Part I
Time and Space

1

Mediated Interactions

1.1 Introduction

On 31 December 1999 the BBC, in conjunction with 50 other broadcasters, launched what it described as a 'complex and ambitious project on a scale never before attempted in the history of television':[1] a 25-hour live broadcast during which more than 2,000 cameras around the world, operating across some 60 satellite paths, would follow the build-up to midnight country-by-country as the international dateline moved west across the globe and the year 2000 dawned in each location in turn. The aim of this project, repeatedly declared by the BBC over the course of its live coverage, was to deliver to its viewers nothing less than 'the world' in its entirety:

> In one revolution, in one day, we bring you the world. Wherever you are, whatever you're doing, the BBC is with you on *2000 Today*.

The *2000 Today* broadcast was by no means without its problems. Presenters tossed to correspondents around the world who were not on camera, could not be heard or were out of sync; live feeds failed to materialize, or were lost in mid-segment. The broadcast, furthermore, seemed destined to disenchant, given its goal of delivering the moment of midnight over and over again around the world. If midnight on New Year's Eve is a liminal moment, a transitional moment made meaningful for individuals through the particularities of their own time and place, their own cultural and personal context, then the serial manifestations of midnight in Kiribas, in Auckland, in New York, in Paris and elsewhere could only serve to depreciate the value of the viewer's own midnight; a risk which the presenters seemed to recognize in their repeated and persistent references over the course of the day to their 'own millennium midnight', the upcoming 'real' moment in the UK.[2] Viewed from this perspective, the subsequent comment in one British newspaper lambasting the 'wreckage of the BBC's millennium coverage' seems easy enough to comprehend.[3]

To view the *2000 Today* programme simply in terms of its limitations, though, may be to miss the point. As one of the anchors remarked, some twelve hours

---

[1] BBC Press Office and WGBH National Promotion, 6 October 1998.

[2] On liminality, see Turner (1969: 95), who likens liminality 'to death, to being in the womb, to invisibility, to darkness, to bisexuality, to the wilderness, and to an eclipse of the sun or moon'; see also Marriott (2001).

[3] *The Guardian*, 4 September 2000.

into the coverage, there was at least one perspective from which the programme could be regarded as extraordinary:

> When I first started television … I'm sure this would've been absolutely totally inconceivable; there's no way this could have happened at all. I mean, about 50 miles was as far as you could go, wasn't it, and now in a very short period of time you've got this absolute miracle … it's extraordinary to think you can actually encircle the globe like this electronically.[4]

The presenter Michael Parkinson's comments here make clear the nature of the rhetoric which underpinned the *2000 Today* broadcast: a celebration of what Carey (1989: 139) has referred to as the 'electronic sublime', the dizzying imaginings of the transformations in human communication which can be delivered by technologies of communication. Whilst *2000 Today*, in other words, may have run the risk of dispossessing the domestic television audience of their 'own millennium midnight' through the incessant reiteration of other midnights in other places, it had a different variety of enchantment to offer instead: the enchantment of 'the world', everywhere simultaneous and everywhere articulated through electronic communication. The liminal moment which television sought to capture in its transmission of this event was thus not simply the moment of midnight as it would manifest itself in the UK some hours into the live broadcast, but also the transition between one millennium of technological progress and the next, marked by the 25-hour flow of images and talk from a multiplicity of simultaneous elsewheres. Self-reflexively, repeatedly, television was celebrating the transformations it had wrought in human communicative potential.

In this book I will be examining the repercussions of these transformations, with a particular view to the shifts that the development of electronic forms of communication has brought about for our experience of time, space and interactivity. Several writers (e.g. Hjarvard, 1994; Stevenson, 1995) have noted that such transformations in our experience of time and space have received little in the way of detailed attention. It is this 'chronically under-researched area' (Stevenson, 1995: 114) which this book will investigate, via a consideration of the defining features of electronic communication, and of live television in particular.

The choice of live television as the subject of this book may appear to require a degree of justification, given the consequences of media convergence. At the time of writing, individual television programmes can be consumed via digital download, and live television can similarly be streamed to a computer or a mobile phone. Why, at such a point in time, would a book on television be a worthwhile expenditure of energy, for either the reader or the writer? The answer, in part, is that the book concerns itself with television as a medium, and not with the apparatus itself; the arguments therefore hold equally true whether we are considering the consumption of a live broadcast in a living room via a traditional television, on a computer screen in an office or on the street via any one of a number of personal media.

---

[4]Michael Parkinson, BBC1, 31 December 1999.

Why television, though? And why live television in particular? The analysis in this book will reveal that live television is quite extraordinarily complex when viewed from the perspective of the transformations that it brings about in our experience of time and space. Previous writings on broadcast communication have tended to focus on two significant aspects of the way in which both television and radio deliver their messages. One of these is the institutional context in which television and radio are produced (Scannell, 1991, 1996), which has as its consequence the organizational involvement in the communicative event of parties other than the immediate participants; the other is the production of broadcast talk for an overhearing audience (Heritage, 1985: 99; Heritage and Greatbatch, 1991: 96), and more specifically for what Scannell (1996: 174) refers to as 'anyone-as-someone', an audience which is typically unknown but which is nevertheless addressed in its particularity. As this notion of the overhearing audience makes clear, broadcast communication is produced primarily for a hearer who is absent from the place in which the talk is produced, and who may be distant in time as well; but the space-time complexities which arise from this distanciation of the audience from the site of the communicative event have rarely been discussed in detail. When compared to radio, television is particularly interesting from this point of view because of the visual context for the communication, shared by television performer and remote audience. This is neither trivial nor insignificant in its range of possible effects; where radio messages are dependent on and constituted by aural communication alone, television involves a range of additional extralinguistic contextual factors: the physical make-up of the place in which the talk is produced, the physical behaviour of the speaker, visual access to other places with their own aesthetic and spatial characteristics. With live television, furthermore, co-temporality – the sharing of the same present moment – exists as a perpetual ontological possibility, realized whenever the broadcast is 'live' in the full sense that the time of the event, the time of television creation and the time of transmission and reception are one and the same (Heath and Skirrow, 1977: 53). Taken together, these characteristics – the institutional character of the broadcast event; its production for a non-specific but momentarily particularized absent audience; the presence of a perceptual space which may be more or less shared between speaker and absent audience but which minimally involves some visual domain that is known by the performer to be available to the viewer; the potential for the instantaneous transmission of information across distance – define the context in which televisual communication is produced and received. Jointly and severally they give rise to a range of interesting phenomena, which it will be the task of this book to investigate.

