This is a pre-copyedited, author-produced PDF of a chapter published in de Vignemont F, Serino A, Wong HY & Farnè A (eds.) T*he World at Our Fingertips: A Multidisciplinary Exploration of Peripersonal Space*. Oxford: Oxford University Press, pp. 215-230.

Sameness of Place and the Senses

Alisa Mandrigin and Matthew Nudds

Abstract

When we watch a film at the cinema we typically experience the speech we hear as coming from the mouths of the actors depicted on the screen, rather than from the loudspeakers. This is an everyday example of the spatial ventriloquism effect. In this chapter we are interested in the question of what it is for things that we are aware of through different senses to appear to be in a single space, or even—as in spatial ventriloquism—at the same place.

The answer may seem trivial and obvious: all that is required is that we pick out places in the different senses in the same way. However, as Millikan (1991, 2000) has argued, representing a single location in the same way is not the same as representing sameness of location. What we need, either instead of, or as well as, sameness of reference frame, is for sameness of place to be a part of the content of experience.

Empirical evidence suggests that there exist peripersonal representations that encode multisensory information about the region of space that immediately surrounds the body, that contribute to goal directed actions and that play a role in mechanisms that protect the body. The existence of peripersonal representations generates a puzzle for accounts of perception: namely, what is the relation between peripersonal representations that figure in the empirical discussions and our everyday perceptual experience of ourselves and the world? Here we examine whether peripersonal space representations might play a role in out conscious awareness of the spatial relations between entities experienced in vision, audition and touch.

Keywords

Peripersonal Representation; Perceptual Experience; Multisensory Perception; Multisensory Interactions; Ventriloquism; Spatial Location; Egocentric Reference Frames

1 Introduction

The spatial ventriloquism effect occurs when there is spatial discrepancy between roughly synchronous visual and auditory stimuli.[1] The effect is produced when, say, a ventriloquist produces speech sounds without moving her lips, at the same time as manipulating the mouth of a dummy. Various aspects of the effect have been studied in the lab, with three main findings. Two involve localisation responses: there is mislocalization of both visual and auditory stimuli in the direction of one another; and there is a similar mislocalization aftereffect for visual or auditory stimuli presented

[1] Here after, for the sake of brevity, called the ventriloquism effect or ventriloquism.

independently of one another after a period of consistent spatial discrepancy between audio-visual stimulation (see Bertelson 1999 for an overview).

Vision and audition ground awareness of their objects as having locations in space. Given the patterns of localization responses in studies of ventriloquism, we might think of the effect as merely a case in which where we hear a sound to come from and where we see a material object to be each shift as a result of a multisensory interaction between vision and audition. Visual information about the location of a material object impacts auditory processing of information about the location from which a sound has emanated. Auditory information about the location from which a sound has emanated impacts visual processing of information about the location of a material object.

But, the consequence of ventriloquism is not simply a shift in the apparent location of what we see and hear, as measured by our spatial responses. The change is such that what we hear appears to come from what we see.[2] The speech produced by the ventriloquist appears to come from the dummy's mouth. One way to understand that is spatially: what we hear appears to come from the same place as what we see or, if we think we represent sound sources in audition, the source of what we hear appears to be at the same place as what we see.

Illusions such as the ventriloquism effect can serve to highlight features of everyday experience that would otherwise go unnoticed. It's not just in the bad (illusory) case that what we hear might seem to be at or to come from the same place as what we see. In the ordinary case when I watch and listen to you speaking, I see the movement of your lips and I hear the sounds you produce. The sounds appear to come from the place I see your lips to be. Touch also grounds awareness of its objects as having locations in space. The things that we see, hear and touch can appear to us to be in or coming from the same place.

What is it in general for two things to appear to be in the same place across modalities? A first thought might be that it is to experience places in such a way that we can draw the following inference, where F is an auditory feature, and G is a visual feature:

> That (visual) place is G.
> That (auditory) place is F.
> Therefore that (visual) place is F and G.[3]

[2] There may be still more to the illusion: it might be claimed that, additionally, ventriloquism produces an illusory cross-modal experience of a speech event, or an illusory cross-modal experience of causation.
[3] In the case of ventriloquism this is, of course, not true: that visual place and that auditory place are not in fact the same place, so the inference is not valid.

That is, if we visually experience a place as having some visual feature, and auditorily experience the same place as having an auditory feature, we can immediately draw the conclusion that the place has both auditory and visual features. But how is this possible? To say that we can immediately draw the conclusion means that no collateral information, no further premise to the effect that the place we see is identical to the place we hear, is required to make the inference. We have two distinct applications of a demonstrative concept—two judgements to the effect that 'that place is *x*'—from which we can conclude that 'that place is F and G'. And the thought might be that I have two experiences that ground the application of the same demonstrative concept. Hence, I can draw the conclusion about *that place* on the basis of my visual and auditory experiences of it.[4] This is certainly one way that we can think of the connection between our experiences of places across the senses.

But consider again the ventriloquism effect. The third measure used in studies of the ventriloquism effect is that of 'perceptual fusion'. Subjects are asked to indicate when they have "an impression of common origin" (Bertelson 1999, 348). That is, they are asked when they have the impression that the visual and auditory stimuli are at the same place. In many studies subjects report that the visual and auditory events *appear* to them to have occurred in the same location, despite the two stimuli being spatially disparate (Bertelson & Radeau, 1981, Choe et al. 1975; Jack & Thurlow, 1973; Thurlow & Jack, 1973; Witkin et al., 1952).

