

# Qualitative Research

<http://qrj.sagepub.com/>

---

## **Multimodality, multisensoriality and ethnographic knowing: social semiotics and the phenomenology of perception**

Sarah Pink

*Qualitative Research* 2011 11: 261

DOI: 10.1177/1468794111399835

The online version of this article can be found at:

<http://qrj.sagepub.com/content/11/3/261>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Qualitative Research* can be found at:**

**Email Alerts:** <http://qrj.sagepub.com/cgi/alerts>

**Subscriptions:** <http://qrj.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://qrj.sagepub.com/content/11/3/261.refs.html>

---

# Multimodality, multisensoriality and ethnographic knowing: social semiotics and the phenomenology of perception

Qualitative Research

11(3) 261–276

© The Author(s) 2011

Reprints and permission: sagepub.

co.uk/journalsPermissions.nav

DOI: 10.1177/1468794111399835

qrj.sagepub.com



**Sarah Pink**

Loughborough University, UK

## Abstract

In this article I consider the potential for combining multimodality and anthropologically informed sensory ethnographic methodologies. I focus on a comparison between anthropological and multimodality approaches to the senses, the relationships between images and words, and ethnography. In doing so I reveal some of the tensions and fundamental differences between these approaches before then considering if and/or how these might be reconciled.

## Keywords

multimodality, phenomenological anthropology, sensory ethnography, visual methods

## Introduction

In this article, building on arguments developed in my recent book *Doing Sensory Ethnography* (Pink, 2009), I consider the potential for combining multimodality and anthropologically informed sensory ethnographic methodologies. I focus on a comparison between anthropological and multimodality approaches to the senses, the relationships between images and words, and ethnography. In doing so I reveal some of the tensions and fundamental differences between these approaches before then considering if and/or how these might be reconciled. Before continuing I make two disclaimers. First, anthropology is a broad discipline, the practitioners of which cannot be said to ascribe to a single approach. The branch of anthropological thought I discuss here is a phenomenological anthropology influenced by the work of Tim Ingold (e.g. 2000, 2008). This anthropological approach is particularly interesting to discuss in relation to multimodality scholarship precisely because it is founded on such different philosophical principles. However, that is not my sole reason for selecting it, for it also offers a very viable way of understanding

---

## Corresponding author:

Sarah Pink, Department of Social Sciences, Loughborough University, LE11 3TU, UK

Email: s.pink@lboro.ac.uk

human experience, knowledge and practice. Second, I acknowledge that my engagement with literature about multimodality scholarship is restricted, rather than attempting to be a full review of the field. My intention here has been to engage specifically with both influential works in this area and recent work that incorporates multimedia. My aim is to thus focus the discussion on how these texts deal with questions of the senses and (audio) visual media, because these are key points of contact with the conceptualization of a sensory ethnography. While these two points delimit the intellectual exercise of the article, they do, however, permit the making of a valid comparison.

To develop this discussion I reflect on areas over which anthropologists and multimodality scholars coincide in their interests, if not in their critical perspectives or theoretical and methodological commitments. First, given that it is my aim to investigate how a sensory ethnography might be combined with multimodality scholarship, I discuss recent approaches to the senses – sensory categories and sensory perception – in anthropology and by multimodality scholars. Because for multimodality scholars the senses are associated with what are referred to as ‘modes’, it is necessary to consider what these mean in relation to the related concept of ‘media’. Therefore, following from this, and because sensory ethnography often involves the use of multiple media, I trace continuities between the way the senses have been understood in each discipline and the concomitant treatments of the relationship between writing and images. Having established the fundamental principles that inform these approaches, I turn the discussion to explore how these inform the ways ethnographic methodologies are conceptualized in each discipline. It will be apparent throughout this discussion that the two approaches are based on fundamentally different theoretical premises and methodological approaches. Yet it is the brief of this special issue to consider how they might usefully work together. Thus, finally I will consider how such different types of academic knowing might be understood in relation to each other.

## Multimodality, the senses, words and images

Both social semioticians interested in multimodality and social anthropologists have come to acknowledge that an understanding of (multi)sensoriality is essential to understanding aspects of society and culture. However, their understandings of both the relationships between the senses and the way a ‘sensory’ approach might be mobilized as a methodology are derived from quite different starting points. Such understandings also inform how scholars from each discipline can correspondingly interpret how texts communicate. First, I outline how these questions have been treated in the multimodality literature. I then cast a critical perspective on the assumptions that underlay this interpretation based in phenomenological anthropology.

Multimodality scholars tend to understand communication on two levels, and as happening through the relationship between what they call ‘modes’ and ‘media’, both of which come in diverse forms. As Gunter Kress states it:

I use the term ‘mode’ for the culturally and socially produced resources for representation and ‘medium’ as the term for the culturally produced means for distribution of these representations-as-meanings, that is, as messages. (Kress, 2005: 6–7)