The investigation in this book will centre around the work of mediation which live broadcasting carries out in its transformation of the stuff of the world into the material of the broadcast event. It is the mediation of the event which bestows upon it the spatial and temporal complexity that will be discussed in what follows. In realizing the event, live television restructures the world both in space (the relations between the place in which stuff happens, the place from which the event is spoken and the places in which it is received) and in time

(the temporal architecture of the event, in which the everywhere-simultaneous world is structured into convoluted and dynamic reconfigurations of past and present in the real time of the live broadcast).

Such restructurings, as I will argue in Chapter 2, do not bring about a change in the nature of the universe. Space does not shrink, when electronic communication permits absent participants to engage in immediate interactions with each other; nor does time run backwards, re-loop or reconfigure itself. Rather, the transformations of time, space and patterns of interactivity which will be examined in this book are shifts in our *experience* of the world, permitting us to encounter events from temporal and spatial perspectives other than those available to us in a face-to-face interaction.

It is this transformation of our encounter with the world which will be crucial to the arguments developed here, and as a consequence the book will adopt a phenomenological approach, with a focus on the way in which live television, through its real-time mediation and realization of the situation, performs the event as a particular kind of object of experience for its audience. Central to this endeavour will be the approach advocated by the phenomenological sociologist Alfred Schutz, insofar as the book will seek to 'bracket' (Heritage, 1984: 41; Schutz, 1962: 105) or temporarily put to one side taken-for-granted views of space, time and interactivity in order to construct a bottom-up analysis of the phenomena under investigation. The purpose and value of such 'laborious analyses' (Schutz, 1962: 100) should become clear as the book proceeds with its construction of a grounded underpinning for the thesis that it will be developing.

I will begin this task in the remainder of this chapter with a consideration of the differences between face-to-face and electronically mediated interactions, which will permit us to develop an account of the spatial characteristics of instantaneous communication across distance. This will be taken forward in Chapter 2 with a similar consideration of time.

1.2   The Face-to-Face Encounter

Mediated encounters – encounters in which some or all of the relevant participants are absent from the immediate situation, dealing with each other instead via an intervening medium – differ in significant ways from face-to-face interactions. In the face-to-face or 'canonical' situation (Lyons, 1977: 638), participants are in the same place, and can potentially see, hear, touch, smell and even taste the same objects. Avery and McCain (1986: 122) define seven sensory modalities, all of which can potentially come into play in a canonical encounter: audio verbal (hearing or listening to speech), audio non-verbal (hearing non-speech sounds), visual verbal (reading), visual pictorial (looking at images), olfactory (smelling), tactile (touching) and taste. The possibility of jointly accessing all of these modalities will depend, of course, on the size of the place in which individuals are co-present: at a large-scale event such as an open-air concert or festival, a person several hundred metres away from me will not be able to read what I can read (unless we are both reading it from a video screen) or smell what

I can smell (unless the aroma of fast-food is omnipresent in our surroundings). If we consider a more finite and bounded space such as a room, however, it becomes clear that objects in the immediate vicinity will offer themselves to the entire range of sensory modalities for all those co-present. A poster on the wall can potentially be seen and read by anyone in the room; speech, unless it is whispered, will be available to anyone who is able to listen; my companion's recent lawn-mowing will be available as both a visual and an olfactory stimulus (the flotsam and jetsam from the garden sticking to his or her shoes; the smell of freshly cut grass). Furthermore, as 'embodied subjects' (Merleau-Ponty, 1962: 318) – individuals who are necessarily in the world and of the world – the more of our senses that we can bring to bear upon an object, the more fully we can realize it. Take as an example the kind of optical point-of-view shots that occasionally turn up in films made for IMAX screens. The huge IMAX theatre at Niagara Falls in Canada, for instance, features a film on the history of the Falls in which a camera mounted on the prow of a boat permits spectators the illusion that waters are surging over their heads. This may induce a degree of claustrophobia or dizziness in the audience at the apparent immediacy of the experience; nevertheless, one need only step out of the cinema doors to encounter the Falls themselves and experience the difference between being 'at grips with the world' (Merleau-Ponty, 1962: 303) in all its sensorily experienced fullness on the one hand and the thinness of seeing and hearing a torrent rush towards one on a screen on the other. To be present in the immediate vicinity of an object, then, is to have the fullest possible sensory access to it.

Full co-existence with objects and individuals, furthermore, as Husserl argues, involves not only the kind of totalizing unity of the senses that Merleau-Ponty talks about but also the engagement of the 'kinesthetically functioning living body' (Husserl, 1970 [1936]: 106), with its experience of its own 'body position, movement, weight, muscular tension and suchlike' (Hammond et al., 1991: 159): an ineffable and irreducible sense of ourselves and of our own immediate embodied situatedness in space and time. The individual who comes blinking out of the movie theatre and into the sunlight at Niagara not only will hear and see the spectacle, taste and feel and smell the spray, but will be possessed as well by an ineluctable sense of their own *hereness* in the particular environment in which they find themselves and in which other objects and circumstances – the waterfall, the crowd, the sunshine – are co-present. In this context we can usefully draw on Evans' (1982: 167) argument that individuals cannot conceive of two separate places simultaneously as 'here'. Developing an elaborate conceit around the idea of a remotely controlled submarine on the seabed which is equipped with limbs and excavators, Evans asserts that even though the observer might, in such cases, be able to reference a remote object as 'this' or the place where it was picked up as 'here', this is by no means the same as actually conceiving of oneself as being there:

> The subject can *play* at being where the submarine is ('Here it's mucky'); he can *play* at having that mechanical contrivance for his body ('I'll pick up that rock'). But really *he* is (say) in the bowels of a ship on the surface of the water. This is not just one

view he can adopt if he likes; it is the view to which everything in his thinking points. (Evans, 1982: 166, original emphases)

Individuals, Evans continues, must conceive of themselves as actually being *somewhere*: we cannot at the same time perceive the world from two points of view (1982: 168). However caught up we may momentarily be in the stimuli offered by a situation remote from us, then – even in the case of immersive virtual environments – we nevertheless always persistently find ourselves to be in the location where we are watching, or listening, or speaking: our maximal commitment, both sensory and kinaesthetic, is to the place *where we are*.

