectly) see the hotness of the stove *in* its color. Both the directly and indirectly accessed information is somewhat concurrently present in experience: one can see both color and warmth. For SSDs to be interpreted as indirect perception, users would have to be able to perceive both the sounds and the shapes in the sounds. However, there is no scientific evidence to corroborate that both sounds and shapes are perceived at the same time. Before training, individuals using a visual-to-auditory device like the vOICe merely indicate perceiving noises through headphones. They then become able, through practice, to perceive shapes through the device and, when they do so, they are no longer paying attention to the sounds they first heard. By contrast, people who indirectly see the temperature of the stove in its color are still paying attention to the color - and perceiving it as such - independently of perceiving its temperature.

Similarly, with visual-to-tactile conversion systems, trained users no longer feel the tactile stimuli on their skin. In a much quoted report, Guarniero, a trained user of the TVSS (who was also a graduate in philosophy), explains:

> "Very soon after I had learned how to scan, the sensations no longer felt as if they were on my back, and I became *less and less* aware that vibrating pins were making contact with my skin. By this time objects had come to have a top and a bottom; a right side and a left; but no depth - they existed in an ordered two dimensional space" (Guarniero, 1974, p. 104 - quoted by Leon, 2011, p. 165; Heil, 1983; 2003, p. 228, 2011, p. 288; Peacocke, 1983, p. 15).

The main problem now comes from capturing the exact way in which training allows users to move beyond the experience of stimulation to the location where it is transmitted (such as the tongue or the ears) and enables them to to gain relevant information about the object-source. Recent studies have illuminated this transition, dubbed "integration". We think it is necessary to distinguish carefully three aspects of integration often associated in the literature. Unlike cochlear implants for instance, SSDs are worn intermittently and are not therefore fully anatomically integrated. The term "integration" is rather meant to capture the effects of familiarization with the device. It encompasses three different aspects that will subsequently be described in more detail: distal attribution, direct acquaintance and generalization.

Distal attribution

First, training with SSDs has been reported to result in what is called "distal attribution", and what some philosophers would be happy to call "intentionality": the received information is taken to be *about* an object which appears to be independent of the perceiver[1]. The expression, as much as the correlated philosophical notion, has to be handled carefully. Distal attribution is often considered to be synonymous with *spatial projection.* It is thought to occur when the sensations felt on the skin or in the ears are projected onto a distant perceived object, which should correspond to the one captured by the camera. Being perceived as distant in space is nonetheless not necessarily conceptually equivalent to being perceived as mind-independent (e.g. objective, see the debates around Strawson, 1959, for instance in Evans, 1980 and Burge, 2010). First, for certain SSDs, the sensor needs to be in contact with the object to be perceived, so that distant objects cannot be accessed. The Optacon, for instance, and its predecessor, the Optophone, functions more like a scan, as pages of written text need to be put in close contact with the sensors before being converted into tactile Braille stimuli. Second, SSDs could be designed to compensate for the loss of a contact sense. This has actually been the case for instance with SSDs compensating for the loss of touch (e.g., see Bach-y-Rita et al., 2003) or pain (Brand & Yancey, 1993), which do not give access to remote objects.

Leaving the variety of current and possible devices aside, finally, what seems to be crucial for the most well-known visual-to-tactile and visual-to-auditory SSDs is not so much the projection in distant space, but the transition to an allocentric representation of space. In other words locations are not just represented with respect to users' perspective, as it is classically assumed to be the case for touch, but within an external frame of reference which is independent of their actual position (see also Arnold et al., forthcoming for a discussion). As Guarniero (1974) reported, the objects seem to the user to have “a left and a right”, and not just to be felt relative to her own left and right. What distal attribution brings is first and foremost the perception of objects as being independent from the observer.

Direct access and transparency

Integration also involves a second transition: users become “directly aware of the distal object” (Siegle & Warren, 2010, p. 209). What this means is that the information about the independent objects appears to be delivered immediately, e.g. literally not through the mediation of a sensational episode. This aspect is mainly documented in subjective reports and remains hard to interpret further at this stage. Most often, sensations in the intact sensory modality are said to somehow be merged into a new percept, from which they become inseparable. People perceive the shapes of objects, not mere sounds, when they use the vOICe.

The integrated use of SSDs may then count as a form of *transparent* access: perceiving the object is no longer dissociable from what it is like to use the device, in the same way as perceiving blue is not dissociable from what it is like to see blue (Tye, 2002). There is however an alternative interpretation, under-estimated in our sense, according to which sensations remain present, but are less attended to (see for instance Noë, 2004).

[1] Although note that there are currently no unproblematic experiments that provide evidence of distal attribution with SSDs with methods other than explicit inference (see Hartcher O’Brien & Auvray, submitted, for a review).

In summary, immediacy, directness and transparency are hard to define (see Smith, 2002, for a discussion). They are also difficult to separate and assess empirically (Auvray et al., 2005, 2007-b; Siegle & Warren, 2010).