If we think of this in terms of the actual production of texts then, for example writing is a mode, while a book is a medium (see Kress, 2005: 7). Other examples of things that might be considered modes are given by Bella Dicks and her co-authors who write: 'obvious ones include writing, speech and images; less obvious ones include gesture, facial expression, texture, size and shape, even colour' (2006: 82). Dicks et al. (2006) mention texture, which, while it does not refer to a sensory modality as it would be defined in the anthropology of the senses (it would be touch), does refer to a quality of an object that a modern western subject could understand through a conventional sensory category. Therefore, on the basis of this I am assuming that modern western sensory categories come into play in studies of multimodality as informing the way in which 'modes' (e.g. texture) might be perceived (e.g. through touch). Indeed, the senses have never been far from thinking about multi-modality and (as I also outline in *Doing Sensory Ethnography* [Pink, 2009]), have been acknowledged in some of the founding work of Kress and Theo van Leeuwen in this area. In Kress's definition of multimodality, a specific understanding of the senses as differentiated informs his understanding of human perception. He sees 'sight, hearing, smell, taste and feel' as 'each being attuned in a quite specific way to the natural environment, providing us with highly differentiated information' (2000: 184). More explicitly Kress and van Leeuwen propose that 'the sense of sight gives access to the world differently from the senses of touch, smell, taste' (2001: 127). Thus, according to vision a special status. Within this paradigm there is also some recognition of a relationship between the senses, in that Kress suggests that 'none of the senses ever operates in isolation from the others' and that this 'guarantees the multimodality of our semiotic world' (2000: 184). Thus, from these points I would deduce that it is our assumed ability to perceive the world around us – and as such the modes of communication that produce meanings/representations in the form of media – through the five (differentiated) senses that is pivotal for multimodality scholars. In that it facilitates a connection between how we perceive the environment as different forms of sensory 'information', how these understandings as sensory categories are thus understood as communicative gestures, textures, smells and so on, and thus how they are subsequently represented in textual forms and human performances (or at least forms that might be read *as if text*). While (to my knowledge) the existing multimodality literature does not specify exactly how this process works, a comment made by Dicks et al., when writing about the qualities of ethnographic photographs gives a clue:

Photographs allow us to see modes that are visual: colour, shape, size, position, light. What they do not show us are modes that operate through the other senses – of touch, smell, hearing and taste – such as bodily movement, texture, three-dimensional shape, sounds. (Dicks et al., 2006: 88)

The above passage is therefore illustrative of the dependence that the multimodality approach has on the five-sense model. Different modes can be associated with or seen to pertain to particular categories of sensory information, or sensory channels (I will return to this below). However, it also leads into the next area for discussion – the question of how multimodality scholars understand images and writing and their potential for communication.

The tendency to separate out different modes and media of communication is not only reflected in the way the senses are understood as each 'providing highly differentiated information' (Kress, 2000: 184), but also in Kress's (2005) suggestion that we should draw a distinction between the way writing and images communicate. McDonagh et al. (2005) draw (critical) attention to two passages where Kress outlines his perspective:

Because words rely on convention and on conventional acceptance, words are always general, and therefore vague. Words being nearly empty of meaning need filling with the hearer/reader's meaning. (Kress, 2005: 15)

And

... Unlike words, depictions are full of meaning; they are always specific. So on the one hand there is a finite stock of words – vague, general, nearly empty of meaning; on the other hand there is an infinitely large potential of depictions – precise, specific, and full of meaning, (Kress, 2005: 15–16)

Kress proposes what I find a rather surprising distinction between words and depictions. However, before discussing this further I first draw a series of comparisons between anthropological and multimodal understandings of the senses.

In the remainder of this section I outline how both the senses and the relationship between writing and images are understood rather differently from an anthropological perspective. The discussion is based in the anthropology of the senses, which as a subdiscipline was established around the early 1990s, and owes much to the work of founding scholars including especially David Howes (e.g. 1991, 2003) and Paul Stoller (e.g. 1989, 1997), both of whom remain leading figures in this area. While there is some disagreement among anthropologists who have written on the senses concerning the aims and approaches that a sensory anthropology should take (these are discussed in detail in Pink, 2009), it is fair to make two points that will inform the discussion here. First, although they might not agree precisely on how human perception should be theorized, perception is central to the work on anthropologists of the senses (see for example, Feld, 2005; Howes, 2005; Ingold, 2000). Second, there is a common recognition that the modern western five-sense sensorium is a cultural construct (see Geurts, 2003; Howes, 2005). My presentation of recent anthropological thinking is admittedly very dependent on a critical approach to the anthropology of the senses developed by Ingold (see especially, Ingold, 2000), which is increasingly influential among scholars in anthropology and other disciplines who attend to the senses (as outlined in Pink, 2009). This is in part because I find it a useful and convincing paradigm. Yet this is not the only reason why it is applicable here, since it also allows us to consider the role of the concept of 'affordances' in each approach.

The first fundamental difference between understandings of the senses in multimodality scholarship and the anthropology of the senses focuses on the question of the five-sense modern western sensorium. Anthropologists of the senses now commonly recognize that the notion that there are five senses attached to five sense organs is a modern western construct, not necessarily applicable in other cultures (see, Geurts, 2003; Howes, 1991, 2005). If this is the case, then it would seem problematic to use this construct as a basis upon which to build a universal theory about how communication happens. Indeed, if the