This primary orientation to conditions of presence reveals itself most clearly in *deixis*, the linguistic system that permits us to indexically identify places, objects and individuals relative to our own situatedness in the world.[5] The most basic use of expressions such as 'here', 'there', 'this place', 'that place', for instance, is to index places in relation to a zero-point centred upon the speaker's own corporeal engagement with the world: 'here' is where I am at the moment in which I utter it, in 'this place'; 'there' is 'that place', where I am not.

Deixis, furthermore, in its most fundamental uses, appears to require not only presence but also *co-presence*. As Lyons remarks:

> There is much in the structure of languages that can only be explained on the assumption that they have developed for communication in face-to-face interaction. This is clearly so as far as deixis is concerned. Many utterances which would be readily interpretable in a canonical situation-of-utterance are subject to various kinds of ambiguity or indeterminacy if they are produced in a non-canonical situation … if the participants cannot see each other, or cannot see what the other can see.
>
> (Lyons, 1977: 638)

The use of deictic expressions, in other words, presupposes that the indexed element is intersubjectively available to all relevant participants; and there are a range of uses where only co-presence, apparently, will do for establishing the relevant common ground. An earthbound individual cannot felicitously refer to an object ten miles away as 'on my right', because the object will be out of visual range and so inaccessible to those co-present; but a plane pilot, producing a commentary for passengers, can advise them to look out of their windows and view an entire town ten miles away 'on the left'. Conversely, we can point to the moon, when it is visible, and expect any co-present individuals to understand us when we claim that 'no-one ever really landed *there*', but we cannot point up the motorway from London and refer to Glasgow as 'there'.[6] 'What does *this smell* remind you of?', addressed to a friend on the phone, is likewise assured of an infelicitous outcome, as is the question 'does my bum look big in *this*?' directed at an individual in the next room.[7]

---

[5] See Bühler, 1982; Fillmore, 1975, 1982; Levinson, 1983; Lyons, 1977; Rommetveit, 1968, 1973.

[6] Looser constraints attend the 'symbolic' use of deictic expressions; cf. Fillmore's example of a telephone conversation in which the speaker asks whether Johnny is 'there' (Fillmore, 1975).

[7] Although the latter may well have a particular speech act function, that of requesting that the hearer come into the room *where I am* to see which outfit I'm referring to.

So-called *gestural* uses of deictic expressions (Fillmore, 1975: 40), further-more, require more than simple co-presence. 'Does my bum look big in *this*?' may well be doubly infelicitous, requiring not only a co-present audience but one who is actively monitoring the speaker's indexical behaviour as she gestures in the general direction of an item of clothing or a part of her anatomy. All that would stand between blindfolded listeners and a disaster in such a situation would be whatever reserves of tactfulness – or irritation – they could muster. When deictics are used gesturally, being co-present isn't sufficient; relevant participants need to see what is being indicated if they are to make sense of what is said.

Co-presence is thus a fundamental condition for the optimal interpretation of a range of deictic uses of language. It is the shared vicinity, in all of these instances, which supplies the relevant common ground for identifying the indicated object, individual or event; and a *co-operative* speaker (Grice, 1975) will reserve their use of deictic expressions for those circumstances where other relevant participants will share the appropriate common ground, whether this be the locale-at-large or a particular demonstratively identified object.

The cardinal property of the canonical situation – the potential for all co-present individuals to jointly access the full range of sensory modalities on offer – thus has as one consequence the use of a range of linguistic expressions that permit participants to identify and locate elements in their vicinity – sounds, smells, sights, tastes – relative to the sensory and kinaesthetic zero-point which is *where they are*. Other behavioural and interactional features are similarly consequent; but whilst the gestural use of deictics involves the *physical* indication of some local element, other sets of cues, both non-verbal and verbal, operate at a more symbolic and expressive level. Participants who can see each other and – more to the point – who know themselves to be seen, can indicate via facial expressions what their attitude is to objects, events or other individuals in their immediate locality (turning up their noses at an unpleasant smell, for instance, or smiling when a friend enters the room); they can respond with a steady flow of visual feedback, positive or negative, to activities and conversations going on around them (making intermittent eye contact with whoever is holding the floor to indicate that they are attending; looking ostentatiously away when a speaker is trying to engage their attention to indicate boredom or dislike; squeezing their face into an interested or alert expression, or raising their eyes to heaven in a face-threatening display of indifference or disbelief). If they are the speaker, they can make use of similar sets of kinesic cues to manage turn-taking during conversation: indicating by means of a head movement or eye contact that they are willing to cede the floor (Levinson, 1983: 302), perhaps to a specific targeted individual, or avoiding such behaviours precisely as a means of prolonging their conversational turn. Verbal behaviour of various kinds can also come into play: participants who know they can be heard by co-present others can choose to produce a range of 'back-channel' cues (Goffman, 1981: 14) to indicate conversational support ('uh-huh', 'then what?'); once they are themselves in possession of the floor they can select the next turn-taker via a direct address to some or all of those co-present (consider that staple of the seminar situation, 'what do *the rest of you* think?', uttered by a tutor to the seminar group at large, often

when one participant has been dominating the discussion). In a shared locale, prosodic cues, too, have their part to play: an increased speech rate or louder voice will drown out someone who is trying to interrupt; particular intonation patterns can be used to indicate the speaker's affective orientation towards their own contribution (surprise; amusement; grief; fury…); paralinguistic features such as particular tones of voice can indicate to familiars that the speaker wishes their utterance to be taken ironically, or at face value, or with a pinch of salt. Taken together, in face-to-face encounters prosodic, kinesic and linguistic cues can function to structure interactions and to deliver a range of expressive signals, contextualizing what is said and permitting the disambiguation of potentially face-threatening acts. As Thompson puts it:

> Participants in face-to-face interactions are constantly and routinely engaged in comparing the various symbolic cues employed by speakers, using them to reduce ambiguity and to refine their understanding of the message. If participants detect inconsistencies, or cues that do not tally with one another, this can become a source of trouble which may threaten the continuation of the interaction and cast doubt on the sincerity of the speaker.