In ventriloquism I *experience* the source of the sound as at the same place as the visually perceived object. Sameness of place is part of the content of experience. Experience presents the places as one and the same, such that no drawing of the conclusion is required. We don't need to make the inference because perceptual experience is such that the identity of place will be apparent. So, we need an account that does more than explain how it is that we can perceive two things or two features in audition and vision such that it's transparent to thought that the same place is being experienced. We need to explain how it is that their shared location is perceptually apparent to us: we need to explain how experience presents visible, audible and tactual features as features of a single place.

[4] In such a case, it will always be possible for someone to fail to reflectively appreciate that the place picked out in one experience is identical with the place picked out in the other experience, perhaps due to inattention. Thank you to an anonymous reviewer for drawing our attention to this way in which we might fail to appreciate identity of place. Still, the suggestion is that they will nevertheless have reasonable grounds to draw the inference: if they considered the matter, it would not be rational for them to fail to draw the conclusion that the place they visually experience is identical with the place they auditorily experience.

2 A common reference frame

When we experience the same place in vision and touch, what is required for it to be apparent to us that it is the same place? One answer is that sameness of place will be apparent if we pick out or identify places in the same way irrespective of the sense modality of experience.

When we pick out places we must do so relative to something. That is, we need a frame of reference within which, or relative to which, the place can be picked out. The frame of reference provides a means for picking out a place by locating it relative to some other thing or things. If a place is picked out on two occasions in the same way relative to the same frame of reference, then it will be the same place. Applied to perceptual experience, the idea is that experiences that present places must do so relative to a frame of reference. So places are presented in experience, and have their identity, relative to that frame of reference. One way in which two features can be presented in experience as being at the same place is by being presented as being in the same relative place within the same frame of reference.

If our perceptual experiences in different sense modalities present places in a frame of reference that identifies places in the same way, then it might be tempting to assume that when two experiences identify the same place, it will be apparent to the perceiver that it is the same place.

According to Evans' account of spatial content (1982, 1985), spatial locations are represented in perceptual experience in an egocentric frame of reference. To label a reference frame as 'egocentric' is to say that it individuates locations relative to the perceiver's body.[5] For Evans, perceptual experience has egocentric spatial content in virtue of its relations to behaviour. To perceive the cup's egocentric spatial location is to be disposed to engage in a repertoire of behaviours that are appropriate, given the spatial location of the cup relative to oneself. That is, to be represented as having a location is to be represented as having a property whose significance is cashed out in terms of relations to a certain behavioural repertoire.[6]

Now, according to Evans, because "there is only one behavioural space" (1985, p. 390), there is therefore only one egocentric space. If both vision and audition represent things that are in the same

[5] The term 'egocentric' has come to be used across a variety of disciplines to talk about spatial representations at different levels of explanation. For example, psychologists and cognitive neuroscientists suggest that spatial information is encoded in early sensory processing in a number of different egocentric reference frames, each taking a different part of the body as its point of origin. This is a claim about sub-personal representation. Philosophers have also claimed that conscious perceptual experience locates objects and events by means of an egocentric reference frame. But this is a personal level claim: the claim that conscious perceptual experience has egocentric content.

[6] Campbell (2005) suggests that a natural way to understand Evans' claim here is in terms of affordances (Gibson 1979). Egocentric spatial content is constituted by its implications for behaviour in that identifying the egocentric location of an object is identifying the current affordances of the object given its location relative to one's own body. Put simply, places are tied to responses.

place—just to my right, say—they each represent their respective objects as reachable in the same way, that is, as requiring the same movement in order to reach them. Since the behavioural repertoire required to reach an object will be the same irrespective of modality, and since for Evans egocentric content identifies the locations of things by means of the bodily behaviours that they afford, by being represented at the same egocentric location, the location of something that is seen and something that is heard will be represented in the same way.

If we understand egocentric spatial content in this way, then we will endorse the following. When I hear the ringing of the doorbell and see the steam rising from the kettle, my perceptual awareness of the spatial locations of each of the events will be cashed out for each event in terms of dispositions to perform bodily movements. When I hear the whistle of the kettle and see the steam rising from the very same kettle, my awareness of the spatial location of these events is to be explained in terms of my dispositions to perform the same set of bodily movements. I see that place (in front of me) is F, and hear that place (in front of me) is G. If Evans is right, then I'm certainly in a position to judge that that place (in front of me) is F and G. I am able to think 'that place in front of me is F', and 'that place in front of me is G', and I can conclude that 'that place in front of me is F and G'.

But it's tempting also to think that, if places are individuated in visual and auditory experience in the same way, it will be apparent that it is the same place that is F and G. Can we explain how sameness of place can be experientially apparent simply by appealing to the idea that experiences in different sensory modalities individuate places in the same way?

According to Peacocke (1986), when a location is represented in experience in the same way twice, then the identity of the location will be experientially apparent to the perceiver. This will be so whether both things are experienced within the same modality, or across modalities. In discussing perceptual experience of distances, he says that:

> "we should require that if μ is the manner in which one distance is perceived and μ′ is the manner in which a second distance is perceived by the same subject at the same time, and μ = μ′, then the distances are experienced as the same by that subject" (Peacocke, 1986, 5).