Generalization to new objects

The third aspect of integration comes from what is called “generalization”. With training, users are able to recognize objects or to navigate in environments which differ from the ones they have been trained with. This crucially demonstrates that their learning can be transferred to unfamiliar situations and therefore does not just reduce to memorized routines about familiar cases (Auvray et al., 2007-a; Kim & Zatorre, 2008). For instance, having learnt to recognize a plant during their training, users of the vOICe are subsequently quite fast in learning to identify another plant. In addition, with tactile devices, once trained with the tactile matrix on a body surface, performance transfers to another body surface when the matrix is displaced (Arnold & Auvray, submitted). Although generalization to new objects remains limited and deserves further study, it rules out the hypothesis that SSD-users proceed only on the basis of associative memory: users do not only learn to associate a specific pattern of sensory stimuli (a certain sound, a certain pattern of tactile vibrations) to a specific object or kind of object. Identification goes beyond the memorized associations and, as we just said, users are able to identify another pattern of sensory stimuli as being a new object of the same kind.

Here the data need to be very carefully interpreted: background knowledge certainly plays a role in SSD-use, as it does in everyday navigation or identification. A study by Siegel and Warren (2008) showed that, when using a minimalist visual-to-tactile device, people who are told to attend to proximal stimulation have worse performance than people who are told to directly attend to the distal source of stimulation. Having active training noticeably helps with the generalization of learning. But, as the evidence stands, nothing rules out that conceptual knowledge and explicit rules do not provide users with other advantages, or good-enough substitutes. How much knowledge *that* and how much knowledge *how* – or practical mastery – enter in the integration of SSDs is certainly a hot topic for discussion; and above all, investigation. Anyhow, the fact that there is a form of spontaneous generalization, and the fact that there is a transfer of learning can be taken as similar to perceptual learning, where for instance people do not need training to recognize new textures or new shades of color.

As we can see, the evidence gained from the integrated use of SSDs encourages the intuitions that SSD-use is quasi-visual: Gaining information in an immediate way, about new independent and possibly distant objects is almost like seeing. But is the status of SSDs thereby clarified?

3. In which sensory modality do SSDs get integrated?

The contention remains between the "same but less than vision" theories (also known as deference thesis) and the "not-vision" theories (usually associated to a dominance thesis). Overcoming this ambiguity requires explaining how and where to draw the line between seeing, touching and hearing. As Nudds (2011) rightly wrote, the individuation of the senses is itself an intuitive matter, calling for further clarification:

> "The distinction we make between the five senses is universal. Instead of saying in a generic way that we perceive something, we're talking about a perception in one of these well defined five modes: we see, hear, touch, smell and taste things. But what exactly is our distinction, when we identify these specific ways of perceiving things? What is, in other words, a sensory modality?" (Nudds, 2011, p. 311).

The vagaries surrounding the status of SSDs may reflect the vagaries of our concept of a sense. Instead of discouragement, this brings some hope that clarifying our theorizing about SSDs will shed some light on the distinction between the senses.

Four criteria to distinguish the senses

Among the various attempts to analyze our intuitive distinction between the senses, the most famous one came from Grice. In his 1962 paper, he proposed a list of the main criteria along which this distinction could be drawn. They can synthetically be specified as ranging over characteristic phenomenology, intentional objects, kinds of stimuli and dedicated processing.

The most immediate distinction between our senses comes from experience. Seeing usually does not feel like hearing or touching. Each sense corresponds to a specific kind or family of feelings; what Grice called "special introspectible characters".

> (Phenomenological criterion). Two senses, S1 and S2, are different as far as they give rise to experiences with different kinds of subjective feelings or conscious qualitative characters.

This criterion is reminiscent of Descartes' (1637/1985) remarks: what makes us think that blind people have a form of vision when using their canes, results from an analogy with our own phenomenology or experience. *What it is like* for us to be moving in a room with a stick could be confused for *what it feels like to see*[2], and therefore, the same might be true of blind persons.

The first comparison needs of course to be made with caution: what does it mean to say that touching a stone with a stick is subjectively similar to seeing the stone? Here though, it might only mean that one recognizes the object explored with the cane, and when identifying it as a stone through her background knowledge, recalls how a stone looks like. The phenomenological criterion is not supposed to apply to this form of mental imagery: If the experience in a sensory modality S1 is accompanied by a vivid form of mental imagery pertaining to another modality, S2, this should not lead to say that S1 is the same as S2[3].The criterion is supposed to apply to the initial or proper phenomenal modes of presentations, or kinds of experiences.

The difficulty to apply this criterion is illustrated here by the difficulty of distinguishing between "feeling like seeing" and "feeling like one is acquainted with an object (strongly) associated to a visual image".

[2] Certainly, in our case, because of the strategic or automatic recruitment of visual imagery (see Spence & Deroy, 2013 for discussion).

[3] See for instance, what happens with synaesthesia (Auvray & Deroy, in press; Keeley, in press).

This first criterion is strongly tied up with a second one, spelled out in terms of the objects or properties which are perceived.

> (Object criterion). Two senses, S1 and S2, are different as far as they give access to different characteristic kinds of objects or properties.