> (Thompson, 1995: 83)

To see and be seen, to hear and be heard also facilitates *dialogue* (Thompson, 1995: 82). While we can easily conjure up a range of face-to-face interactive situations which involve one participant holding the floor in a lengthy monologue (institutional settings such as lecture theatres, classrooms and churches being obvious examples[8]), face-to-face encounters, as the previous paragraphs' brief delineation of turn-taking mechanisms suggests, also allow for a fully bilateral exchange of information in which more than one co-present participant can take the floor successively, responding to remarks made by previous individuals and initiating new topics, monitoring feedback and being monitored in turn. Conversation, 'that familiar predominant kind of talk in which two or more participants freely alternate in speaking' (Levinson, 1983: 284), is always potentially on offer when two or more individuals are in the same place.

1.3 Communicative Affordances

In canonical encounters, *space* (the set of relations between relevant objects and individuals) is fundamentally linked to *place* (the arena in which these relations are structured), so that the relevant physical relationships between conversational participants, and between those participants and the objects around them, are constrained by the boundaries of the locale in which they find themselves. Two or more people in the same place thus loosely share a set of spatial relations to their immediate environment and to each other. If they are in a garden, then a tree, say, might loom over all concerned if they are seated beneath it, or will

---

[8] See Scannell (1996: 18ff.) for a discussion of the structural features attendant upon participation in institutional occasions.

block the sun if they are lying on loungers with the tree overhead. Similarly, if I am watching television with three of my friends then one of them may be on my left, on the second of two sofas which face in the general direction of the television set, and the other two may be on my right, sharing the other sofa with me; the television itself will be broadly 'in front' of all of us, though each individual's spatial orientation to it will be slightly and uniquely different, depending on where they are sitting, whether they are leaning forward or back, whether they are slumped in a position where they are forced to look upwards in order for their gaze to encounter the screen, are alert and upright and gazing directly ahead, or are reaching for another handful of snacks, with their gaze and their primary bodily orientation directed elsewhere in the room.

It is this fundamental association of space with place in the canonical encounter which gives rise to the interactional characteristics that I discussed in the previous section: the ever-present possibility of conversational encounters taking place; the use of deictic expressions, and of gestural deictics in particular, to index elements in the immediate vicinity; and the availability of linguistic, paralinguistic, prosodic and kinesic features which cue co-present listeners as to who is holding the floor, what their communicative intentions are and how they are responding to objects, events and individuals around them. In a situation where space maps onto place, participants can make use of expressions that either index features *of* the environment (*this smell*; *that sound*; *those apples*) or express a spatial relation *to* the environment ('put it *over there*'; 'bring that *here*') with reasonable confidence that their listeners will understand them; the full and extensive use of linguistic, extralinguistic and paralinguistic cues is similarly dependent upon a set of circumstances in which you and I both have at least the potential to be spatially orientated towards each other in the same place.

Such a situation only holds good, of course, if all relevant conversational partners are co-present. In mediated interactions, where one or more participants are physically absent and communicating via an intervening medium rather than through the air in the place in which they jointly find themselves,[9] spatial relations can no longer be contained within a single bounded arena:

> In pre-modern societies, space and place largely coincide, since the spatial dimensions of social life are, for most of the population, and in most respects, dominated by 'presence' – by localised activities. The advent of modernity increasingly tears space away from place by fostering relations between 'absent' others, locationally distant from any given situation of face-to-face interaction.

> (Giddens, 1990: 18–19)

---

[9] 'Primarily' is an important word here: face-to-face encounters may involve limited kinds of intervening media, such as the piece of paper used by students in a classroom to exchange scribbled notes with each other, or the megaphone used by speakers to address a crowd at a rally. These will not count as mediated interactions for our purposes here as the definition in 1.2 specifically makes reference to physical absence. Cf. Thompson (1995: 83) who similarly situates the mediated interaction in terms of 'the use of a technical medium … which enables information or symbolic content to be transmitted to individuals who are remote in space, in time, or in both'.

I will return in the next chapter to consider the wider question of modernity which is raised in this quote from Giddens, and to discuss its implications for understanding both space and time in mediated interactions. For the moment, however, I wish to concentrate on the kind of 'dislocation' (Moores, 1997: 238) of space and place which Giddens also broaches here.

This notion of dislocation we can usefully investigate via a further examination of the typology of interactional characteristics which we examined in 1.2, with reference to the extent to which particular electronic media offer the full range of sensory modalities to the embodied individual. This discussion, in turn, will provide a basis for classifying various kinds of remote encounter along a cline stretching from 'thin' at one extreme (encounters with objects which provide only the most limited kind of sensory and/or kinesic access) to relatively 'thick' encounters at the other, where particular media appear to offer the individual a relation to the mediated world which approximates to full co-presence. This will permit us to undertake a functional analysis of the relationship between particular communicative situations and the interactional characteristics that they do or do not facilitate.

The idea of the *communicative affordance* will be of central importance to this discussion. The concept of affordances was originally developed by Gibson (1982), and refers broadly to the way in which the functionality of an object does or does not enable particular kinds of actions on the part of an individual engaging with it. Hutchby (2001: 26) takes up and expands upon this idea to talk about the notion of the *communicative* affordance, the 'possibilities for action that emerge from … given technological forms' (Hutchby, 2001: 30). The concept of communicative affordances has already implicitly permeated the discussion in this chapter. To talk about the range of interactive structures and behaviours that are characteristic of the canonical situation is to talk about the communicative affordances of face-to-face engagements, about the kinds of encounters that they allow or disallow.

The notion of communicative affordances, as Hutchby is at some pains to point out, is not technologically deterministic. Although a particular form or medium may proscribe certain kinds of behaviour and seemingly prescribe others, users may choose not to interact with it in the anticipated way. As an example of the manner in which technological forms may be appropriated for purposes other than those originally intended, Hutchby cites Grint and Woolgar on the development of telephone technology:

> [T]elephone technology was used originally to broadcast concert music. It was not axiomatic to its design that the telephone system would ultimately be restricted primarily to two-way personal communication … The original use of telephone technology, and indeed its use now, was and is the result of interpretation and negotiations, not determinations.

