1. Video analysis: interactional coordination in movement and speech

If society is conceived as interaction among individuals, the description of the forms of this interaction is the task of the science of society in its strictest and most essential sense.

Simmel 1950, pp. 21-2

This work is part of a program of work undertaken several years ago to explore the possibility of achieving a naturalistic observational discipline that could deal with the details of social action(s) rigorously, empirically, and formally. . . . Our analysis has sought to explicate the ways in which the materials (records of natural conversations) are produced by members in orderly ways that exhibit their orderliness and have their orderliness appreciated and used, and have that appreciation displayed and treated as the basis for subsequent action.

Schegloff and Sacks 1973/1974, pp. 233-4

I want to argue that however rich a researcher’s imagination is, if he uses hypotheticalized, typicalized versions of the world, he is constrained by reference to what an audience, an audience of professionals, can accept as reasonable. That might not appear to be a terrible constraint, except when we come to look at the kinds of things that actually occur. Many of the objects we work with would not be accepted as a base for theorizing if they were urged as imagined. We can then come to see that a warrant for using close looking at the world as a base for theorizing about it is that from close looking at the world we can find things that we couldn’t, by imagination, assert were there. One wouldn’t know they were “typical.”

Sacks, quoted in Jefferson 1983b, pp. 17-18

Introduction: medical interaction and video analysis

In recent years there has been a growing interest amongst both medical practitioners and social scientists in communication in the consultation. In the United Kingdom it is especially within general practice or primary health care that we find a growing concern for the relationship and interaction between the doctor and patient. It is now widely recognized
2  

that the everyday practice of medicine, the process of diagnosis and prognosis and restoring persons to health and normality, is thoroughly bound up with the ways in which doctors and patients communicate.

Though the significance of communication in general practice had been noted for many years, with discussions on the bedside manner and related subjects, it was perhaps Michael Balint in his classic essay *The Doctor, His Patient and the Illness* (1957) who more than anyone else brought to the profession’s attention the importance of communication in the consultation. This is not to suggest that many general practitioners formed or participated in “Balint groups” or were directly influenced by his work. Rather his powerful demonstrations of unexplored illness and the criticalness of communication to diagnosis and treatment permeated the profession and gave support to the growing arguments for postgraduate training and research in general practice.

Amongst the social sciences it is perhaps in sociology that we find the greatest commitment to the analysis of the consultation or, more generally, social interaction in medical settings. As far back as 1935, Henderson, drawing from Pareto, describes physician–patient interaction in terms of the constituent parts of a social system. More important, in 1951 Parsons published his classic *The Social System*, a chapter of which is devoted to an analysis of modern medical practice in relation to the maintenance of social equilibrium, an analysis which provides a rich conception of the mutually compatible roles of physician and patient. This chapter alone not only revealed the significance of doctor–patient interaction to sociological inquiry, but is widely accepted as forming the beginnings of medical sociology itself. However, it was the lectures and essays of Everett Hughes at the University of Chicago in the 1950s which led to the emergence of a wealth of empirical work, largely ethnographic, concerned with social interaction in medical settings. Studies by Becker et al. (1961), Davis (1960, 1963), Glaser and Strauss (1965), Goffman (1961), Roth (1963), Strauss et al. (1964), and many others provide a substantial body of findings and an array of insights concerning the organization of everyday medical practice and the interaction between the profession and its clientele.

Since these studies there has been an immense variety of research conducted on interaction and communication in medical settings. The ethnographic tradition has continued with studies by Emerson (1970), Sudnow (1967), and more recently Strong (1979), and a range of other theoretical and methodological perspectives, both qualitative and quantitative, have been used to examine the behaviour of doctor and patient in the consultation. An important development over the past decade or
so has been the use for analysis of audio recordings of actual medical encounters. The recordings provide the researcher with access to detail unavailable to more traditional modes of inquiry such as fieldwork, observation, and interview. The year 1976 saw the publication of the classic *Doctors Talking to Patients*, by Byrne and Long, a detailed study of the verbal behaviours of general practitioners in more than two thousand consultations. Subsequently we have seen the emergence of a range of empirical studies concerned with the details of doctor–patient interaction, especially talk in the consultation, conducted from a variety of perspectives by both social scientists and medical practitioners; see for example the studies reported in Atkinson and Heath 1981; Fisher and Todd 1983; Pendleton and Hasler 1983; and Tanner 1976.