This second criterion is very largely reminiscent of the old Aristotelian way of drawing the distinction between the senses in terms of "proper objects". For instance, vision gives access to colors whereas audition gives access to sounds. It is important to note here that the distinction remains a psychological distinction: proper objects are intentional objects that do not necessarily correspond (or at least not exactly) to physical entities. Saying that vision's proper objects are, for instance, colors and visual shapes, does not presume of what colors and shapes are, metaphysically or physically speaking. This said, it is obvious that this criterion considers sensory modalities to be a way of accessing information: the question is what kind of information is *presented* to us, and how.

This way of conceptualizing the senses leaves aside the fact that senses are also organs and physiological circuits; e.g., they correspond to ways in which this information is *delivered*.

Two additional criteria are needed to capture this aspect. The first one appeals to a difference in terms of range of responses. Different senses are sensitive to different changes in the environment:

> (Stimulation criterion) Two senses, S1 and S2, are different as far as they respond to different kinds of physical stimuli.

Furthermore, the different stimuli serve as inputs to different and independent kinds of processing, which can be specified in computational and neurological terms. This criterion is sometimes also put in terms of sense-organs or kinds of receptors.

> (Processing criterion) Two senses, S1 and S2, are different, as far as they correspond to independent processing and neurological channels, or to different sense-organs.

A perceptual episode, for example, is visual if it is initiated by the stimulation of the retina by light-waves, which results in the activation of the optic nerve and brain areas V1 to V5 in the occipital lobe. Another episode is auditory if it begins with the vibration of the eardrum, generated by acoustic waves, and extends through the activation of the auditory nerve and another part of the cortex, the temporal lobe.

Adapting Grice's distinction

Our definitions of the four criteria depart from Grice's in two important aspects. First, the definitions are all formulated in "as far as" conditions. This is crucial to enable a graded application, needed to account for "quasi" cases like SSDs. Absolute criteria, such as those used by Grice, mean that each condition needs to be perfectly met before categorizing something as a certain sense. In other words, the phenomenology, object, kinds of stimulation and/or kinds processing in SSDs must be *perfectly* identical to what they are in sighted people, for them to count as seeing. However, this requirement can be criticized as being too "anthropocentric", the objection being that other animals might be granted a sense of vision in the absence of such perfect identity. Dogs have a sense of vision, although their experiences and neurological wiring might only be partially like ours. Likewise, we argue, it is also detrimental to the various realizations of vision (and other modalities) in human perceivers.

A second difference with Grice's view comes from the relation between the criteria. Grice's agenda was to show that the four criteria are all closely connected, although some have more importance than others (i.e., intentional objects and characteristic experiences). Here, we want to remain neutral about their connection or independence. Our concern is indeed to see whether or not they help clarify the status of SSDs, be it in combination or in isolation.

4. Why SSDs challenge the ecumenical conception of the senses

As was stressed by Grice, the common-sense distinction between the five senses mixes all four criteria. It encompasses broad ideas about access as well as delivery. Vision for instance is taken to involve a certain kind of experience, to be about a certain range of objects and properties, to be obtained through the visual system and to start with specific stimuli. Problems arise when one wants to make a more precise articulation of these different criteria.

A precise combination of the four criteria creates immediate problems for SSDs, as they do for other cases (for instance, in animal perception, see Matthen, 2007). For example, a device like the vOICe comes out being both visual and non-visual depending on which criterion is applied. *Phenomenologically* speaking, the experiences are not canonically auditory, but they are not canonically visual either. Yet the *objects* that are perceived definitively resemble visual objects. In terms of *stimuli*, both auditory and visual ones seem to be involved in the functioning of the headphones and the camera respectively. Yet the sense-organ (i.e., *processing* criterion) which is used is certainly not visual. This at least explains that the ambivalent intuitions about the visual status of SSDs come from the contradictory results of these criteria. In terms of access (e.g. how it feels to access such and such proper object), a device like the vOICe can be characterized as neither visual nor auditory, but might still emerge as strongly visual if one considers the object as a visual one. In terms of delivery (e.g. what kind of stimuli and processing are involved), the vOICe counts both as visual and auditory.

If one wants to maintain that four criteria are needed to identify a sensory modality, it seems necessary to at least accept that not all of them are equally important. But which one of the four should prevail in the definition of a sense? Several factors might encourage philosophers to privilege the criteria of intentional objects and stimuli. These two criteria are the safer ones because they do not rely on subjective reports (like the phenomenological criterion) or on what can seem still to be unsettled or at least fast-changing neurophysiological accounts (like the processing criterion). Moreover, widespread representationalist tendencies support the criterion of intentional objects over the other criteria, whereas the externalist sympathies (also frequent) tend to stress the criterion of environmental stimuli. Now, these two criteria will make SSD-use come out as visual; which is often what is defended in the literature. A way to acknowledge that the other criteria do not go in the same direction, and to put them further down the list, is add some nuance to the philosophical assimilation of SSD-use to vision, and talk about almost seeing or other quasi-visual cases as we discussed above (e.g. Heil, 1983).