five-sense sensorium is a cultural construct, and entails a set of categories that are used by modern western subjects as ways of ordering their world, rather than being a universal truth that can be applied to any context, this renders it as an object for analysis itself. In this sense it would be inappropriate to treat the five sense sensorium as representing a set of universal analytical units or categories, unless this is done with the recognition that these are not 'natural' categories or the substance for the establishment of universal truths (or, as such, universal theory building) but rather they are convenient units through which to analyze sets of experiences (see also Pink [2009] where their utility as analytical categories is discussed). When working with modern western subjects who use the same categories then this might be non-problematic. Yet, as cross-cultural studies have revealed, they are not always applicable, as I discuss elsewhere (Pink, 2006) this is shown especially well in the work of Kathryn Linn Geurts, through her ethnographic research in Ghana (e.g. Geurts, 2003). Geurts, a North American anthropologist worked with the Anlo Ewe people of Ghana. In seeking to understand their sensory embodied experiences, she found a lack of direct correspondence between the categories of vision, smell, taste, sound and touch that she would use to describe sensory experience and those of the participants in her research. In contrast, she identifies (and develops a discussion of) a series of interrelated categories that Anlo Ewe people employ. Some of these, such as audition, balance, kinaesthesia and synaesthesia do not map directly onto the modern western five-sense sensorium, and moreover, those senses we might more easily identify with, such as seeing, tasting, olfaction and hearing are not necessarily constructed as separate from each other (see Geurts, 2003: 37–69 for a detailed discussion of these categories). Moreover, as recent research in modern western contexts has shown, people often employ additional sensory categories to refer to sensory experiences that cannot be accommodated by the five-sense model. For example, my own research about the home showed how the category of 'freshness' is used to explain some elements of the home (e.g. Pink, 2009) and Åsa Bäckström has shown how Swedish skateboarders also surpass categories offered by the five-sense sensorium when they describe their embodied experiences (Bäckström, 2009). The implication of these ethnographic findings for multimodal analysis is that they question assumptions about human perception that divide modes up as functioning through specific sensory routes. These 'extra' sensory categories enable participants in research to talk about sensory experience without having to allocate an experience exclusively to one sensory category. So, for example, when Dicks et al. (2006: 88) suggest that photographs only permit us to 'see modes that are visual' they do not consider that vision might *not* be an exclusive category of experience. This leads me to my next point.

Anthropologists of the senses do not necessarily understand the senses as differentiated and simply operating in relation to each other. Rather, they account for the relationship between the phenomenology of perception and the cultural constructedness of sensory categories. These ideas draw on Maurice Merleau-Ponty's work on the phenomenology of perception, and have been appropriated by a number of anthropologists working in his field. As Geurts puts it, drawing on Csordas's (1990) work on embodiment:

Merleau-Ponty's phenomenology involves a rejection of the empiricist model that suggests external objects stimulate our internal organs such that we register sensory data and instead embraces the idea that perception begins in the body and ends in objects (Csordas, 1990: 8–9). (Geurts, 2003: 74)

Indeed, the cultural categories represented in the modern western and other sensoria are only culturally constructed resources that we use to enable us to communicate about human sensory perception. Ingold's contribution in this area of thought is particularly important for the discussion here. There is no space here to go into the detail of the trajectory of his argument (see Pink [2009] where this is discussed in more detail), but it can be summed up. Ingold also draws on the work of the philosopher Merleau-Ponty. Of particular relevance for the discussion here, is his reference to the argument that the body is not 'a collection of adjacent organs but a synergic system, all of the functions of which are exercised and linked together in the general action of being in the world' (Merleau-Ponty, 2002 [1962]: 234). Therefore, writing specifically about the relationship between sound and vision, Ingold proposes that 'the eyes and ears should not be understood as separate keyboards for the registration of sensation but as organs of the body as a whole, in whose movement, within an environment, the activity of perception consists' (2000: 268). He also builds on similar points made by the ecological psychologist James Gibson, writing that '... the perceptual systems not only overlap in their functions, but are subsumed under a total system of bodily orientation ... Looking, listening and touching, therefore, are not separate activities, they are just different facets of the same activity: that of the whole organism in its environment' (Ingold, 2000: 261).

This invites us to re-think Kress's notion of 'sight, hearing, smell, taste and feel' as 'each being attuned in a quite specific way to the natural environment, proving us with highly differentiated information' (2000: 184). Instead, we might take as a starting point the idea that a rather less culturally structured flow of neurological information becomes differentiated into categories that we call the senses. As such it is not so much that 'none of the senses ever operates in isolation from the others' and that this 'guarantees the multimodality of our semiotic world' (Kress, 2000: 184). Rather, it is that we tend to communicate linguistically about our embodied and sensory perception in terms of sensory categories. However, because one category is never enough to express exactly what we have actually experienced, the illusion of the 'separate' senses operating *in relation to* each other is maintained. An approach that sees the senses as interconnected, rather than differentiated, also challenges Dicks et al.'s (2006) understanding of the limits of photography. Whereas they suggest that photographs cannot tell us about 'modes that operate through the other senses – of touch, smell, hearing and taste – such as bodily movement, texture, three-dimensional shape, sounds' (p. 88), scholars working in visual anthropology and film theory regard the image as having great potential for representing/evoking other sensory experiences. In particular (as I outline in Pink, 2009) the phenomenological writings of the anthropological filmmaker David MacDougall (1998, 2005) and the notion of 'haptic cinema' advanced by the film scholar Laura Marks (2000) demonstrate the possibilities of communicating about tactile experience, and texture in film. However, working with these ideas requires researcher engagements that go beyond observation and data collection to attend to the ways in which we might reflexively draw on our own existing biographical experiences (as researchers and film viewers) in order to imagine and recognize our sensory embodied responses to other people, objects, textures and more in film and video. Therefore, for example, if I was (re)viewing a video of a research participant running hot water through her fingers, into a foaming bowl of detergent for washing up, I would not be simply interested in what I could observe and how this relates to different modalities. Rather I would seek to use my

own prior experiences to comprehend the skilled sensory process that she is engaging in by monitoring the temperature of the water, the feel of the foaming water, the soapiness of the cloth and the cleanliness of the dishes. I would be therefore trying to understand her ways of knowing and regulating this process – ways of knowing in which the senses remain interconnected. I would also be interested in how she would later interpret these experiences in terms of different sensory modalities. Yet this would not be so that I could understand how separate actions of touching and seeing combine to make meaning, but, rather thinking about this the other way round, I would be interested in how she divided this experience up into the categories of what is known by seeing and what is known by feeling because this would help me to understand how she gives culturally constructed meaning to these activities.