As yet, however, there have been relatively few studies concerned with the visual aspects of behaviour in the medical consultation. One explanation is that it is only recently that a cheap, reliable, and relatively unobtrusive technology for recording vision as well as sound, namely video, has become widely available. More important perhaps is that it is largely psychology and social psychology which have developed empirical research in the area of visual behaviour. The main thrust of this work, but by no means all, is experimental and has necessarily been conducted under laboratory conditions. Consequently studies of naturally occurring behaviour in particular habitats such as the medical consultation are relatively few. An important exception, though it is perhaps inappropriate to consider it in terms of traditional studies of visual communication, is the major body of research which had its beginnings in the work of Bateson, especially Bateson and Mead 1942 and Ruesch and Bateson 1951 and the interdisciplinary collaboration between Bateson, McQuown, Hockett, Birdwhistell, and others in the early 1950s at the Institute of Advanced Study in Palo Alto. The collaborative research at Palo Alto was never published, but it did form the background to Birdwhistell’s studies (1970) of body motion and Schefflen’s important analysis of psychotherapy sessions (1963, 1966, 1973) and influenced directly and indirectly a range of other empirical work on visual aspects of human behaviour in natural settings, such as Condon and Ogston 1966, 1967, 1971 and Kendon 1967, 1972, 1974a, b, 1977.

In general, sociology, unlike psychology, social psychology, and social anthropology, has been slow to take up the opportunity afforded originally by film and now video. Though there are disadvantages to video in comparison with film, it does provide a cheap and unobtrusive means of recording both the vocal and visual behaviour of human beings in situ and subjecting it to close and detailed scrutiny. It provides the facility
of making repeated viewings of a fragment of human interaction and the possibility of identifying features of behavioural organization previously unavailable to scientific observation. Moreover video allows the researcher to make available raw data to the scientific community and provide others with the opportunity of evaluating observations and findings, at least in public presentations. On these grounds alone it might be imagined that the emergence of video would have a significant impact on sociology, if not lead to a scientific revolution akin to the impact of the microscope on biology. As yet, however – and we are now into a decade or so of relatively cheap and efficient video technology – sociology has not shown a substantial interest in using the medium for research.

Whatever the financial difficulties of universities both in the United Kingdom and abroad over the past few years, it is unlikely that they explain the near-total absence of video in sociological analysis. The ability to record both the visual and vocal facets of human behaviour may well be irrelevant to certain modes of investigation and sociological concerns; for example, many forms of quantitative analysis might well find no advantage in video technology and the potential it affords. Yet in sociology there is a strong ethnographic tradition, an approach which in various ways emphasizes the importance of grasping the perspectives of the participants and examining social interaction in natural settings. In fact some documentary programmes shown on television are themselves fine examples of ethnography, providing rich insight into the social organization of a particular setting or activity. Sadly, however, ethnographers, save for some important exceptions such as Erickson and Schultz 1982 and Gumperz et al. 1979 which have in general tended to emerge from social anthropology rather than sociology, have fallen behind their colleagues in the media, rarely using video even to supplement the more conventional modes of gathering data. There are of course difficulties of access and recording in some settings of interest to sociologists; more important perhaps is the lack of an analytic framework for handling data collected on video. The theories and concepts conventionally used in ethnography, though finely suited to data generated through fieldwork, observation, and interview, may not, at least as they are traditionally conceived, be applicable to video recordings of actual activities and settings. In examining video recordings of naturally occurring activities, the researcher is faced with a level of detail in human interaction that renders our more familiar sociological concepts and analytic devices somewhat inappropriate save in a very crude sense.