We believe that even this nuanced assimilation of SSD-use to vision is still biased. The privilege given to the criteria of intentional objects and environmental stimuli is not appropriate when dealing with SSDs. In systems such as the vOICe or the TVSS, the human-machine interface is designed to provide access to "visual objects" (e.g., objects and properties which are equivalent to objects and properties normally accessed by the sighted). The interface is also designed to start from inputs that are traditionally visual. Prioritizing the criteria of intentional objects and stimuli offers a trivial reformulation of the initial goal: SSDs are designed to function as vision, therefore they are visual. This reformulation does not make a distinctive contribution in defending the real visual status of SSDs. The problem is not “what are SSDs used *for?*” but rather “*how* are they used?”. It is a descriptive, and not a design or functional, question. The former question remains unanswered.

5. SSDs challenges to single criterion approaches

If the combination of criteria raises problems, adopting a single criterion could lead to better results. Each of the four Gricean criteria has been thought, in turn, to be sufficient to draw the line between the senses; be it phenomenology (Leon, 1988; Lopes, 2000; O’Dea, 2011), stimulation (Gibson, 1966; see also Heil, 1983, 2011), objects (Roxbee Cox, 1970; Everson, 1999), or processing (see for instance Milner & Goodale, 2008, for a discussion in the domain of vision). Do single criteria perform better in clarifying the quasi-visual status of SSDs?

The phenomenology criterion

The application of the first criterion introduces a methodological problem: this criterion cannot be investigated from a third person perspective. It is impossible to specify what the experience of trained blind people is really like, for instance what a user of the vOICe really feels. We can only apprehend the associated phenomenology by using the device ourselves. However, even if we do so, there is no certainty that our experiences would be qualitatively similar, or even comparable, to those of other users, and especially to those of blind users. Taken at face value, the most recent studies tend to stress the inter- and intra-individual variations in the experiences enjoyed by trained users of SSDs.

One standard way of investigating other's conscious experiences, is to ask them what these experiences resemble. This methodology was adopted in a series of interviews in which two blind users of the vOICe surprisingly reported visual experiences, involving colors, but this phenomenology was limited to certain kinds of objects (Ward & Meijer, 2010). One of them for instance describes:

> "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 realized Oh! I can see down. I can see depth."

Later on, she claims:

> "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: 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 grayish black. Something I know such as grass, tree bark, leaves, my mind just colors it in" (from Ward & Meijer, 2010, p. 496).

It is very difficult to know what to infer from these reports. The origin of the visual, color imagery is crucial for the application of the phenomenological criterion. It is important to know whether or not the visual phenomenology is truly perceptual, or just given by imagination or memory, and then simply added to a more confused set of non-visual qualities obtained when using the SSD (see Auvray & Farina, forthcoming, for a discussion). Another related thing to be checked is what kind of visual imagery exactly is at stake. Blind people may not be the best ones qualified to report on this point, and mere reminiscence of crossmodal imagery might be at stake (e.g. Spence & Deroy, 2013). In a similar vein, recent studies have found a correlation between activation of the occipital cortex and the occurrence of new phenomenological experiences of phosphene in trained (late-blind) users of visual-to-tactile SSDs (Ortiz et al., 2011). Such novel phosphene experiences are not documented in congenitally blind individuals.

In their study, Auvray et al. (2007-a) asked sighted participants what it felt like to perceive with the vOICe. The replies varied importantly. While some of the participants claimed that their experience was close to visual experiences, others claimed it was closer to auditory ones. The hypothesis that the phenomenology does not come with a clear modal signature, and is neither exactly visual nor exactly auditory, can be confirmed by the fact that some thought that their experience was best described as tactile or olfactory. Some even reported that their experiences felt more like a sonar sense, suggesting that they were trying to make sense of something quite new. Furthermore, the phenomenology differed depending on the task that was performed: most of the participants gave different descriptions of their qualitative experience for localization tasks and recognition tasks. One of the participants for instance felt that her experience was visual when she was locating an object in space, but auditory when she was recognizing the shape of the object.

The lesson of this survey is twofold: From a methodological point of view, it questions users' ability to establish classes of similarities between their experiences, and thus to apply the phenomenological criterion in a rigorous or consistent way. For the present purpose, it shows that the phenomenological criterion is not sufficient to determine whether SSD-users see or not. The phenomenological criterion is by itself not stable or accurate enough to go beyond the intuition that using a SSD is quasi-visual.

The object criterion

The object criterion proves easier to assess. Trained SSD-users are able to navigate in new environments, and to make coarse but still accurate, judgments about shapes and distances of remote objects. They rely on something that audition or touch cannot usually provide, and have acquired something closer to vision. But the distinction in terms of kinds of objects does not prove more helpful when it comes to moving beyond this general characterization.