Peacocke uses the term 'manner' to speak of the way in which an object or property is presented in perceptual experience. To give the content of a perceptual experience, he suggests, we must give an account of the manner or way in which the things that are perceived are perceived to be.

Applied to the experience of places, if *x* and *y* are at the same place and if *x*'s location is represented in vision in the same manner or way (and at the same time) as *y*'s location is represented in audition, then, according to Peacocke, *x* and *y* will be experienced as at the same location. Indeed, Peacocke

stipulates that spatial locations must be represented in the same way across modalities on the grounds that it is this that allows us to be perceptually aware of the sameness of direction, size, and so forth, across sense modalities:

> " […], there is a second requirement on the manners. This is the requirement that they be *amodal*, in the sense that the same manner can enter the content of experiences in different sense modalities. You may hear a birdsong as coming from the same direction as that in which you see the top of a tree: we would omit part of how the experience represents the world as being were we to fail to mention this apparent identity. It also makes sense to say that something feels roughly the same size as it looks." (1986, 6)

We agree with Peacocke that to fail to mention that experience presents the identity of places across the sense modalities would be to leave out part of how experience presents the world as being. But, with Millikan (1991, 2000), we want to deny that sameness of reference frame will deliver a representation of sameness of place. Millikan (1991, 2000) points out a number of mistakes that can result from not adequately distinguishing between the personal and the subpersonal, and, in particular, attributing a property at one level because we have reason to attribute it at the other level. We earlier introduced the notion of an egocentric reference frame at the personal level, but talk of egocentric reference frames can also be pitched at the subpersonal level. To say that we have evidence of the existence of egocentric reference frames at the subpersonal level is to say something about the neural vehicles of representational content. To claim that egocentricity is a personal-level feature is to make a claim about the representational content of experience: it is a claim about what is represented. The mistakes that Millikan has in mind stem from confusing the content of experience—what is conveyed to the subject by her experience—with properties of the vehicles of representation, and *vice versa*.

One of the mistakes that Millikan draws attention to is that of confusing the sameness of the vehicles of representation with the representation of sameness (2000, 2). That is, we make the assumption that, because there are two instantiations of the same vehicle at the same time, the relation of sameness that holds between the vehicles of representational content will itself feature as part of the content of perceptual experience. In general, it's the mistake of assuming that all properties of the vehicles of representation will equally be part of the content of experience. Since vehicles are real, concrete things then, if they are the same, they will bear the sameness relation to one another. The mistake arises if we then project, without argument, the sameness of the vehicles into the content of perceptual experience. That is, if we don't keep in mind that vehicles of representation are distinct from their representational content, then we might make the mistake of assuming, without offering any argument, that sameness will be part of the content of experience. To use

Millikan's terminology, we're guilty of 'externalising' the sameness of the vehicles—the sameness of the sub-personal neural vehicles of representation—into the content of experience. The claim, then, is that there is a difference between perceptually representing a location twice in the same way, and perceptually representing the identity of the location experienced in two sense modalities.

This can be backed up by considering how relations between places are represented. Suppose that your hand is hidden beneath a table, and we project a spot of light onto the surface of the table. Would it be possible for you to tell where the spot is in relation to your hand? In other words, would you be able to tell that the spot of light is to the left or right of your hand? One way you could tell would be to judge where the spot is in relation to some part of your body—your nose, perhaps—judge where your hand is in relation to that same body part, and then work out where the spot is in relation to your hand. For example, you might judge that the spot is 30˚ to the right of your nose, that your hand is 20˚ to the right of your nose, and work out that the spot is 10˚ to the right of your hand. In this case we are supposing that both locations are given in the same way, but you still have to work out the non-egocentric relation between the two things, between the spot and your hand. Millikan's point becomes clear when we notice that it would make no difference to this example if the spot and your hand are in the same place. You would still need to make a judgement about the egocentric location of your hand and the spot of light, and then work out that they are in the same place. The working out might seem to be a trivial step, but it's still a step.

So, we want to argue, being represented in the same way is not in general sufficient for us to experience sameness; some further conditions must be met. If the location of *x* is represented in vision, and the location of *y* is represented in audition, and *x* and *y* have the same location, it doesn't follow merely from the fact that their location is represented in the same way that it will appear that x and y are in the same location. Sameness of representational content does not automatically produce an experience of sameness.

Even if we can show that we are aware of locations in such a way that it is transparent when two things have the same location—i.e., that they are represented in such a way that we are in a position to draw the inference that they have the same location—that still falls short of what we want to explain. In ventriloquism cases, what we hear appears to come from the same place as what we see. We don't have to reflect on our experience in order to judge that this is the case, just as we don't have to reflect on our experience when we see two things as in the same place. Sameness of location is part of the content of our experience. We need to explain how that is so. Appealing to the *way* locations are represented does not do that.

So, the conclusion we can draw from this is not that we do not experience sameness of location when locations are represented in the same way. Sameness of location is part of the content of our experience and we need to explain how that is so. What we should conclude, though, is that simply appealing to the *way* locations are represented does not help us to provide that explanation.

3 Cross-modal Experiences

If all experience gave us was a place presented in the same manner, or reference frame, we would only be able to become aware that the source of a sound is in the same place as something that we see when we engage in an act of reflection. But there is an alternative and stronger sense in which I might be aware of sameness of place across modalities: I don't have two experiences—visual and auditory—of the same place, but rather have a cross-modal experience of a place. Rather than making two distinct demonstrative judgements from which I can draw the conclusion that *that place is F and G*, I have a single cross-modal experience that grounds the judgement 'that place is F and G', where F and G are features perceived with different senses.