I now turn to Kress's distinction between words and depictions. As I noted above, Kress's insistence on the specificity of images and the precision of their meaning seems out of tune with arguments made in several disciplines. For instance in a critical response to Kress's (2005) essay, McDonagh et al. (2005: 85) point out that 'to accept Kress' argument that images are inherently filled with precise meaning would require that we ignore most practice in visual arts and design, nearly all of which is grounded firmly in the realm of the symbolic'. Kress's understand of visual meanings is also opposed to that advanced by scholars working in the area of (feminist) art therapy, where the meanings of images are seen as being shifting and contingent. Indeed, this malleability of visual meanings is seen as essential to the therapeutic process (Hogan and Pink, 2010). The distinction Kress draws is also quite the opposite to that proposed in earlier works by visual anthropologists, who have emphasized the subjectivity of filmic interpretation (see Pink [2007a] for a discussion of the relationship between film and text in visual methods). Given these existing strands of interdisciplinary scholarship that have for some time emphasized the contingency of the meanings of images (and material objects) Kress's assertion seems unfounded. Indeed, as I have argued elsewhere (Pink, forthcoming), and following Ingold's work on the commonalities between writing and drawing (Ingold, 2007) it seems more profitable to seek to explore the synergies and connections between writing and images than construct them in terms of binary oppositions.

The differences between these approaches may also be understood through a focus on their differential conceptualizations of the notion of 'affordances'. As noted above, Ingold's work is informed by Gibson's ecological psychology. In particular Gibson's notion of affordances contributes to the way Ingold formulates human perception. Central to this, for Ingold, is Gibson's 'insight that the information picked up by an agent in the context of practical activity specifies what are called the "affordances" of objects and events in the environment (Gibson, 1979: 127–43)' (Ingold, 2000: 166). According to this perspective, thus, affordances are not fixed in objects or events to be perceived uniformly. Rather they are determined through the nature of the 'action in which the perceiver is currently engaged' (Ingold, 2000: 166). Paul Prior's (2005) critical essay offers a way to think about the differences between Ingold's and Kress's approach through the notion of 'affordances'. Prior identifies '... Kress' treatment of "affordances" as highly determinative, mutually exclusive, and binary'. By contrast, he points out that 'Gibson's (1979) basic notion of affordances was, in fact, intended to avoid turning objective properties of things into such hard categories' and 'stressed that affordances are relational, ecological, and tendential (not determinative)' and recognized



the ‘fuzziness of categories’. Kress’s approach, Prior writes, is very different in that it ‘proposes a set of hard binary distinctions between words and images. Words in his account are finite, sequential, vague, conventional, authored, narrative and/or causal, and open to critique. Images are infinite, spatial, specific, natural and transparent, viewed, and available only for design’ (Prior, 2005: 26).

Phenomenological anthropology appreciates the importance of human perception and experience and seeks to understand this as inextricable from the constitution of cultural categories and meanings. Thus understanding ‘affordances’ as contingent on perceptual processes rather than as providing predetermined pre-categorized sense-data. Multimodality is in contrast built on an understanding of culture as a set of ‘readable’ representations that can neatly be placed in mutually exclusive categories with their own characteristics, and that are perceived through differentiated channels of sensory information.

The discussion in this section has made it clear that there are certain fundamental incompatibilities between these two approaches. Phenomenological anthropology builds on the phenomenology of perception and ecological psychology to understand the senses as not simply interconnected, but as part of a system in which they are not so easily distinguishable. While they are spoken about in culturally constructed categories (which indeed vary from culture to culture) human perception is not divided up into these categories, neither are the modern western sensory modalities of vision, sound, touch, smell and taste exclusively attached to the sensory organs we would conventionally associate them with. Similarly, the binaries made between writing and images in the multimodality paradigm are not compatible with either the notion of affordances that informs anthropological understandings, or with the understandings that cut across visual anthropology and a number of other disciplines that suggest that images only become meaningful in the context of their viewing, and as such do not ‘carry’ precise or universal meanings that can be read from them.

## Multimodality and ethnography

The contrasting foundations of phenomenological anthropology and multimodality outlined above have consequences for how scholars with commitments to either approach can *do* and understand ethnography. I now explore the implications of this.

There is a vast literature that seeks to define ethnography and how it is ‘done’; however, for the purposes of this discussion two recent approaches stand out:

Paul Atkinson (2005) has argued for an approach to ethnography that surpasses what he sees as a recent fragmentation of qualitative research. He writes that:

. . . despite classic ethnographic appeals to holism, context and similar ideas, qualitative, ethnographic research seems to have become increasingly fragmented. As the methodological literature has expanded, it has also diversified. Different authors adopt and promote specific approaches to the collection and analysis of data. Equally, particular kinds of data become celebrated in the process: personal narratives, life-histories and other documents of life; film, video and photographic images; texts and documentary sources; material culture and technological artefacts; spoken discourse. In the process types of data and corresponding types of analysis are elevated to occupy a special status. (Atkinson, 2005: paragraph 2)

Atkinson is arguing for a 'return' to classic ethnography. Yet he does not provide convincing evidence of exactly how recent authors have promoted the fragmentation he claims. For example, he writes that 'The collection and analysis of visual materials tend, unhelpfully, to be treated as a specialist domain' citing my own work in this area among other culprits. Yet, his claim is unsubstantiated: visual scholars have generally not sought to create an approach to ethnography that emphasizes the visual above other elements, but rather one that acknowledges its place in relation to them (see Pink 2007a). Rather, I believe that the difference between Atkinson's approach to ethnography and the new and innovative methodologies that he sees as a sign of fragmentation lies elsewhere. It is instead concerned with two rather different understandings of how culture, experience and meaning might be accessed and analyzed. My reason for making this important distinction here is because it is also essential to understanding the relationship of multimodality to different approaches in ethnography.