There are indeed limits to the application of the object criterion in isolation. Distance is accessed both by sight, audition and other distal senses. At best, therefore, if SSD-users are able to make distal attribution, they can be said to exercise a distal sense; whether it is visual or not remains open. Shape is not the *proper object* of sight either, as it can be accessed through touch as well. None of the above-mentioned aspects of the object criterion is therefore sufficient to declare SSD-use as being characteristically visual. A more promising suggestion comes from combining the two kinds of information, given that the joint access to distance and shapes seems unique to vision. Audition provides information about object distances, but gives only limited and approximate information about an object's size (and even less information about shape); touch provides information about shapes, but not at a distance. But in SSDs, the question remains open as to whether the two kinds of information can be accessed jointly or, as suggested by previous experiments (Auvray et al. 2007-a), they are distributed in two distinct perceptual tasks (localisation and identification).

A second limit is that certain key visual properties are lacking in SSD-users' experiences. Their perception is limited to a smaller number of dimensions. Colors, for example, are recognized as being specific to vision, yet they are not accessed through the vOICe or the TVSS. Levels of grey are not even present in the TVSS. There is no reason to see this objection as decisive. The absence of color is an actual technological limit and should not be taken for a principled impossibility. Current systems could be made to code for color: the relevant sensors are available and the conversion can be performed. SSDs could theoretically give access to all the contents which are specific to vision. The main problem comes from the fact that the device blurs object dimensions. Each dimension in the lacking modality needs to be translated into another dimension in the intact modality, and in this instance it appears a real impossibility to map all visual dimensions simultaneously in modalities which have less dimensions such as touch (see Loomis & Klatzky, 2012, for a review).

The object criterion remains compatible with the two-pronged nature of this dilemma. It is true that SSDs give access to a lesser number of properties than regular vision. This may mean that they are on a continuum with vision, or that they remain in the intact modalities, which share access to these properties. Taken in isolation, the object criterion is finally no more conclusive than the phenomenological one.

The stimulation criterion: starting with electromagnetic waves

The criterion of stimuli provides a clear answer, if one agrees to consider the user and the device as a single system: the stimulus to which the overall system responds to is the same as vision, since the sensors detect electromagnetic differences. The difficulty in applying this criterion comes this time from the way to accommodate the non-visual, intermediate step. What counts as "stimuli" could be either the light-waves received by the sensor or the auditory or tactile stimuli that get converted to by the device.

Note here that the stimulation criterion is distinct from the phenomenological one, about whether users feel auditory *sensations* when using the vOICe. The question concerns the role of the auditory *stimuli*, which undeniably occur. On the one hand, these auditory stimuli only exist because there are based on visual stimuli: both their existence and their nature depend on electromagnetic waves. This strict dependence on visual stimuli makes them somehow dispensable in the understanding of the input-output process. To take an extreme analogy, one does not count the chemical transduction of information through sensory channels as a sign that chemical stimuli are relevant for the functioning of every sense. On the other hand, auditory stimuli are certainly *necessary* for the proper functioning of the device. That is, they are constitutive of the way in which these SSDs have been designed. Switching from visual-to-auditory devices to visual-to-tactile devices for instance requires a completely new design, a unique translation code from vision to touch and a different training. The distinction between the senses in terms of stimuli does not deliver a single answer to the dilemma.

The processing criterion

The answer given by the strict application of the neurophysiological criterion is no less problematic. The distinction between the senses has sometimes been thought to be only a matter of sense organs. Following this line of argument, Morgan (1977) pointed out the similarity between the visual system and the TVSS: in both cases, an image is formed on a bi-dimensional surface (the retina in the case of vision, the lens of the camera for the TVSS), this surface contains discrete elements (rods and cones in the eye, vibrators for the TVSS), these elements are connected to certain regions of the surface (receptive fields) that send electrical signals to the brain, the device (eye or camera). The sensor (eye or camera) can be moved, and its movements introduce changes in the image. Moreover, these two systems are similar in that the source of stimulation is remote and sensing is subject to the effects of occlusion, when an obstacle comes between the source object and the detection system. But are these structural similarities sufficient to conclude that the two are one and the same sense? The animal kingdom provides a perfect illustration of the difficulty: it is commonly said that dogs “see”. These creatures are equipped with detection systems which share some similarities with human vision. It is more parsimonious to say that perceiving through these organs, or devices, is distal and discrete. Human vision, hearing and many other animal senses can be grouped into this category, thereby denying the possibility to decide over the visual status of SSDs on such a criterion.

Can further neurological investigations help? Studies have shown increased activation in visual areas in trained users of visual-to-auditory (De Volder et al., 1999; Renier et al., 2005) and visual-to-tactile conversion devices (Ptito et al., 2005). There is thus a recruitment of visual areas through brain plasticity (see Poirier et al., 2007, for a review). Applying the processing criterion suggests that blind users acquire a form of vision: if perception recruits a channel identified as visual, it counts as visual.

We must however be cautious with this sort of inference: the studies which motivate the claim that the occipital lobe is visual have been performed on sighted, and not blind, people. Interestingly, Ptito et al.’s (2005) study revealed increased activation in the visual cortex only in trained congenitally blind people, but neither in sighted ones reaching the same level of performance, nor in untrained people from both groups. We must also not forget that calling this area visual results from a functional mapping and that a common area may not be functionally similar in two groups of people, noticeably blind and visually impaired people.