The absence therefore of sociological research using video technology has derived in part from the lack of an analytic framework that can guide
the investigation of recordings of naturally occurring actions and activities and their social organization. However, following the major contribution to sociology provided through the work of Harold Garfinkel and in a very different way Erving Goffman, there has emerged in the discipline a form of inquiry that can handle both rigorously and formally the detail provided through audio and video recordings of everyday events; a framework that allows us to explore the social organization of human interaction and the production and coordination of action and activity. Conversation analysis, a development within ethnomethodology, emerged in the 1960s as a result of the pioneering work of the late Harvey Sacks with Gail Jefferson and Emanuel Schegloff. Through their substantial collection of empirical studies, for example Jefferson 1972, 1973, 1974, 1978, 1979, 1980, 1983a, b, c; Sacks 1964–1972, 1972a, 1972b, n.d.; Sacks, Schegloff, and Jefferson 1974; Schegloff 1968, 1972, 1979, 1980, 1984; Schegloff, Jefferson, and Sacks 1977; and Schegloff and Sacks 1973, they have unearthed a hitherto unexplored domain of social organization and provided the methodology and analytic resources to exploit recordings of naturally occurring human behaviour for the purposes of sociological inquiry. Their contribution has given rise to an extensive body of empirical studies concerned with the organization of conversation and the structures of social interaction in a variety of institutional settings. (See for example Atkinson and Drew 1979; Atkinson and Heritage 1984; C. Goodwin 1982; Psathas 1979; Schenkin 1978; Sociological Inquiry 1980; Sociology 1980; Sudnow 1972; and for a general discussion Heritage 1984a, b and Levinson 1983.)

As the name suggests, conversation analysis has largely been concerned with the social organization of naturally occurring talk, but as video technology has become more widely available to a growing number of researchers have begun to investigate the visual as well as the vocal elements of human interaction. As far back as 1964, in his early lectures, Sacks made numerous observations concerning visual behaviour, and in recent years we have seen the emergence of various studies that have used video to explore the social organization of human movement and speech, including Atkinson 1984; C. Goodwin 1979a, 1980, 1981a, 1984, forthcoming; M. H. Goodwin 1980; Goodwin and Goodwin 1982, forthcoming; Heath 1982, 1984a, b; and Schegloff 1984. These studies have begun to reveal a way in which video can be used for the purposes of sociological inquiry and in particular to examine the interactional coordination of social action and activity, whether visual, vocal, or a combination of both.

The chapters collected here are all based upon research that attempts
to use video for the purposes of sociological inquiry. The research draws from the methodological resources and substantial body of findings generated in conversation analysis to examine the social organization of certain actions and activities in the medical consultation. In particular the chapters address the moment-by-moment interactional coordination of body movement and speech between doctor and patient. They are based on the detailed analysis of a large collection of video recordings of naturally occurring, everyday medical interactions, predominantly general-practice or primary-health-care consultations. The corpus of data used for the research consists of approximately five hundred hours of video and includes more than a thousand general-practice consultations recorded in a wide variety of practices throughout the United Kingdom. In the course of collecting data over the past decade or so, recordings of other types of interaction, both formal and informal, have been gathered. These include videos of psychiatric and social-work interviews, team and management meetings, receptionists dealing with clients, and conversations in a variety of settings. This relatively large and versatile corpus of data proves very useful during the analysis; it allows one to build large collections of particular phenomena, not infrequently more than two hundred instances of certain action sequences, and compare and contrast a phenomenon across instances, interactions, and settings.10