Atkinson argues that 'we need to retain a structural, formal sense of the multiple orderings of talk, action, things, places and so on' (Atkinson, 2005: paragraph 19) and following this calls for an ethnography that attends to Clifford Geertz's (1973) notion of 'thick description'. He argues that:

More sophisticated versions refer to the over-determined character of culture, with multiple frames of reference and perspective. My own gloss is to suggest that whatever else 'thick description' could mean, it should include systematic reference to the multiple forms of cultural life, producing cultural descriptions that preserve those distinctive forms. It thus takes GEERTZ's 'textual' approach to cultural analysis seriously, by insisting that the 'texts' need to be analysed in terms of their material and conventional properties. It also transforms the emphasis on 'culture' into an equal stress on social action. (Atkinson 2005: paragraph 21)

In following the 'cultural' approach Atkinson is advocating an ethnography that attends to many media of representation and to the analysis of visible social actions. Yet in my interpretation the analysis remains at the 'observable' surface. It is the analysis of things (e.g. objects, discourses, social actions and performances) that can be seen, heard, touched, tasted or smelled, and their characteristics, and the aim is at least in part, to reveal their cultural meanings – that is, the meanings that are thought to *reside in them*. This approach to ethnography is compatible with the multimodality paradigm as outlined by Dicks et al. (2006). Dicks et al. understand multimodal ethnography as 'a new multi-semiotic form in which meaning is produced through the inter-relationships between and among different media and modes' (2006: 78). They suggest that the notion of multimodality, along with the distinction between modes and media is useful for understanding the layers of representation that the ethnographer encounters in her or his (visual and other) data because:

What we actually observe in the field are the various media in which these modes are produced – marks on the page, movements of the body, sounds of voices, pictures on the wall. (Dicks et al., 2006: 82)

Again, the emphasis is on ethnography as an observational practice. I am not taking this to mean that either Atkinson (2005) or Dicks et al. (2006) consider the observational to

refer solely to what can be *seen* visually. Rather, my point is that their treatment of ethnography is one that involves the collation/production of data based on what can be observed and recorded as 'naturalistic' behavior. This might include note taking, video recording, sound recording, photography and more. The recordings might represent social action, discourses, material configurations, textures and other things that they would define as 'modes'. Dicks et al. also see the medium being decisive in the production of meaning in that they are concerned with 'how different modes are transformed when translated into different media. What semiotic modes do we lose when we use the camera? What meaning potential does speech afford that is closed off by images or writing? By contrast, what common semiotic affordances are produced by both film and writing?' (2006: 93–94). In this formulation thus affordances come prior to perception (as for Kress, 2005), rather than through the action of the perceiver in an environment (as for Ingold, 2000).

Above I have discussed the approach to ethnography that informs a multi-modal ethnographic methodology. I have argued that this is based on a 'classic' (Atkinson, 2005) understanding of ethnography, that is observational in principle and follows a Geertzian (1973) treatment of culture as 'readable' as text because it sees culture as discernable through an analysis of how it is *represented* in social action, material forms and more. Scholars who are developing a multimodal ethnography therefore seek to analyze the aspects of culture and meaning that can be 'observed' and recorded as data. This means recording the media through which the abstract modes are realized. Data then, which are produced through the observation of 'naturalistic' behaviour are understood as thus *representing* the media through which these modes of 'meaning making' (Dicks et al., 2006: 82) they seek to study can be detected. This approach to ethnography is in some ways contemporary in that it recognizes the need for researcher reflexivity, and seeks to understand how meanings are constituted relationally. Yet, its dependence on an observational and culture-as-readable-text approach to ethnography, and its emphasis on the *differences* between different senses, different modes and different media is based on and indeed entails a separating out of the world into sets of discrete components that work necessarily in relation to each other. To identify these involves *looking at* (i.e. observing) the world and the actions of people in it.

By contrast, the second approach to ethnography I outline here would involve *learning in and as part of* the world, and seeking routes through which to *share or imaginatively empathize with* the actions of people in it. It is also congruent with the phenomenological approach to understanding sensory perception as outlined above. If we follow Ingold's definition of anthropology as being a practice like art in that '[b]oth are ways of knowing that proceed along the observational paths of being *with*' (Ingold, 2008: 87), this might also be seen as a specifically anthropological approach to ethnography. While Ingold is keen to insist that 'anthropology is *not* ethnography' (2008), here I appropriate his idea to outline what I see as an anthropological type of sensory ethnography. In doing so I suggest that the anthropology of 'being with' can be likened to the idea of ethnography as a practice that seeks routes to understanding the experiences and meanings of other people's lives through different variations of being with, and doing things with them. The role of the ethnographer as apprentice has long since been important in anthropological practice (see Downey, 2007; Grasseni, 2007; Marchand,

2007; Pink, 2009) and certainly offers routes to an appreciation of how others experience. Recently closer attention has focused on questions of how learning happens in apprenticeship (e.g. Marchand, 2010). In such work it becomes clear that there are certain forms of knowledge that cannot be understood simply through observation. For instance, through a discussion of his work on Brazilian capoeira Greg Downey highlights the roles of unconscious learning and how in the learning of a physical skill the body also undergoes biological changes (Downey, 2010) and Tim Rice discusses the process of 'learning to listen' through a stethoscope during his auditory ethnographic apprenticeship research in a hospital (Rice, 2010). These ethnographers engaged with the practices that they wished to learn (about) beyond observation by becoming (to different extents and in different ways) apprentices in those sensory embodied skills. On the one hand here is a shift between *looking at* and collecting data on to *being in* and engaging in ways of knowing about the worlds and actions of other people.