Increased activity in a given brain area is not sufficient to deduce the presence of a normally associated function. In accordance with this conclusion, Kupers et al.

(2006) used transcranial magnetic stimulation (TMS) in both blindfolded sighted and blind participants' visual cortex before and after training with a visual-to-tactile conversion system (the Tongue Display Unit). When TMS was applied over the visual cortex before training, none of the participants reported any subjective tactile sensation. However, after training, when TMS was applied over the same brain area, some of the blind participants (3 out of 8 early blind and 1 out of 5 late blind) reported somatopically organized tactile sensations that were referred to the tongue, whereas no such sensations were reported by sighted participants. The authors concluded from their data that the subjective experience associated with increased activity in the visual cortex after practice is tactile and not visual (see also Ptito et al., 2008, for similar results with Braille reading).

The application of the sole processing criterion does not deliver a definite answer. At best what can be concluded is that using a SSD is distal (when users move in their environment), which we knew already from the study of SSD-integration; and that it requires some further perceptual processing, which can recruit un-solicited areas like V1 in blind people. Should we conclude that calling SSD information visual or not is just arbitrary? Or, as suggested by Nudds (2004), shall we call it a matter of pragmatic decision which offers no fact of the matter? These two lines, we suggest, can be resisted.

6. Shifting the problem

It is impossible to clearly answer the question of whether blind people "see" or "hear" when using the vOICe through any of the usual accounts, whether they rely on a combination of criteria or a single privileged one. As we reviewed these criteria, it nonetheless became clear that the integration of such devices has noticeable effects on users.

The mistake comes, we claim, from not acknowledging this difference and trying to square SSD-use with one of the existing sensory modalities. The application of the criteria should thus lead to a more qualified re-description of the effects of SSDs integration:

> (Phenomenological criterion). Using a SSD leads to subjective impressions that cannot easily be compared with experiences in existing modalities.
>
> (Object criterion). A SSD can give access to a variety of objects, constrained by the initial design and the number of dimensions in the stimulated modality.
>
> (Stimulation criterion). A SSD uses a series of inter-dependent stimuli.
>
> (Processing criterion). Integrating a SSD requires some supplementary neurological processing.

Altogether, these criteria lead to the conclusion that SSDs introduce something *new*, which, albeit intimately connected to the intact modality, does not reduce to it. It also strongly *depends on* the missing modality – which governs the design, e.g. the selection of stimuli and kinds of conversion. What appears is an experience which is significantly different from both, while it can also be reminiscent of each modality, at times. Several recent accounts point toward the "novelty thesis", and help climb above the dilemma and pragmatic renunciations which have been discussed above. As detailed in Auvray and Myin (2009), SSDs are better compared to mind-enhancing tools than to perception. Following Clark's (2003) lines, such tools provide means to carry out cognitive functions in ways that would have been impossible without them, given the intrinsic or initial properties of the system. Accordingly, SSDs provide cognitive extensions to the existing senses, which might then be integrated to these senses.

Starting from a different perspective, Nagel et al. (2005) designed a SSD for entirely new intentional objects. They equipped participants with a "magnetic belt" which gave them some brand new information about changes in the magnetic field. With training, users were able to track their orientation relative to the cardinal points. The authors' main point was to establish that trained users had acquired a new "modality", e.g. a new sensorimotor skill. They nonetheless suggested that their results form a transformation of the existing spatial perception:

> "Strictly speaking, the changes in perception indicate not a genuinely new modality, but modification of the meta-modality of spatial perception. The term 'metamodality' is used to reflect the fact that normal spatial perception is fueled by visual, auditory and somatosensory information. The ability to infer spatial information from these pooled 'primary' modalities may have thwarted our objective to create a completely new sensory modality. Instead, the acquired sensorimotor contingencies lead to a transformation of this already existing meta-modality" (Nagel et al., 2005, p. 23).

The most noticeable aspects of SSDs, according to this account, come from their non-reducibility to the existing senses, and their dependence on them. As suggested by Nagel et al. (2005), the successful use of a SSD would not have been possible without the existing resources, in this case in terms of intentional objects (spatial orientation) and stimuli (tactile vibration). Their non-reducibility and dependence on existing senses are constitutive features of SSDs.

This leads us to refine the idea of a novel ability. The best way to capture this novelty is to think about the integrated use of a SSD being something closer to a new set of automatic recognition abilities arising from a trained rewiring between other sensory modalities, and eventually other pre-existing capacities. Unlike other senses, whose developments occur in a parallel way, SSDs are modes of access and delivery which fundamentally arise – technologically and cognitively – get crafted on the existing senses.

This novel perspective radically changes the way integration should be understood, that is with SSDs fitting, so to say, among the existing senses. What we want to suggest is that there is no reason to think about sensory substitution as being comparable to exercising of a single sense in the first place. SSDs are an augmentation and depend on the pre-existing senses, in a similar way that reading capacities for instance depend both on audition and vision (see Deroy & Auvray, 2012, for a full account). That is, reading brings a new capacity in the form of a single modality (vision) but actually depends on the existence of a first modality (auditory speech perception). SSDs do not introduce new kind of objects or rules in a modality which already exists, but bring a new capacity which must be analyzed as depending on two modalities, such as crossmodal capacities.