It is hoped that these chapters, in exploring the social organization of movement and speech in the medical consultation, will contribute to our understanding of doctor–patient communication and the growing body of research, in various disciplines, concerned with human interaction. The chapters address various substantive areas within the medical consultation, including the physical examination, leave-taking, and maintenance of involvement, and attempt to cast some light on particular aspects of the interaction between the doctor and the patient. Underlying this interest in the medical consultation is a more general concern with the social organization of movement and speech and the systematics and practical considerations which inform the production and recognition of a range of actions and activities. Consequently it is likely that some of the observations and findings generated in relation to the materials drawn from the consultation hold for social interaction in other settings, both formal and informal. In directing attention towards the medical consultation, I wish to show that though there are features particular to this type of interaction, patients and doctors rely upon and use interactional resources that are not specific to the setting or categories of persons in question. Thus the studies here may be relevant to a range of other work concerned with social interaction not only from within so-
Video analysis

... and in particular an utterance may be coordinated with the visual behaviour of the co-participants. It examines the ways in which speakers may attempt to gain the attention of others through gestures and various forms of body movement and goes on to explore the design of visual behaviour used to maintain a common focus of involvement. These themes inform Chapter 4, where we look at how persons encourage each other to take notice of a particular phenomenon in the local milieu, whether it is a bruised hand or an elaborate gesture. The analysis investigates how people use visual behaviour to fashion the responsibilities and obligations that fellow participants have towards the activity at hand. It attempts to show how involvement in interaction is in a continual state of flux, accomplished moment by moment within the topic or business of the consultation, and reveals the essential contribution that visual behaviour plays in focusing and sustaining our mutual attention.

In exploring the nature of involvement in the consultation, particular interest is given to the ways in which a speaker, be it doctor or patient, can encourage fellow interactants to participate in a certain action or activity. Chapter 5 is rather different. Staying with the themes of involvement and participation, it explores the behaviour of doctor and patient during the physical examination. Far from encouraging each other to heighten their involvement in the business at hand, in the physical examination we find both doctor and patient attempting to disattend to a range of potentially disturbing actions and activities. Temporarily the participants distance themselves from each other and their doings, yet surreptitiously keep an eye on the proceedings and coordinate their visual and vocal action.

Chapter 6 is concerned with the process through which doctor and patient progressively step out of a state of mutual engagement and involvement. It examines the organization of physical leave-taking and...
shows how it is systematically coordinated with the utterance-by-utterance movement out of the business of the consultation. In the final chapter, a postscript, the opportunity is taken to address some of the more unusual aspects of the interaction between doctor and patient and to discuss the implications of a little of the foregoing to everyday professional conduct. It discusses the use of medical record cards during the consultation and how reading and writing the records can serve to undermine the patients’ ability to disclose information and render the doctor insensitive to the moment-by-moment demands of the interaction. The second part of the chapter briefly examines the use of computers during the consultation. It shows how the use of computers generates difficulties not dissimilar to those found with the records, but in addition finds that computers compete for the attention of the doctor and lead to some rather peculiar problems in speaking for both participants.

In a variety of ways therefore the chapters collected here are concerned with two seemingly unrelated themes. On the one hand they address the social organization, the partnership of body movement and speech in the interaction between doctor and patient, and on the other they explore the nature of involvement and the fashioning of participation in the medical consultation. By discussing a range of substantive and analytic issues I hope to demonstrate how the one theme is thoroughly bound up with the other.

The rest of this chapter is concerned with describing a few of the methodological assumptions which underlie this research and its observations.

A methodological note: sequential relations in movement and speech

Conversation analysis rests upon the principle that an utterance can be regarded as an action or activity, produced and recognized in and through a social organization. It has developed the idea, introduced by J. L. Austin (1962) and subsequently elaborated in speech-act theory (cf. Searle 1969), that particular types of utterance, originally referred to by Austin as “performatives,” can be said to be doing things with words: “If a person makes an utterance of this sort we should say that he is doing something rather than merely saying something.” In numerous empirical studies concerned with the interactional organization of naturally occurring talk, conversation analysis has demonstrated how an immense variety of utterances are found to accomplish social actions and activities and that there is no reason a priori to assume that doing things with words is limited to certain types of utterance. Moreover, in
contrast to speech-act theory and other forms of language analysis, it has been cogently shown that an utterance and the action it performs can only be understood with regard to the context in which they occur.