> (Grint and Woolgar, 1997: 21, cited in Hutchby, 2001: 21)

Conversely, we can consider the early arguments around the appropriate uses of radio technology, which 'was first conceived as a means of point-to-point communication' (Peters, 1999: 206), but where the inherently public and

unconstrainable nature of the signal aroused anxieties concerning the privacy of communication, leading to its eventual manifestation as a broadcast medium decades later. Radio waves, as Peters (1999: 195) points out, can potentially be used either as 'a central exchange for many voices' (radio broadcasting) or as 'a means for point-to-point contact' (ham radio); the same holds true for telephone technology (party lines on the one hand; mobile phones on the other). That the one developed primarily for interpersonal communication and the other for broadcast has to do with a dialectic between the communicative affordances of each form on the one hand and the history of its contexts of appropriation on the other.

The developers, manufacturers and managers of particular technological forms, in other words, cannot altogether prescribe what their equipment will be used for. Nor can they control the kinds of behaviours which will accompany its use. Peters, for example, notes the entertaining comments of one American journalist concerning the audience's lack of listening decorum:

> Bruce Bliven noted in 1924 that most political orators, if aware of 'the ribald comments addressed to the stoical loud-speaker' of the home-receiver, would seek other jobs. 'The comments of the family range from Bill's "Is *that* so!" down to Howard's irreverent "Aw, shut your face, you poor hunk of cheese!"' Home listening allowed oratory to be received in a mood of chronic flippancy.
>
> (Peters, 1999: 213)

The array of uses to which a particular form will be put, then, cannot be entirely anticipated or controlled; nor can the range of attendant behaviours. We should not, however, allow the anti-deterministic thrust of the argument to blind us to what will be the central issue for the rest of this chapter. If particular forms can give rise to unintended uses and/or modes of reception, it is nevertheless part and parcel of the communicative affordances of any given form or medium that there are certain activities which it *disenables*, which it does not allow the user to pursue.

We can take as an initial starting point here Peters' contention that there are certain sense faculties that are not amenable to communication at a distance:

> Of all the senses, touch is the most resistant to being made into a medium of recording or transmission. It remains stubbornly wed to the proximate; indeed, with taste, it is the only one sense that has no remote capacity … Touch defies inscription rather more than seeing or hearing.
>
> (Peters, 1999: 269)

Why should it be the case that the objects of touch and taste – and smell, for that matter – cannot be accessed via a remote encounter? The short answer is that all three are faculties that depend for their functioning upon an immediate and corporeal interaction with the world. To taste something, our tongues must physically engage with it; to touch, our hands or bodies must do so; to smell, we must be in the immediate vicinity of the olfactory object; and self-evidently, we cannot employ tongues, hands, noses or our own bodies at a distance from ourselves. As Merleau-Ponty (1962: 316) puts it, 'Tactile experience … adheres

to the surface of our body'. Vision, he adds, 'presents us with a spectacle spread out before us at a distance, and gives us the illusion of being immediately present everywhere and being situated nowhere' (Merleau-Ponty, 1962: 316); the same would seem to hold true of hearing. On the face of it, therefore, no medium will be able to afford us access to touch or taste or smell at a distance because each of these requires immediate proximity in order to operate.

But if part of the answer relates to the ways in which our sense faculties function, then another important element has to do with the question of communicative affordances. We cannot see or hear anything beyond our immediate vicinity either: once something is too far away for me to see it, or out of the range of my hearing, then I will not have direct access to what it looks or sounds like any more than I will to its touch or taste. When we listen to a voice on the telephone or interact with a friend over a video-conferencing link we do not reach directly into their remote location by some species of teleperception (eyes or ears on stalks); rather, we engage, *where we are*, with a *representation*, a filtered and transformed variant of the thing: a voice on the phone which sounds sultrier than it is because of the suppression of high frequencies (Peters, 1999: 196); a video-conferenced homunculus on our computer screen which is pixilated and out-of-sync with the words it is speaking. Even if the image on the screen were to be maximally high-resolution, or the voice on the phone were indistinguishable from the unique vocal qualities of a particular individual, though, we would still be encountering a representation: we cannot remotely encounter the real.[10]

It is the 'technical forms' that Hutchby talks about which deliver these representations to us: projectors, telephones, computers, televisions, radios, VCRs, DVD players and the like. Hence the difficulty as far as taste, touch and smell are concerned. Even if we were prepared to lick or paw the apparatus to achieve gratification, thus overcoming the problem of proximity, nothing could be oozing out of it to meet us. Taste or touch do not simply require immediate contact, but immediate contact with an appropriate perceptual object. The apparatus would have to transmit something that we could lick or paw in the first place, and it is entirely unclear what species of medium would permit the transmission of the kind of three-dimensional entities that we would need to directly engage with in order to touch or taste them. We could, it is true, get hold of a card from a television station which would permit us to lick particular squares in order to approximate to the taste of a food just then being consumed on the television screen; the ill-fated experiments with scratch'n'sniff cards in cinemas spring to mind here. But an encounter with a scratch'n'sniff card is in principle no different from the individual who sniffs their beloved's letter, hoping to find a trace of their perfume on it. Such an encounter does not transmit to us, *from a distance*, something that we could smell, but only allows us to engage, *where we are*, with the piece of card or paper that the scent is imprinted on. It is possible, of course, to consider the Proustian case where a representation interacts with memory in such a way as to cause us to have something akin to

---

[10]Whether we can directly encounter the real under any circumstances is a question that is beyond the scope of this book; but see Chapter 5 for a further discussion of this matter.

a momentary apperception of a similar object which we have encountered in another time and another place; such an apperception, though, does not deliver the perceptual object itself to the *place where we are* but only reminds us of what it felt like to encounter it elsewhere. It is also, perhaps, possible to imagine a technological form which could waft odours for us to encounter in the privacy of our own room at dramatically appropriate moments, making the sense of smell a kind of intermediate category in this discussion. Were such an apparatus to teleport 3D objects for us to touch or taste, however, then those objects would inevitably, once teleported, be local rather than remote: although they would be coming at us from afar, we would encounter them face-to-face.