Yet apprenticeship is not the only way that closeness can be sought. Thus it does not have to mean that the researcher is involved actively in or actually learning the *same* practice that she or he is seeking to research. Instead, being with others can involve interviewing, walking with them, discussing images with them, listening to music or other sounds with them and more. The practice of sensory ethnography involves the researchers' empathetic engagement with the practices and places that are important to the people participating in the research. And by association it does *not* therefore principally involve the collection of data *about them* that can later be analyzed. Rather it involves the production of meaning *in participation with them* through a shared activity in a shared place. For example, in my own research this has included the development of video tour and video (re)enactment methods, whereby the researcher and participant actively collaborate in ways that enable the participant to 'show' their material and sensory environment and to demonstrate and discuss the experience of performing everyday practices (see Pink, 2007a, 2007b). Therefore, in *Doing Sensory Ethnography* (Pink, 2009) I outline an approach to ethnography that is informed by the anthropology of knowledge (e.g. Harris, 2007) and the anthropology of the senses. Here I summarize selected key points from that discussion through a focus on the idea of researching the practices of everyday life. The practices I focus on coincide with some of those identified by Michel de Certeau in his (1984) *The Practices of Everyday Life*, in that I discuss talking and walking. These are near universal human activities, which although they might have different meanings in different contexts, are often shared between ethnographer and research participant.

Talking has traditionally been the mainstay of the 'ethnographic interview'. In ethnography, interviews might range from the form of a more casual conversation to sitting down with an audio recorder to discuss specific issues in a focused way. Whatever the context, I understand the interview less as a data collecting exercise than as a shared conversation through which new ways of knowing are produced. However, for the method I discuss here talking is combined with walking and image making with research participants. This is a method I have used in my own work and discussed at length in several other publications (e.g. Pink, 2007b, 2009), and that is increasingly used by researchers in various forms and across academic disciplines (Irving, 2007, 2010; Myers, 2010; O'Neill and Hubbard, 2010). I mention it here, at the risk of being repetitive,

because it brings to the fore two key issues. First, the idea of ethnography as producing knowledge with others, in movement and through engagement with/in a material, sensory and social environment. Second, the use of (audio)visual recording as a way of representing elements of this experience and the memories and imaginaries related to it. This involves uses of video and photography rather different to those proposed by Dicks et al. (2006) whose work I have discussed above.

Synergies between walking and anthropological fieldwork practice have been identified by Lee and Ingold (2006), and anthropological attention to walking has increased significantly in recent years (e.g. Ingold and Lee Vergunst, 2008; Irving, 2010; Lund, 2006). Understandings of walking with others as ways of knowing with them represented in these works are also congruent with the idea of learning and knowing as something that happens in movement. In *Doing Sensory Ethnography* I draw from Mark Harris's (2007) work to stress the situatedness of knowing and that neither humans nor knowing are static. As Harris puts it, someone would not '*stop* in order to know: she continues' (2007: 1 original italics). Therefore I have suggested that 'Knowing is continuous and processual, it is situated and it is bound up with human engagement, participation and movement (Harris, 2007: 4)' (Pink, 2009: 41). A focus on walking and talking thus offers us a basis from which to explore knowing in a range of contexts, one of which is the context of doing ethnographic research. If these practices also involve the use of video and/or photography then they also offer us the opportunity to consider the relationship between words and images in the ethnographic process.

In video and photographic tours, image making becomes a form of ethnographic note taking – rather than a way of visually recording data. It also involves producing images that are commonly used to represent the research experience. The anthropologist Andrew Irving has produced a series of accompanied walks during which the participants in his research have their narratives audio recorded and photographs are taken. In one article, Irving (2010) represents a walk that he shared in New York with Alberto – a man who had previously been diagnosed as HIV positive. The walk involved re-tracing Alberto's journey to and from the clinic on the day of his diagnosis. Irving audio recorded the spoken narrative as he describes this experience, and photographed key locations encountered during the walk. These photographs can be seen as being subjective framings of locations that are imbued with meaning for a specific person. These personal meanings are articulated verbally in the transcripts published alongside the sets of images. But neither the written narratives nor the images should be seen as data from which cultural meanings can be read. Rather they participate in the evocation of the sensory and affective dimensions of location as experienced through the subjectivity of the research participant and as brought into a public domain through the mutual engagement of the researcher and participant. As Irving expresses it:

I try not to pre-figure Alberto's experience or over-interpret his narrative by providing too much information or explaining the exact relationship between the images and the text. Instead we rely upon Alberto's own words and the reader's imagination to understand how other 'people immediately experience space and time, and the world in which they live' (Jackson, 1996: 12). (Irving, 2010: 26)

In my own work I have used the method of 'walking with video' (Pink, 2007b, 2009) to explore material and sensory environments with research participants. By asking a research participant to guide one around a particular locality (in my work this has included homes, a garden and a town) that holds meaning for him/her, and in which he/she is engaged practically on a regular basis, enables the researcher to move *through* and be *in* and *part of* an environment with the participant. When viewing the subsequent video recording the researcher is thus re-experiencing a route through a material, sensory and meaningful world, as already seen through the viewfinder. This is rather different from the perspective of *looking at* and reading from video-as-data from which cultural meanings can be interpreted/read.