What does the idea that we have cross-modal experiences of the same place (that is, of a single cross-modal experience that grounds the judgement 'that place is F and G', where F and G are features perceived with different senses) imply about hearing and seeing something at the same place? There are two ways we might develop that idea.

The first is that we think of places as individual things, so hearing and seeing something as the same place is a matter of hearing and seeing the same individual place. It turns on the question of whether we can experience the same individual as such with two or more senses. Of course, we can experience the same thing in more than one sense: I can see a cube that I hold in my hand, for example. But the question is whether it is somehow perceptually apparent to me that it is the same thing that I both see and touch. If we treat individuals as places, then we might think that the issues that arise for this question will be the same as those that arise with respect to our experience of sameness of place. We'll set this question aside here (see Bayne 2014; Nudds 2014; O'Callaghan 2016 for discussion of cross-modal experiences of objects).

The second is that hearing something as coming from the same place as something that you can see involves being aware of the spatial relation that exists between them. For two things to be in the same place is for them to be spatially related to each other in a certain way (e.g. next to each other, on top of each other), or in a shared spatial relation of a certain kind to a third thing (e.g. on the table, in the box). So to perceive two things as in the same place might instead involve perceiving the

spatial relation between those things. Can we experience the spatial relations between things perceived across the senses?

If two objects are located using an egocentric frame of reference, it is possible to compute where they are relative to one another. So, if object A is represented as being 20˚ to my right, and object B is represented as being 30˚ to my right, then I can work out that object B is 10˚ to the right of object A. If asked the colour of the pen to the left of the coffee cup, I can project my own left and right onto the cup and use that frame of reference to identify the relevant pen. For us to work this out we engage in some reasoning. However, it might be that in many cases reasoning about the locations of objects relative to one another isn't necessary, because experience represents the spatial relations between things by representing both egocentric and allocentric spatial locations.

The idea here is that there is a contrast we can draw between our working out the spatial relation between two things we experience, and experience representing the spatial relation between those two things. This matches the contrast we have drawn above between working out sameness of place on the basis of two applications of the same demonstrative concept, and sameness of place being perceptually apparent to us. The suggestion is that there are processes or mechanisms that encode the spatial relations between the objects that we experience in a way that makes the information available in some distinctively experiential way.

There is empirical evidence that the visual system encodes not only the egocentric location of objects, but also their allocentric location (that is, the location that objects stand in to one another). Using a series of reaction time experiments Gordon D. Logan (1995) has shown that we are able to direct visual attention to an object by computing the spatial relation between one object and another. In Logan's experiments subjects were presented with a visual array and a cue that directed their attention to an item in the array based on its spatial location relative to the cue.[7] Subjects were required to respond by indicating which item in the array had been cued. There was no way of performing the task successfully other than by using the cue to select the object that satisfied the relation between the cue and the target, or by guessing (1995, p. 119). Logan found that reference frames "can be rotated and translated across space according to the intentions of the observer and they can be aligned with the intrinsic axes of attended objects" (1995, p. 169). The speed and accuracy of subjects' responses indicate that subjects can direct their attention without engaging in reasoning. Instead, the computations required to specify which location is to be attended are carried out by the visual system. This suggests that *within vision* allocentric spatial relations can be

[7] Logan's paradigm was based on experiments by Eriksen & Hoffman (1972).

represented.

There are, then, two ways in which we can be aware of the spatial relations between two things that are given egocentrically in experience. It is possible to derive information about the spatial relations between objects from information about their egocentric location. We can derive this information non-perceptually by calculation or reasoning, but it's also possible for this information to be derived perceptually via the operation of sub-personal computational processes.

A perceptual system may or may not carry out the computational procedure. For example, it may be that while both vision and audition represent egocentric information about location, only vision computes intrinsic (i.e., allocentric) locations, so only vision represents the spatial relations between the things it represents. Whether computations are performed by a system is an empirical matter.

We can draw the same contrast between working out the spatial relations between two things and experiencing the spatial relations between those things for two entities experienced in different sense modalities. On this approach, experiencing sameness of place across the senses will just be a version of experiencing the spatial relations between two things experienced in different sense modalities.

And just as we can ask whether the visual system or the auditory system computes allocentric locations, we can ask whether the spatial relations between objects experienced in different sense modalities are computed. That is, do the perceptual systems generate multisensory spatial representations? The notion of a multisensory spatial representation takes us beyond the idea of sameness of reference frame across the senses and of places being represented in the same way. It doesn't follow simply from the fact that locations are given egocentrically in vision and in audition that we are perceptually aware of the relations between the entities we experience in the two senses. What is required is that the perceptual systems compute the allocentric locations of things experienced in the different sense modalities.

What would multisensory spatial representation get us that representing space in the same way in the different sense modalities does not? Our answer is that it would ground a certain kind of experience of unity across the senses that goes beyond judgements of sameness of place. Multisensory spatial representations would play a distinctive experiential role that isn't provided by any of the senses operating on their own. But are there representations of this kind?