7. Conclusions

The common-sense intuition that SSD-use is somehow visual, lead to a dead-end given that SSD-use requires processing capacities from two senses (the substituted and substituting). As we have demonstrated, without both it is not possible to qualify the novel skills. Furthermore, trying to fit SSD-use into a single sense results in a philosophical dilemma which proves, in itself, impossible to overcome: Are trained users of the vOICe or the TDU exercising a form of vision, or some capacity which is only deceptively visual and remains in the intact modality?

Although many hope that clarification could come from an agreement on the individuation of the senses, we have argued that the problem is actually independent of the way one chooses to define vision or the other senses. The problem of classifying SSDs, in other words, arises however one chooses to individuate the senses. Contrary to Gray (2011) - we don't think that the impossibility to fit SSDs into the existing classification of the senses should lead us to abandon the hope to fit them into our typologies of the mind. What needs to be abandoned is only the project of fitting them among the canonical sensory modalities.

As the new capacities brought by SSDs do not and cannot figure on the same list as one of the natural or existing senses, SSDs need to be understood at the intersection of two senses:F or example, as analogous to the acquisition of reading. This alternate view avoids philosophical confusions and offers better ways to understand and push forward and higher the philosophical and empirical investigation of sensory substitution.

Acknowledgements
We would like to thank Dustin Stokes, Mohan Matthen; Stephen Biggs, and Jess Hartcher-O'Brien for their careful reading and comments on earlier drafts of this manuscript.
Ophelia Deroy is funded by an AHRC grant 'Rethinking the Senses'.

References

Arnold, G., & Auvray, M. (submitted). Perceptual learning: Tactile letter recognition transfers across body surfaces. *Multisensory Research.*

Auvray, M., & Deroy, O. (in press). How do synaesthetes experience the world? In M. Matthen (ed.) Oxford Handbook of Philosophy of Perception. Oxford: OUP.

Auvray, M., & Farina, M. (forthcoming) Patrolling the boundaries of synaesthesia : A critical appraisal of transient and artificially-induced forms of synaesthetic experiences. In O. Deroy (Ed.) Sensory blending: New essays on synaesthesia. Oxford: OUP.

Auvray, M., Hanneton, S., Lenay, C., & O'Regan, J. K. (2005). There is something out there: Distal attribution in sensory substitution, twenty years later. *Journal of Integrative Neuroscience*, *4*, 505-521.

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

Auvray, M., & Myin, E. (2009). Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cognitive Science*, *33*, 1036-1058.

Auvray, M., Philipona, D., O'Regan, J. K., & Spence, C. (2007-b). The perception of space and form recognition in a simulated environment: The case of minimalist sensory-substitution devices. *Perception*, *36*, 1736-1751.

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature, 221*, 963-964.

Bach-y-Rita, P., Tyler, M. E., & Kaczmarek, K. A. (2003). Seeing with the brain. *International Journal of Human-Computer Interaction*, *2*, 285-295.

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

Brand, P., & Yancey, P. (1993). *Pain: The gift nobody wants.* New York: Harper Collins.

Burge, T. (2010). *Origins of objectivity.* Oxford: Oxford University Press.

Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence.* New York: Oxford University Press.

Collins, C.C. (1985). On mobility aids for the blind. In D.H. Warren and E.R. Strelow (eds.) *Electronic spatial sensing for the blind,* Boston: Martinus Nijhoff, 107-120.

Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: a new perspective on sensory substitution. *Frontiers in Psychology*, *3*, 457, doi: 10.3389/fpsyg.2012.00457.

De Volder, A. G., Catalan Ahumada, M., Robert, A., Bol, A., Labar, D., Coppens, A, Michel, C., & Veraart, C. (1999). Changes in occipital cortex activity in early blind humans using a sensory substitution device. *Brain Research, 826*, 128-134.

Descartes (1637/1985). *Philosophical writings.* J. Cottingham, R. Stoothoff & D. Murdoch (tr.), Cambridge: Cambridge University Press

Evans, G. (1980). Things without the mind: A Commentary upon chapter two of Strawson's Individuals. In Z. Van Straaten (ed.). *Philosophical subjects: Essays presented to P. F. Strawson,* Clarendon, Oxford, 76-116.

Everson, S. (1999). *Aristotle on perception.* Oxford: Oxford University Press.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gray, R. (2011). *On the nature of the senses* In F. MacPherson (ed.) The senses, Oxford: Oxford University Press, 243-260.

Grice, H. P. (1962). Some remarks about the senses. In R. J. Butler (Ed.), *Analytical Philosophy (First Series).* Oxford: Basil Blackwell, 248-268. Rep. In F. MacPherson (2011) *The senses*, Oxford: Oxford University Press, 83-101.

Guarniero, G. (1974). Experience of tactile vision. *Perception, 3*, 101-104.