### **Making connections (by way of a conclusion)**

The sort of ethnography that is theoretically and methodologically coherent with the multimodality paradigm is a 'classic' ethnography. Therefore, it would not be fair to argue that multimodality and ethnography are incompatible per se. Indeed, the appeal of 'classic' 'Geertzian' approach to ethnography for multimedia scholars is clear since both are guided by semiotic principles. Yet, in attaching multimodality scholarship to the particular conceptualizations of culture, meaning and experience that inform the 'classic' approach, an important and more recent critical literature around ethnographic methodology is by-passed. I appreciate that it is only to be expected that if a scholar were seeking to combine ethnography with multimodality then the obvious choice would be to match this with an approach to ethnography that is theoretically compatible with the multimodality approach. Yet, in doing so, she or he would ignore a whole slice of the historical trajectory of critical debate in anthropology about ethnography and ethnographic knowing. This is one of the great limitations of interdisciplinary borrowings and critiques (see Pink, 2006), and of course I realize that I am equally in danger of having committed such an omission in my rendering of multimodality scholarship in this article.

It is interesting, however, that while multimodality scholars are making overtures to ethnography, there is no correspondingly great interest among anthropologists in the multimodality approach. The theoretical principles that informed Geertz's own work on perception, experience and meanings have, in anthropology, been critically debated (e.g. Throop, 2003; see Pink, 2006). Indeed, even in the 1980s, Geertz's (1986) understandings of experience were contested. Marilyn Strathern (2005) has proposed that we understand anthropology as a discipline as a 'community of critics' constantly re-shaping what is anthropological knowledge and theory thorough debate. When anthropological ideas are extracted from debate and appropriated beyond the discipline they can become increasingly less appealing to anthropologists for whom their meaning within disciplinary debate may have shifted. As I have shown above, an increasingly influential phenomenological anthropology, with an emphasis on experience, has developed new ways of understanding human perception, as well as a focus on the senses and affect. The 'innovative' ethnographic methods that are currently emerging have an emphasis on mobility, affect, empathy, and knowing. They have departed from an observational

approach to seek to find routes through which the ethnographer might engage with other people and their experienced realities in these ways. In doing so, they engage with multiple media and methods adapted to specific circumstances, persons and projects.

Surely, if multimodality scholars are seeking to attach their work to ethnography, such an endeavour should engage beyond a simple return (or recourse) to a classic approach. Would it not be more informative and exciting to engage with new conceptualizations of ethnography, ethnographic knowing and empathetic research practices?

To end, my conclusion, perhaps surprisingly is not to argue that multimodality and sensory ethnography are completely irreconcilable (although while writing this I did begin to wonder if it would be). Instead, my proposal is that if ethnography is to become a useful – and by useful I also mean active and *critical* – tool for multimodality scholars, then it has a dual role to play. First, ethnographic research can indeed enable a greater understanding of practices, experiences and more. Second, a sensory ethnography that challenges the pre-set categories of multimodal analysis and breaks down the binaries between image and text can surely also create a self-critical and reflexive strand within multimodal analysis.

## References

- Atkinson P (2005) Qualitative Research-Unity and Diversity [25 paragraphs]. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 6(3): Art. 26. Available at: <http://nbn-resolving.de/urn:nbn:de:0114-fqs0503261>
- Bäckström A (2009) Making sense of skateboarding – on sensory ethnography and affective embodiment. Paper presented at the Ethnography and Education Colloquium at the Faculty of Education, Gothenburg University.
- Csordas TJ (1990) Embodiment as a paradigm for anthropology. *Ethos* 18: 5–47.
- De Certeau M (1984) *The Practice of Everyday Life*. Berkeley and Los Angeles, CA, and London: University of California Press.
- Dicks B, Soyinka B and Coffey A (2006) Multimodal ethnography. *Qualitative Research* 6(1): 77–96.
- Downey G (2007) Seeing with a ‘sideways glance’: visuomotor ‘knowing’ and the plasticity of perception. In: Harris M (ed.) *Ways of Knowing, New Approaches in the Anthropology of Experience and Learning*. Oxford: Berghahn.
- Downey G (2010) ‘Practice without theory’: a neuroanthropological perspective on embodied learning. *Journal of the Royal Anthropological Institute* 16: S22–S40.
- Feld S (2005) Places sensed, senses placed: towards a sensuous epistemology of Environments. In: Howes D (ed.) *Empire of the Senses: The Sensory Culture Reader*. Oxford: Berg.
- Geertz C (1973) *The Interpretation of Cultures*. New York: Basic Books.
- Geertz C (1986) Making experience, authoring selves. In: Turner V and Bruner E (eds) *The Anthropology of Experience*. Urbana, IL: University of Illinois Press.
- Geurts KL (2003) *Culture and the Senses: Bodily Ways of Knowing in an African Community*. Berkeley, CA: University of California Press.
- Gibson J (1979) *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
- Grasseni (2007) Communities of practice and forms of life: towards a rehabilitation of vision. In: Harris M (ed.) *Ways of Knowing, New Approaches in the Anthropology of Experience and Learning*. Oxford: Berghahn.
- Harris M (2007) Introduction: ways of knowing. In: Harris M (ed.) *Ways of Knowing, New Approaches in the Anthropology of Experience and Learning*. Oxford: Berghahn.