4 Peripersonal Space Representations

While researchers working on the science of perception have historically tended to study each sensory system in isolation from the others, there is a large and growing body of evidence indicating that the sensory systems interact with one another. Various multisensory effects, including the spatial ventriloquism effect, indicate that, for example, processing in regions traditionally associated with vision can impact processing in regions traditionally associated with touch, and so forth. What is more, some areas of the brain involved in the processing of sensory stimulation seem to be multisensory: some neurons within these areas respond to stimulation in more than one sense. A number of regions have been identified as being multisensory in this way, including the superior colliculi (see, e.g., Meredith & Stein 1986), and parts of the parietal and frontal cortex involved in what have come to be known as peripersonal space representations.

The term 'peripersonal space' refers to the area immediately around the perceiver, usually extending not more than 30cm from the body (Rizzolatti et al. 1981a, 1981b). Evidence of multiple multisensory representations of peripersonal space comes from a number of sources: single cell recordings in monkeys; neuropsychological studies in humans with disorders of spatial attention resulting from lesions in the parietal and frontal cortex; and, neuroimaging of neurologically healthy humans.

For example, Rizzolatti and colleagues (1981a, 1981b) recorded the activity of single neurons in ventral premotor areas of the macaque brain. They found that some neurons responded to both tactile stimuli on specific parts of the body and to visual stimulation. Importantly, these 'bimodal' neurons were responsive to visual stimuli presented within the peripersonal space of the particular body part that elicited a tactile response, but not to visual stimuli presented outwith peripersonal space. The tactile and visual receptive fields of these neurons are aligned with one another, and remain in register even when, for example, the limb of the monkey was passively displaced (Graziano et al. 1994). Bimodal neurons in these areas are not only responsive to visual and tactile stimuli: Graziano and colleagues (1999) found that some neurons in premotor area F4 respond to both tactile stimulation on a body part and auditory stimuli presented within peripersonal space.

Evidence of the existence of multisensory representations of peripersonal space is not limited to neurophysiological studies involving animals. For example, neurological patients with lesions to the frontal and parietal cortex often display contralesional extinction. That is, when two stimuli are presented at once, one in each hemifield, the stimulus presented on the contralesional side is not detected. This is so even though a single stimulus presented on either side can be detected. Contralesional extinction is found within a sense modality (i.e., when the two stimuli are both visual), but also across the senses. This is taken as evidence of multisensory peripersonal space representation because crossmodal extinction is stronger when the visual stimulus is presented

within peripersonal space compared to when it is presented in far space (di Pellegrino et al. 1997; Làdavas et al. 1998a, 1998b).

Di Pellegrino and Làdavas conclude that:

> "Collectively, these studies reveal that the primate brain constructs multiple, rapidly modifiable representations of space, centered on different body parts (i.e., hand-centered, head-centered, and trunk-centered), which arise through extensive multisensory interactions within a set of interconnected areas in the parietal and frontal cortex." (di Pellegrino & Ladavas 2015, p. 127)

Should we think of these peripersonal space (PPS) representations as multisensory representations of space of the kind that would allow for sameness of place across the senses to be perceptually apparent to us?

It's worth noting that the term 'representation of space' will typically be used to mean different things in perception science and in the philosophy of perception. All the perception scientist might be committed to in positing a representation of space is the existence of a neuron or set of neurons that respond differentially to the spatial location of a stimulus. So, a neuron that responds differentially when a visual stimulus is within 10cm of the perceiver's hand and when it is outwith 10cm of the perceiver's hand will count as representing space (or spatial location). So, on the perception scientist's understanding of the term, PPS representations will certainly count as representations of space.

But what we are interested in is something more. We want a structure that is capable of delivering information about the spatial relations that a number of different objects stand in to one another. In particular, we want a structure that delivers information about the spatial relations that objects perceived through different sense modalities stand in to one another. What reasons might we have to extrapolate from the empirical work to the claim that peripersonal space representations are representations of space?

Peripersonal space representations are typically described as being map-like, or as mapping the space around the body (see, for example, Grivaz *et al*. 2017, p. 603 and Graziano & Gross 1993, p. 107). A map represents a number of places, as well as the distance and direction of each of those places from each of the other places that is represented, and perhaps also such that for any two places represented, every place in between them is represented. A nucleus is typically described as being map-like if the neurons that make up the nucleus are arranged topographically, so that there is

an isomorphism between the spatial relations between particular neurons and the spatial relations between the receptive fields of those neurons.

But having this kind of structural organisation doesn't seem to be sufficient for a representation to count as a representation of space, unless there is some way for the perceptual system to be able to extract information about the spatial relations between places from the spatial relations between neurons in the structure in question. In thinking of representations of space as map-like, what matters is the *kind* of information that is represented—spatial relations between places being encoded, for example—rather than the structure of the representation itself. What makes something a spatial representation is that it has the *functional* properties that maps have.

Gareth Evans (1982, 1985) raises this issue in his discussion of egocentric spatial content. He questions how it is that sensory inputs which convey spatial information only in so far as they stand in some kind of systematic relation to other possible sensory inputs, come to have spatial significance for the perceiving subject. According to Evans:

> "[…] an egocentric space can exist only for an animal in which a complex network of connections exists between perceptual input and behavioural output. A perceptual input—even if, in some loose sense, it encapsulates spatial information (because it belongs to a range of inputs which vary systematically with some spatial facts)—cannot have a spatial significance for an organism except in so far as it has a place in such a complex network of input-output connections." (1982, p. 154)

Evans' proposal is that what connects a visual or auditory representation with physical space or environmental locations is, in part, dispositions to move in certain ways with respect to locations in physical space. In the absence of such connections, we lose our grip on the idea that visual or auditory experience has spatial content at all. So, one way in which we might try to establish that peripersonal representations should count as multisensory representations of space is by establishing that they bear the right kind of connections with motor behaviour.