1.4   The Mediated Encounter

With touch and taste – and to a considerable extent, smell – taken out of the equation, we are left with just four of Avery and McCain's seven sensory modalities to consider: the audio verbal, the audio non-verbal, the visual verbal and the visual pictorial. Only the faculties of sight and sound, in other words, have the potential to be engaged when we interact remotely with the world. The sense of touch, it should be noted, will not re-enter the picture even in the case of written communication. It is, of course, perfectly possible for an individual in receipt of a missive from the beloved not only to sniff the piece of paper but to touch it and perhaps even to rustle it as well as actually reading the contents. This scenario suggests an encounter with the medium which seems little short of the thickness of actual co-presence, with sight, sound, smell and touch all entering the equation; but the approach is fundamentally flawed. It is not the encounter with the *medium* which will be of relevance here, but the encounter with the absent individual or individuals which the medium affords. The example of the letter suggests the thickness of co-presence because co-presence (of the recipient with the letter) is the actual state of affairs.

If it is the case, then, that the only sense faculties that remote encounters permit us to bring into play are the auditory and the visual, then it immediately becomes clear that no mediated encounter is going to be able to deliver anything other than a relatively thin engagement with the world. This conclusion is entirely in keeping with the argument developed earlier in this chapter. If, as Merleau-Ponty argues, it is only in full co-existence with a phenomenon that we can be said to comprehensively engage with it (Merleau-Ponty, 1962: 318), then remote encounters, with their fundamentally restricted access to the range of sense modalities on offer in the canonical situation, will be able to bring the world to us only in the flat slices that are all that remain after the majority of information has been cut away.

What, though, does this discussion have to tell us about the affordances of different electronic media of communication, as far as encounters with absent others are concerned? Thompson, in his account of mediated encounters, draws a distinction between three types of interaction: face-to-face interactions, mediated interactions and what he refers to as the 'mediated quasi-interaction'

(Thompson, 1995: 82). Four characteristics are used to define these categories: whether they involve a local or a remote encounter; whether the communication is aimed at specific or non-specific others; the dialogical or monological affordances of the situation; and the range of symbolic cues which they do or do not afford participants (1995: 85). The example of the letter which we have been exploring here, for instance, would constitute, in Thompson's terms, a remote encounter (the recipient is in another place at another time) with a specific other, which is potentially dialogical in nature (the recipient can write back), and has a particular and limited range of symbolic cues on offer: for Thompson, the letter would count as a *mediated interaction*. These are characterized by the dislocation of space and place; the limited availability of symbolic cues; an orientation towards a specific other or others; and the potential for a bilateral encounter. As a more detailed example of this category we can consider video-conferencing.

Both sight and sound – the two faculties that we have left at our disposal – would appear at first glance to be readily facilitated during video-conferencing. I can see, on my computer screen, the remote individual I am interacting with, and I can presume that they can similarly, in their own context of co-presence, see me; and I will be able to make a similar assumption about our mutual ability to hear each other.

Video-conferencing, by this description, should thus afford individuals the range of interactive structures and behaviours that characterize the canonical encounter, as these are primarily dependent upon an ability to see and hear fellow participants. At first glance, this would appear to cast doubt on Thompson's assertion that mediated encounters involve a restricted range of cues. In a situation where you and I believe ourselves to be simultaneously audible and visible to each other we can certainly engage in dialogue; and many of the other features I discussed in 1.2 would seem to be equally accessible. Gestural deictics, for example, which depend for their felicitous use on my hearer being able to monitor my indexical behaviour, seem to be warranted by a situation where my hearer can see me; ditto, the use of the array of linguistic, paralinguistic and kinesic cues which permit the two of us to structure the interaction and convey an assortment of affective states and responses. Just as in the canonical encounter, I appear to be able to presuppose that my listener, if they know me well enough, will be able to have a crack at decoding paralinguistic features of my utterance in order to recognize that I am intending to be taken ironically when I say, from Glasgow, 'it's a lovely day again here'; will recognize that particular kinds of vocatives are intended to elicit a response from them and thus to pass them the floor ('what's the weather like there, Jane?'); will appreciate that my intermittent uttering of back-channel cues (wow, really?) indicates that I am attending to and gripped by their contribution, and so on. We appear, in other words, to be able to engage in conversation, with the full range of resources available to us.

The above discussion needs to be qualified in a number of ways, however. The webcam that relays my image to an absent other can take in only a certain amount of my immediate environment, generally centred around my face and upper body; whilst I can indeed use gestural deictics, my use will be constrained by what I can hold up to the webcam or can reasonably assume to be sufficiently proximate to

fall within its visual range. I cannot conceivably indicate objects that are behind the computer – a picture on the wall, say – unless I swivel the webcam round and cease temporarily to be within the frame of my own interaction. This kind of remote encounter, in other words, does not afford participants the kind of 360 degree access to the perceptual field which they can readily obtain in a canonical encounter. There will likely be similar restrictions governing your ability to *hear* what is going on in my vicinity. You will almost certainly have access to my voice; but depending on the quality of the microphone, I may or may not be able to felicitously draw your attention to the sound of the telephone ringing in the next room by saying 'hang on, I have to answer *that*'.

Further difficulties abound. I have already noted the reduced and pixilated representation of the individual which video-conferencing affords. Hutchby adds to this list of potential drawbacks the issue of download speed:

> Even the most powerful desktop computers have trouble displaying real-time internet video at a data transfer rate sufficient to eliminate observable jerkiness and 'frame-dropping', largely because of internet bandwidth restrictions. This is the case even with images displayed in a box only a couple of inches square. Attempts at a full-screen resolution result in a 'blocky' image in which the nuances of phenomena such as gaze direction and facial expression are easily lost.