- Hogan S and Pink S (2010) Routes to interiorities: art therapy, anthropology & knowing in anthropology. *Visual Anthropology Review* 23(2): 1–16.
- Howes D (ed.) (1991) *The Varieties of Sensory Experience: A Sourcebook in the Anthropology of the Senses*. Toronto: University of Toronto Press.
- Howes D (2003) *Sensing Culture: Engaging the Senses in Culture and Social Theory*. Ann Arbor, MI: The University of Michigan Press.
- Howes D (ed.) (2005) *Empire of the Senses: The Sensual Culture Reader*. Oxford: Berg.
- Ingold T (2000) *The Perception of the Environment*. London: Routledge.
- Ingold T (2007) *Lines: A Brief History*. London: Routledge.
- Ingold T (2008) Anthropology is not ethnography. *Proceedings of the British Academy* 154: 69–92.
- Ingold T and Lee Vergunst J (eds) (2008) *Ways of Walking: Ethnography and Practice on Foot*. Aldershot: Ashgate.
- Irving A (2007) Ethnography, art and death. *Journal of the Royal Anthropological Institute* 13(1): 185–208.
- Irving A (2010) Dangerous substances and visible evidence: tears, blood, alcohol, pills. In: Ingold T and Lee Vergunst J (eds) (2008) *Ways of Walking: Ethnography and Practice on Foot*. Aldershot: Ashgate.
- Kress G (2000) Multimodality. In: Cope B and Kalantzis N (eds) *Multiliteracies: Literacy Learning and the Design of Social Futures*. London: Routledge.
- Kress G (2005) Gains and losses: new forms of texts, knowledge, and learning. *Computers and Composition* 22(1): 5–22.
- Kress G and van Leeuwen T (2001) *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Cassell.
- Lee J and Ingold T (2006) Fieldwork on foot: perceiving, routing, socializing. In: Coleman S and Collins P (eds) *Locating the Field. Space, Place and Context in Anthropology*. Oxford: Berg, 67–86.
- Lund K (2006) Seeing in motion and the touching eye: walking over Scotland's mountains. In: Bendix R and Brenneis D (eds) *Etnofoor. Anthropological Journal* 18(1): 27–42.
- McDonagh D, Goggin N and Squier J (2005) Signs, symbols, and subjectivity: an alternative view of the visual. *Computers and Composition* 22(1): 79–86.
- MacDougall D (1998) *Transcultural Cinema*. Princeton, NJ: Princeton University Press.
- MacDougall D (2005) *The Corporeal Image: Film, Ethnography, and the Senses*. Princeton, NJ: Princeton University Press.
- Marchand T (2007) Crafting knowledge: the role of 'parsing and production' in the communication of skill-based knowledge among masons. In: Harris M (ed.) *Ways of Knowing, New Approaches in the Anthropology of Experience and Learning*. Oxford: Berghahn.
- Marchand TH (2010) Making knowledge: explorations of the indissoluble relation between minds, bodies, and environment. *Journal of the Royal Anthropological Institute* 16: S1–S21.
- Marks L (2000) *The Skin of the Film*. Durham, MD: Duke University Press.
- Merleau-Ponty M (2002 [1962]) *The Phenomenology of Perception*. London: Routledge.
- Myers M (2010) 'Walk with me, talk with me': the art of *conversive wayfinding*. *Visual Studies* 25(1): 59–68.
- O'Neill M and Hubbard P (2010) Walking, sensing, belonging: ethno-mimesis as performative praxis. *Visual Studies* 25(1): 46–59.
- Pink S (2006) *The Future of Visual Anthropology: Engaging the Senses*. Oxford: Routledge.
- Pink S (2007a) *Doing Visual Ethnography*. London: Sage.
- Pink S (2007b) Walking with video. *Visual Studies* 22(3): 240–242.
- Pink S (2009) *Doing Sensory Ethnography*. London: Sage.



- Pink S (forthcoming) Drawing with our feet (and trampling the maps): walking with video as a graphic anthropology. In: Ingold T (ed.) volume based on the *Redrawing Anthropology* conference, Ashgate.
- Prior P (2005) Moving multimodality beyond the binaries: a response to Gunther Kress' 'Gains and Losses'. *Computers and Composition* 22(1): 23–30.
- Rice T (2010) Learning to listen: auscultation and the transmission of auditory knowledge. *Journal of the Royal Anthropological Institute* 16: S41–S61.
- Stoller P (1989) *The Taste of Ethnographic Things: The Senses in Ethnography*. Philadelphia, PA: University of Philadelphia Press.
- Stoller P (1997) *Sensuous Scholarship*. Philadelphia, PA: University of Pennsylvania Press.
- Strathern M (2006) A community of critics? Thoughts on new knowledge. *The Journal of the Royal Anthropological Institute* 12(1): 191–209.
- Throop J (2003) Articulating experience. *Anthropological Theory* 3(2): 219–241.

### Biographical note

Sarah Pink is Professor of Social Sciences at Loughborough University, UK. Her research is usually interdisciplinary, involves the use of visual methods and media and where appropriate engages with non-academic audiences. Her books that are relevant to discussions about qualitative methodology include *The Future of Visual Anthropology* (2006), *Doing Visual Ethnography* (2007) and *Doing Sensory Ethnography* (2009).