On first glance, things might seem promising. Many of the visual-tactile neurons in PPS areas respond not only to visual or tactile stimulation, but also during motor action (Brozzoli et al. 2012, p. 451; de Vignemont 2018). And it has been suggested that the connections between peripersonal space representations and the motor system indicate that one of the main functions of PPS representations is to facilitate interactions between the perceiver and objects in her environment (Rizzolatti et al. 1991b; Brozzoli *et al*. 2012; Brozzoli *et al*. 2014).

But while there is evidence that areas responsive to sensory stimulation in PPS are closely connected with the motor system and that peripersonal representations have a role to play in the guidance of motor action, it's not clear that they will also underpin conscious perceptual experience in the way that would be required for peripersonal representations to provide for an awareness of sameness of place across the senses.

First, some of the areas involved in peripersonal space representation that are not part of the premotor cortex are part of the dorsal stream of visual processing (Brozzoli *et al*. 2012). Recent empirical work has been taken to support our making a functional distinction between vision for perception—subserved by ventral processing—and vision for action—subserved by processing in the dorsal stream. Evidence of the impact of lesions in either the dorsal or ventral streams in macaques (Ungerleider and Mishkin 1982), from neurological patients with visual form agnosia or blindsight on the one hand and optic ataxia on the other hand (Goodale & Milner 1992; Milner and Goodale 1995), and differences in the verbal and motor responses made by healthy human subjects to illusory stimulation have been taken to indicate that visual information is processed differentially depending on the task to which it is to be put. Taken together these cases appear to offer us a double dissociation between these two functions, indicating that the processing on which each supervenes is localized in different parts of the brain. The dorsal visual stream is (now) typically conceived of as being 'for action'—devoted to the programming and control of motor acts—while the ventral stream in vision is thought of as being 'for perception', bringing to awareness information about objects required to identify and remember them (Goodale & Milner 1992).

The dual visual systems hypothesis is controversial. Yet, the claim that we have an unconscious, online system dedicated to the processing of visual information for the guidance of motor actions, and the fact that it is within this system that peripersonal space representations are to be found, should alert us to the possibility that peripersonal representations do not provide for our perceptual awareness of the spatial relations between objects experienced in different sense modalities.

Secondly, the connections to motor action that have been uncovered look to give us, at most, perception of egocentric location only. What we seem to have is a representation of the spatial location of a visible object relative to the perceiver's body. But we have been arguing that representation of spatial location relative to the perceiver is not enough to give us a representation of sameness of place across the senses. We need something further.

What makes my Ordinance Survey map a useful tool for understanding the spatial relation between two geographical features is not that there exists a piece of paper with two marks on it, each of which I can use to guide motor actions towards the respective geographical feature that the mark

represents. What allows me to grasp the spatial relation that the two geographical features stand in to one another is that I am able to see where the marks on the paper are in relation to one another. Similarly, what will make a topographically organised layer of neurons a representation of space of the kind we are interested in, is the existence of a mechanism that extracts the spatial relation between cells that are active in that layer so that the spatial relation that the two relevant stimuli stand in to one another can itself be represented. In fact, the topographically organised layer of neurons in isolation wouldn't count as a representation of space: it would be the layer of neurons *together with* the mechanism that gathers and compares relative spatial positions of active neurons that would be considered a representation of space of the sort we are after. Evidence of the existence of PPS representations, even if they have a topographic arrangement, does not give us a structure that can support our conscious awareness of the unity of space unless we have evidence of an associated mechanism that extracts relevant information about the spatial relations between neurons.

Finally, it's not clear that peripersonal representations exhibit the spatial specificity for them to underpin our conscious awareness of the spatial relations between entities experienced through different sense modalities. For example, the tactile receptive fields of bimodal neurons in premotor area F4 are considered to be large, encompassing in some cases whole body parts such as the hands, and many have overlapping receptive fields. For this reason, PPS neurons are thought to form only "a crude map of the body surface" (di Pellegrino & Làdavas 2015, p. 127).

We therefore have a structure that may underpin our awareness of only the rough or approximate spatial relations that the objects we perceive stand in to one another. For example, investigation has revealed a distinction between PPS cells that respond to visual stimuli that are within 10cm of the macaque's hand, and those that respond to visual stimuli that are between 10cm and 30cm of the macaque's hand. If we assume the existence of a mechanism that compares responses in these two types of cells, this would allow for presentation as of a visible object being anywhere within a region of space that is between 10cm and 30cm from an object near to the subject's hand.

In response to this, it might be pointed out that many PPS neurons exhibit superadditive effects. If the macaque is presented with both a visual and tactile stimulus within the receptive fields of a particular PPS neuron, then activity in the neuron will supercede the sum of the activity produced by the same neuron in response to a visual stimulus alone and a tactile stimulus alone. Could this at least underpin an awareness that a seen object and a felt object are in the same place? Superadditivity, the thought might be, could encode for sameness of place across vision and touch.