Talk in conversation and throughout a range of formal environments, though by no means all, is organized locally, utterance by utterance, through a systematics which provides for the transition between and allocation of turns at talk one at a time. Within this turn-by-turn, speaker-by-speaker organization, it is found that each next utterance is addressed by its speaker to the local environment of activity and in particular to the immediately preceding action(s), unless, that is, a device is used to display specifically that the utterance is directed to other talk. Speakers design their utterances with regard to prior action(s), and hearers rely upon this local design of actions in order to understand a speaker’s particular contribution. Moreover actions are not only designed with reference to preceding actions, but themselves preserve and contribute to the context, advancing the interaction and forming the framework to which subsequent action will be addressed. As Heritage (1984a) has suggested, a current speaker’s action is both “context-shaped and context-renewing.” Consequently the character of an utterance, an action, or an activity can only be determined, both by participants and analysts, with reference to its location within the local framework of action. As Schegloff and Sacks suggest:

That is: a pervasively relevant issue (for participants) about utterances in conversation is “why that now,” a question whose analysis may also be relevant to find what “that” is. That is to say, some utterances may derive their character as actions entirely from placement considerations. For example, there do not seem to be criteria other than placement (i.e. sequential) ones that will sufficiently discriminate the status of an utterance as a “statement,” “assertion,” “declarative,” “proposition,” etc., from its status as an “answer.” Finding an utterance to be an “answer,” to be accomplishing “answering,” cannot be achieved by reference to phonological, syntactic, semantic, or logical features of the utterance itself, but only by consulting its sequential placement, e.g. its placement “after a question.” (1973/1974, pp. 241–2)

In addressing the interactional organization of “naturally occurring” talk, conversation analysis has focused upon the organization of structural aspects of social actions and activities. In particular attention has been directed towards the sequential relations which pertain between certain types of utterance and the ways in which actions and activities are conventionally or procedurally accomplished. In general it has been found
that almost every action projects a determinate range of possible nexts, providing an opportunity for specific types of subsequent action, and is itself selected from a range of possibilities made relevant by the immediately preceding action(s). The growing body of empirical studies in conversation analysis has identified the sequential relations of a broad range of actions and activities and explored in detail a variety of structural organizations that inform the production and recognition of naturally occurring talk.

Parallel considerations apply to visual behaviour. As with utterances and talk, human movement performs social action and activity. A movement, whether a gesture or postural shift, a nod, or a look, may be used to accomplish particular tasks in face-to-face interaction. Movement performs “locally” and gains its significance through its coordination within the moment-by-moment progression of action or activity, be it vocal, visual, or a combination of both. Moreover there is no reason a priori to assume that doing things visually rather than through speech will be limited to particular types of action or activity, or certain forms of non-vocal behaviour. Rather, as with utterances and talk, it may be fruitful, at least in principle, to consider how the immense variety of movement found in face-to-face interaction may perform social actions and activities. Montaigne captures a flavour of the scope of work accomplished through visual behaviour.

What of the hands? We require, promise, call, dismiss, threaten, pray, supplicate, deny, refuse, interrogate, admire, number, confess, repent, confound, blush, doubt, instruct, command, incite, encourage, swear, testify, accuse, condemn, absolve, abuse, despise, defy, flatter, applaud, bless, humiliate, mock, reconcile, recommend, exalt, entertain, congratulate, complain, grieve, despair, wonder, exclaim. . . . There is not a motion that does not speak, and in an intelligible language without discipline, and a public language that everyone understands. (1952, pp. 215–16)

As suggested, it has been found that utterances gain their character and interactional significance through their position in a developing stretch of talk and in particular with reference to the immediately prior utterance and utterances. So too with action and activity articulated through movement. For example, whatever the fears that visitors to auctions may suffer, it is extremely unlikely that the odd wave, smile, or wink will be treated by an auctioneer as a candidate bid. For a movement to be treated as a bid, it has to be positioned with respect to the preceding action and the step-by-step progression of the activity. In particular a