Hanneton, S., Auvray, M., & Durette, B. (2010). Vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics*, *7*, 269-276.

Hartcher O'Brien, J., & Auvray, M. (submitted). A novel perspective on distal attribution through the prism of sensory substitution studies. *Multisensory Research.*

Heil, J. (1983). *Perception and cognition.* Berkeley: University of California Press.

Heil, J. (2003). *From an ontological point of view*. Oxford: Oxford University Press.

Heil, J. (2011). The senses. In F. MacPherson (ed.) *The senses.* Oxford: Oxford University Press, 284-296.

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

Humphrey, N. (1992). *A history of the mind*. New York: Simon & Schuster.

Jansson, G. (1983). Tactile guidance of movement. *International Journal of Neuroscience, 19*, 37-46.

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

Keeley, B.L. (2013). What's exactly in a sense? In Simner, J. (ed) *Oxford Handbook of Synaesthesia* (pp. 941-958). Oxford: OUP

Kim, J-K, & Zatorre, R. J. (2008). Generalized learning of visual-to-auditory substitution in sighted individuals. *Brain Research*, *1242*, 263-275.

Kupers, R., Fumal, A., Maertens de Noordhout, A., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *PNAS*, *35*, 13256-13260.

Leon, M. (1988). Characterizing the senses. *Mind and Language*, *3*, 243-270.

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2012). Sensory substitution of vision: Importance of perceptual and cognitive processing. In R. Manduchi & S. Kurniawan (Eds.) *Assistive Technology for Blindness and Low Vision*. Boca Raton: FL: CRC Press.

Lopes, D.M. (2000) What is it like to see with your ears? The representational theory of mind. *Philosophy and Phenomenological Research, 60*, 439-453.

Matthen, M. (2007). Defining vision: What homology thinking contributes. *Biology and Philosophy*, 22, 675-689.

MacPherson, F. (2011, ed.). *The senses: classic and contemporary philosophical perspectives.* Oxford: Oxford University Press.

Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering,* 39, 112-121.

Milner, A.D. & Goodale, M.A. (2008). Two visual systems re-viewed. *Neuropsychologia,* 46, 774-785.

Morgan, M. J. (1977). *Molyneux's question. Vision, touch and the philosophy of perception.* Cambridge: Cambridge University Press.

Nagel, S.K., Carl C., Kringe T., Märtin R. & König, P. (2005). Beyond sensory substitution-learning the sixth sense. *Journal of Neural Engineering*, *2*, 13-26.

Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.

Nudds, M. (2004). The significance of the senses. *Proceedings of the Aristotelian Society, 104*, 31-51.

_________(2011). The senses as psychological kinds. In F. MacPherson (ed.) *The senses*. Oxford: Oxford University Press, 311-340.

O'Dea (2011). A proprioceptive account of the sensory modalities. In F. MacPherson (ed.) *The senses.* Oxford: Oxford University Press, 297-310.

O'Regan, J. K. (2011). *Why red doesn't sound like a bell: Understanding the feel of consciousness.* Oxford: Oxford University press.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*, 939-973.

Ortiz T., Poch J., Santos J. M., Requena C., Martínez A. M., Ortiz-Terán L., et al. (2011). Recruitment of occipital cortex during sensory substitution training linked to subjective experience of seeing in people with blindness. *PLoS ONE, 66*, e23264.

Peacocke, C. (1983). *Sense and content: Experience, thought and their relation.* Oxford: Oxford University Press.

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

Prinz, J. (2006). Putting the brakes on enactive perception. *Psyche, 12,* 1-19.

Ptito, M., Fumal, A., de Noordhout, A. M., Schoenen, J., Gjedde, A., & Kupers, R. (2008). TMS of the occipital cortex induces tactile sensations in the fingers of blind Braille readers. *Experimental Brain Research, 184*, 193-200.

Ptito, M., Moesgaard, S. M., Gjedde, A., & Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain, 128,* 606-614.

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C., & De Volder, A. G. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *NeuroImage, 26*, 573-580.

Ross (2001). Qualia and the senses. *Philosophical Quarterly, 51*, 495-511.

Roxbee Cox, J.M. (1970). Distinguishing the senses. *Mind,* 79, 530-550.

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

Siegle J. H., & Warren W. H. (2010). Distal attribution and distance perception in sensory substitution. *Perception*, *39*, 208-223.

Smith, A.D. (2002). *The problem of perception*. Cambridge: Harvard University Press.

Spence, C., & Deroy, O. (2013). Crossmodal mental imagery. In S. Lacey and R. Lawson (eds), *Multisensory Imagery: Theory and applications* (pp. 157-183). Philadelphia: Springer

Strawson, P.F. (1959). *Individuals: An essay in descriptive metaphysics*. London: Methuen.

Ward, J., & Meijer, P. (2010). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition, 19,* 492-500.

White, B. W., Saunders, F. A., Scadden, L., Bach-y-Rita, P., & Collins, C.C. (1970). Seeing with the skin. *Perception & Psychophysics*, 7, 23-27.