Again, the problem is that of specificity. Many PPS neurons respond to both tactile stimulation on a body part, and to visual stimulation within a region of space around the body part in question. So the same neuron will respond when a visible object is directly adjacent to an object that touches the relevant body part of the perceiver, and when a visible object is located 9cm from an object that touches the body part. Depending on the fineness of grain with which we are interested, we might consider both of these two cases to be cases in which something that is seen is in the same place as something that is felt. However, it seems to be at odds with our perceptual awareness, which admits of finer grained distinctions. It is possible, for example, for typical perceivers to visually distinguish a case in which two pencils that are within arms' reach are lying in parallel next to one another from a case in which two pencils that are within arms' reach are lying in parallel but 9cm apart.

PPS neurons therefore don't seem to have the specificity required to account for the determinacy of our perceptual awareness. What is more, many of the neurons investigated respond to somatosensory stimulation on either of two unrelated parts of the body. A single neuron might respond to either touch on the right hand or to touch on the right cheek. Rizzolatti and colleagues therefore conclude that:

> "[…] it is difficult to imagine that neurons that respond equally well to a stimulation of different, spatially segregated parts of the body have as their main function that of informing the organism where the stimulus is located. This is even more the case if they also respond to visual stimuli." (Rizzolatti et al. 1991b, p. 160)

While peripersonal space representations may, in some sense, map the space immediately around a perceiver, we agree with Rizzolatti and colleagues that it doesn't look as though they do so in a way that will underpin the perceiver's perceptual awareness of the spatial relations that the objects she perceives in the different senses stand in to one another.

5 Conclusion

In ordinary, everyday situations we don't take ourselves to be encountering distinct and discrete spaces when we perceive the world through more than one sense modality (Eilan 1993). I might hear a sound whose source I cannot see because it is too far away. Still, I don't take the place at which that sound source is located to be isolated from the visible objects I see in my immediate environment. We are able to answer questions about the spatial relations between the things that

we see, hear and touch. We can, for example, say that the place the high-pitched beeping is coming from is to the left of the flashing light.

From an ecological perspective, all that we need is the capacity to make use of sensory information in order to perform successful motor actions within a suitable time frame. Given that we have multiple sensory systems that process information about the location of objects and features in our environment, there is a problem to be solved: how is it that the spatial relations between things perceived through different modalities come to be resolved so that we can act successfully on the world using sensory information from more than one sense at a time? The problem might be solved in a number of ways. A computation that maps location L (as encoded in one sensory system) onto location L' (as encoded in another sensory system) might be built into the architecture of the perceptual systems. Or there might be a common way of encoding spatial information across the sensory systems, so that no translation is required. Howsoever the problem is solved, it doesn't matter for our survival that we perceive the unity of space. Yet, in some cases, the objects we encounter in different senses appear to us to be located within a single space, or, even, in the same place. The speech I hear seems to come from the same place as the movement of the lips that I see. This is highlighted for us in the case of spatial ventriloquism, when, for example, the voices of the actors perceptually appear to come from the place that we see lips moving on the cinema screen. That is, it perceptually seems to us that the actors' voices come from the same place at which we see the lips move.

Empirical evidence suggests that there exist peripersonal representations that encode multisensory information about the region of space that immediately surrounds the body, that contribute to goal directed actions and that play a role in mechanisms that protect the body. The existence of peripersonal representations generates a puzzle for accounts of perception: namely, what is the relation between the peripersonal representations that figure in the empirical discussions and our everyday perceptual experience of ourselves and the world?

Here we have considered whether PPS representations might underpin our conscious appreciation of the unity of the space we encounter in the different sense modalities. That is, we have examined whether PPRs might play a role in our conscious awareness of cross-modal spatial relations, including sameness of location. Many peripersonal neurons are multisensory, their tactile receptive fields have rough somatotopic organisation, and there is spatial alignment between their tactile receptive fields and their visual or auditory receptive fields. They are often said to represent the space immediately surrounding the perceiver. However neurophysiological studies on macaques reveal that peripersonal space representations have a number of properties that suggest they are ill-suited to

underlie conscious awareness of the unity of space across the senses: the receptive fields for neurons are large and overlapping, a single neuron will often be responsive to somatosensory stimulation on different, spatially separated parts of the body, and the same neuron will respond to both somatosensory stimulation on a body part, and visual stimulation in the space around that body part. Moreover, these bimodal neurons are mostly found in the premotor cortex or in parts of the visual dorsal system, suggesting they play a role in controlling the way we interact with our environment, without impacting on conscious experience. It's still possible that PPS neurons might play a role in our consciousness awareness of the spatial relations between objects perceived through more than one sense modality. There might, for example, be a mechanism that extracts information about the spatial relations between receptive fields of PPS neurons in order to generate a representation of the spatial relations between stimuli. Until this has been investigated, though, we think that we should withhold from judgement about whether PPS representations underpin a cross-modal representation of space.

References

Bayne, T. [2014]: 'The Multisensory Nature of Perceptual Consciousness', in D. Bennett and C. Hill (*eds*) *Sensory Integration and the Unity of Consciousness,* Cambridge, MA.: Massachusetts Institute of Technology Press, pp. 15–36.

Bertelson, P. (1999). Ventriloquism: A Case of Crossmodal Perceptual Grouping. In G. Aschersleben, T. Bachmann & J. Müsseler (eds.) Cognitive Contributions to the Perception of Spatial and Temporal Events. Elsevier Science B.V., pp. 347-362.