> (Hutchby, 2001: 124)

Hutchby notes, too, a further relevant phenomenon. The webcams which permit us to see remote others are not embedded in the computer screen, but are elsewhere, generally on top of the computer monitor (2001: 127). In order to see you, I need to focus my gaze on the window in my screen where your image is available; if I do so, however, then I will not appear to you to be meeting your gaze, for which I need instead to be directing my gaze to the webcam that captures my image for you. To you, my direct gaze at your image on my screen, which represents for me the intimacy of eye contact, will appear as a gaze downwards. Video-conferencing would therefore appear to rule out the possibility of a mutual gaze: for me to appear to you to be making eye contact, I need to be looking into the webcam, in which case I cannot look to see whether you are making eye contact with me.

This combination of circumstances has a number of implications for the degree to which video-conferencing affords its participants the full array of interactive behaviours. As the quote from Hutchby makes clear, facial expressions cannot be reliably used to convey feedback to my speaker, although as a user of the technology I may of course be unaware of this and may be happily contorting my face in a variety of elaborative gestures which cannot be recognized in the window on your screen in which I am currently manifesting myself. Kinesic cues such as head movements to indicate that I am offering you the floor may also be hard to read given the resolution of the image; and the use of gaze for similar ends will be destined, of course, to fail, given the asymmetry of gaze which the situation as it is currently constituted affords. This particular technology, in other words, does indeed restrict the range of cues, both symbolic and otherwise, which a participant can provide for the absent other they are addressing.

Thompson contrasts his category of mediated interaction with what he refers to as *mediated quasi-interactions*. Whilst the latter also involve Thompson's 'separation of contexts' (1995: 85) and a limited set of symbolic cues, they are distinguished by being aimed at an indefinite range of participants and by being fundamentally monological in character. What Thompson has in mind is 'the kinds of social relations established by the media of mass communication (books, newspapers, radio, television, etc.)' (Thompson, 1995: 84). We can take as an example here television's address to the audience at home. Video-conferencing offers its participants the possibility of seeing and hearing along a single pathway: a bi-directional vector stretching from my machine to yours and back again for as long as we are both connected to the Net and the relevant software is running. Television, like video-conferencing, involves the faculties of sight and sound; but the sensory modalities that can be accessed – the audio verbal, the audio non-verbal, the visual verbal and the visual pictorial – are typically said to be asymmetrically available. When I watch television, I can see what there is to be seen on the screen, and hear what is coming out of my speakers, but I can be neither heard nor seen at the other end. If I am watching and listening to an individual on my screen, then, I cannot directly respond, or so the argument goes: unlike video-conferencing, this individual will have no access to my own context of co-presence, and is therefore functionally blind and deaf as far as any communications on my part are concerned. I can see and hear, but not speak; the other, on the screen, can speak but cannot see or hear me.

What are the consequences of this asymmetrical communicative situation? Television performers, for their part, can felicitously carry out many of the activities and behaviours that are associated with face-to-face communication. Just as gestural deictics can be used by either participant in a video-conferencing transaction, so they can be used by the television performer to index elements in their immediate vicinity. The television chef, for instance, can clearly indicate items of food, cooking paraphernalia and the like via a gesture and a demonstrative expression ('*this* is what we'll be preparing today'), secure in the assumption that anyone at home paying attention will be able to determine the appropriate object because they will be able to see the relevant gestures. Individuals on television can, too, make use of particular tones of voice to mark their utterance as ironic or deadly serious; and can indicate via their intonation that they are angry, or neutral, or amazed.

What they cannot typically do, however, on this view, is to make use of some of the interactive mechanisms that characterize a dialogical exchange. Those interactive mechanisms that have to do with *addressing* relevant conversational participants are still available: the gaze to the webcam can be replaced by a gaze to camera, and direct address of a non-specific kind ('on today's programme *you're* going to see') can clearly operate felicitously in a situation where they can be seen and heard by absent others. Performers cannot, however, offer the viewer the floor; nor would there be any point in them producing back-channel cues, as there is no flow of information from the viewer for them to respond to or support.

On this view, television's affordances are thus clearly different, characteristically, from the affordances of the Internet, with its potential for symmetrical communication. We can typically rule out conversation, for a start, which requires a two-way flow of information: even though, as the earlier quote from Peters suggests, viewers can rail and shake their fists at the screen when they do not like what they are watching, the performer at the other end will be blissfully unaware. There is no point in their using gestural deictics, nor indeed is there any point in their having recourse to any of the behaviours that would allow them to interact with the other: they are invisible and inaudible, and cannot communicate with the individual who is addressing them.

This kind of 'mediated quasi-interaction', however, with its absence of one-on-one communication and its preclusion of a direct response on the part of the viewer or listener comprises only one of many possible 'communicative circuits' (Scannell, 1991: 11) in play in broadcast communication. There are a number of routes through which individual members of the remote audience can enter into a bilateral encounter with a performer on television. Other media of electronic communication such as the telephone, for example, can be used to text or to speak directly with someone on the television; and there is, too, the example of the kind of live television programme where cameras are set up in remote locations in such a way that members of the audience can be momentarily seen and heard on the screen.[11] For this reason, in what follows, I will continue to use the term 'mediated encounter' to cover all instances of televisual interaction-at-a-distance. Whilst television, in its address to its audience, clearly and routinely constructs a different set of affordances from those available to someone engaged in a two-way Internet transaction, it also affords other modes of engagement, less asymmetrical in form.

There is another reason, too, for maintaining a single overarching category of the mediated encounter in what follows. This chapter has argued that the interactional structures and behaviours which can arise in particular kinds of communicative situations are directly consequent upon the nature of the context of utterance. Thompson's account does not dispute this. He stresses, for instance, that his category of 'mediated interaction' is stretched across time and space and 'thereby acquires a number of characteristics which differentiate it from face-to-face interaction' (Thompson, 1995: 83); indeed his project is bound up with tracing the social and interactional consequences of the forms of communication characteristic of modernity. Where the two arguments might diverge, however, is in the emphasis that will be placed here on the affordances of particular media of communication, on what they will or will not allow individuals to do. Where Thompson wishes to distinguish mediated quasi-interactions from mediated interactions on the basis of the particular communicative circuit in operation and its implications for dialogical communication, the fundamentally *phenomenological* investigation in this chapter would lead us, rather, to consider such encounters as a routinized consequence of broadcasting's affordances to

---

[11]On the role of other forms of electronic communication in constructing an interactive channel, see Chapter 7; on live television and the articulation of different places, see Chapter 6.

its users. The electronic broadcasting media allow individuals to be seen and heard at a distance via the microphone, the camera and a variety of methods of broadcast transmission (cables, satellites, terrestrial signals) but not to see or hear through these particular artefacts; a one-way address to absent others would seem to be an obvious entailment of this particular set of circumstances, although the capacities of users to transcend the limitations of a particular form or medium are not to be underestimated, as we shall see in due course.