Bertelson, P. & de Gelder, B. (2004). The Psychology of Multimodal Perception. In C. Spence & J. Driver (eds). Crossmodal Space and Crossmodal Attention. Oxford: OUP, pp. 141-177.

Bertelson, P. & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory- visual spatial discordance. *Perception & Psychophysics, 29:* 578-587.

Brozzoli, C., Ehrsson, H.H., & Farné, A. (2014). Multisensory representation of the space near the hand: From perception to action and interindividual interactions. *The Neuroscientist*, 20(2): 122-135.

Brozzoli, C., Makin, T.R., Cardinali, L., Holmes, N.P. & Farnè, A. (2012). Peripersonal space: A multisensory interface for body-object interactions. In M.M. Murray & M.T. Wallace (Eds.) *The neural bases of multisensory processes*. Boca Raton, FL.: CRC Press.

Choe, C.S., Welch, R.B., Gilford, R.M., & Juola, J.F. (1975). The "ventriloquist effect": Visual dominance or response bias? *Perception & Psychophysics*, 18, 55-60.

Campbell, J. (2005). Information-processing, phenomenal consciousness and Molyneux's question. In J.L. Bermúdez (ed.) Thought, Reference, and Experience: Themes From the Philosophy of Gareth Evans. Oxford: Clarendon Press, pp. 195-219.

di Pellegrino, G. & Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia,* 66**,** 126-133.

di Pellegrino, G, Basso, G. Frassinetti, F. (1997). Spatial extinction on double asynchronous stimulation. Neuropsychologica, 35, 1215-1223.

Eilan, N. (1993). Molyneux's question and the idea of an external world. In N. Eilan, R. McCarthy and B. Brewer (eds.) Problems in the philosophy and psychology of spatial representation. Oxford: Blackwell, pp. 236-256.

Eriksen, C.W., & Hoffman, J.E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics*, 12: 201-204.

Evans, G. (1985). Molyneux's Question. In his Collected Papers. Oxford: Clarendon Press, p. 364-399.

Evans, G. (1982). The Varieties of Reference. Oxford: Clarendon Press.

Goodale, M.A. and Milner, A.D. (1992). Separate Visual Pathways for Perception and Action. *Trends in Neurosciences*, 15(1), pp. 20–5.

Graziano, M.S.A., Yap, G.S., & Gross, C.G. (1994). Coding of visual space by premotor neurons. *Science*, 266, 1054–1057.

Graziano, M.S., Reiss, L.A., & Gross, C.G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, 397, 428–430.

Graziano, M.S. and Gross, C.G. (1993). A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Experimental Brain Research,* 97**,** 96-109.

Grivaz, P., Blanke, O. & Serino, A. (2017). Common and distinct brain regions processing multisensory bodily signals for peripersonal space and body ownership. *NeuroImage*, 147: 602-618.

Jack, C.E. & Thurlow, W.R. (1973). Effects of degree of association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills*, 37, 967-979.

Làdavas, E., di Pellegrino, G., Farnè, A., Zeloni, G. (1998a). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *Journal of Cognitive Neuroscience*, 10, 581–589.

Làdavas, E., Zeloni, G., Farnè, A. (1998b). Visual peripersonal space centred on the face in humans. *Brain*, 121, 2317–2326.

Logan, G.D. (1995). Linguistic and Conceptual Control of Visual Spatial Attention. *Cognitive Psychology*, 28: 103-174.

Meredith, M.A. & Stein, B.E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology, 56(3), 640-662.

Millikan, R. (2000). *On clear and confused ideas*. Cambridge: Cambridge University Press.

Millikan, R.G. (1991). Perceptual content and Fregean myth. *Mind*, 100(4), 439-459.

Milner, A.D. & Goodale, M.A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press

Nudds, M. [2014]: 'Is Audio-Visual Perception Cross-Modal or Amodal?', in D. Stokes, M. Matthen and S. Biggs (*eds*) *Perception and Its Modalities*. Oxford: Oxford University Press, pp. 166–88.

O'Callaghan, C. (2016). Objects for multisensory perception. *Philosophical Studies*, 173: 1269–1289.

Peacocke, C. (1986). Analogue Content. *Proceedings of the Aristotelian Society Supplementary Volume*, 60: 1-17.

Rizzolatti, G., Scandolara, C., Matelli, M. & Gentilucci, M. (1981a). Afferent properties of periarcuate neurons in macaque monkeys. I. Somatosensory responses. Behavioural Brain Research, 2, 125-146.

Rizzolatti, G., Scandolara, C., Matelli, M. & Gentilucci, M. (1981b). Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. Behavioural Brain Research, 2, 147-163.

Schneider, G.E. (1969). Two Visual Systems. *Science*, 163(3870), 895-902.

Thurlow, W. R. & Jack, C. E. (1973). A study of certain determinants of the “ventriloquism effect”. *Perceptual and Motor Skills, 36:* 1171-1184.

Ungerleider, L.G. and Mishkin, M. (1982). Two Cortical Visual Systems, in D.J. Ingle, M.A. Goodale, and R.J.W. Mansfield (Eds.) *Analysis of Visual Behavior,* Cambridge, MA: Massachusetts Institute of Technology Press.

de Vignemont, F. (2018). Peripersonal perception in action. *Synthese*.

Witkin, H. A., Wapner, S., & Leventhal, T. (1952). Sound localization with conflicting visual and auditory cues. *Journal of Experimental Psychology, 43:* 58-67.