1.5 Conclusion

I began this chapter by saying that mediated encounters and face-to-face interactions are significantly different from each other. We have now seen at least some of the reasons why this is the case. In a context of co-presence, space maps onto place, and co-present individuals will thus share, subject to their precise orientation, a roughly equivalent set of spatial relations to the objects around them. To be present in a particular place, furthermore, affords individuals access to the full range of sensory modalities on offer. All participants, to the extent that they are in full possession of their senses, will therefore be able to smell, touch, taste, hear and see the same objects. Taken together, these properties of the canonical encounter mean that individuals can identify and locate a range of phenomena with a reasonable degree of security that the objects and individuals they indicate will be identifiable to those around them, and can make use of a full range of prosodic, kinesic and linguistic cues to engage with each other.

In mediated encounters, by contrast, individuals do not inhabit the same context of co-presence. This dislocation of space from place means that they no longer share a spatial orientation to the set of objects in each vicinity. Nor can they touch or taste (or, very likely, smell) the same phenomena; and to the extent that they can see or hear what is going on elsewhere, this is by virtue of a locally available representation only, with all that that entails in terms of image or sound quality. As a consequence, mediated encounters effect a 'narrowing' (Thompson, 1995: 85) of the available range of conversational cues, and may also involve restrictions in the use of certain kinds of context-dependent language such as deictic expressions. Under certain circumstances, furthermore, mediated encounters will only permit asymmetrical encounters, where one or more participants will be unable to interact directly with others.

Mediated interactions may, however, possess a range of mechanisms to compensate for the lost functionalities of the canonical encounter, and it would be a mistake to view them simply in terms of what they lack in comparison to face-to-face encounters. Whilst certain kinds of interactive machinery will not work because of the restricted access to relevant conversational participants which a particular technological form or medium may afford, other mechanisms will exist, or will be developed, which will allow users to overcome with varying degrees of success the limitations of the apparatus. Take, for example, the letter that I discussed earlier. The writer cannot make use of prosodic cues to orient the recipient to the import of the communication, but can, for instance, use

underlining or CAPITALS expressively, and has the option of communicating via graphic cues as well (little hearts, kisses and a variety of affectionate squiggles). As Thompson puts it:

> Communication by means of letters … deprives the participants of a range of cues associated with physical co-presence (gestures, facial expressions, intonation, etc.) while other symbolic cues are accentuated.

(Thompson, 1995: 83)

Or take the example of deictic expressions. We might well expect it to be the case that gestural deictics cannot be used felicitously in written communication, as they depend for their interpretation on other participants being able to monitor the speaker's indexical behaviour. It ought, then, to be the case that I cannot, in a letter or email, expect my recipient to make sense of what I am saying if I ask them to 'look at this', or indeed if I enquire whether my 'bum looks big in this'. This is not, however, entirely the case. An arrow, indexing a drawing in a letter or a photograph in an email, will serve just as well as a physical gesture; similarly, we would have no trouble making sense of a button on a website which urged us to 'press here'. Just as pictorial cues in letters (and the offline or online use of emoticons in computer-mediated communications) permit the writer to add affective overtones of various kinds to what they are saying, so graphic information can be used indexically, in place of the physical indexical behaviour of a co-present participant, to pick out particular features of the environment.

Nor should we necessarily think of these mechanisms invariably as compensatory, as a way of overcoming specific interactional problems entailed by the thinness of remote encounters. This would be to inappropriately privilege the canonical encounter, and although the argument in this chapter takes the canonical situation as ontologically prior (as Thompson (1995: 81) puts it, 'For most of human history, most forms of social interaction have been face-to-face') it is nevertheless clear that more historically recent forms of communication-at-a-distance do not simply seek to mimic the patterns and structures of face-to-face interaction but may in some cases complement or supplant them. The affordances of an apparatus that delivers a remote encounter may be additive as well as subtractive, in other words; and while I will have reason to argue that there may be pressing reasons why particular media of communication might wish to develop mechanisms to simulate the immediacy of the face-to-face encounter, it is also the case that the absence of certain functionalities in certain situations may be a positive advantage. The letter, after all, may be designed to end a relationship in a way which avoids the intimacy and emotional commitment of a face-to-face encounter: the writer may precisely wish to avoid monitoring the recipient's feedback. We can cite also in this respect Hutchby's analysis of Internet relay chat, which affords certain possibilities for conversational turn-taking, such as nominating the individual at whom the writer wishes to target their utterance ('allo Nicky'; 'you got a pic then zazoo') but where the interactive free-for-all constituted by a multiplicity of participants all potentially typing messages at the same time can also be, in Hutchby's terms 'serendipitous' (Hutchby, 2001: 191),

creating ambiguities in turn-taking and address which permit participants a wider freedom to interact with unknown others than they would have in a face-to-face situation where kinesic, linguistic and prosodic cues would likely function to circumscribe the interaction.

All mediated encounters, furthermore, self-evidently afford their users one incalculable advantage over individuals interacting in a canonical situation. Whilst, as we have seen, mediated encounters are perilously thin in terms of their ability to deliver to remote participants the full range of sense objects in the distant situation, thus imposing particular kinds of limitations on the way in which interactive tokens can be displayed, it is nevertheless the case that mediated encounters allow us to engage with others *at a distance*, to stretch our communications across space. Seen from this perspective, it is the canonical encounter which comes to look limited, with its narrow and singular perceptual field and the constraints it imposes upon participants' ability to jointly transcend space in their transactions with each other. To truly get a handle on the extraordinary nature of these taken-for-granted encounters-at-a-distance which now routinely permeate our contexts of co-presence, though, it is necessary to think not only about matters of space and place but also about the related question of *time*. This I will proceed to do in the next chapter, where we will start once more from first principles in order to examine the impact of mediated communication on individuals' experience of the world at the moment in which they encounter it.