

Action and embodiment within situated human interaction[☆]

Charles Goodwin

*Department of Applied Linguistics,
University of California at Los Angeles,
3300 Rolfe Hall, Los Angeles, CA 90095–1531, USA*

Received 21 December 1998; revised version 25 August 1999

Abstract

A theory of action must come to terms with both the details of language use and the way in which the social, cultural, material and sequential structure of the environment where action occurs figure into its organization. In this paper it will be suggested that a primordial site for the analysis of human language, cognition, and action consists of a situation in which multiple participants are attempting to carry out courses of action in concert with each other through talk while attending to both the larger activities that their current actions are embedded within, and relevant phenomena in their surround. Using as data video recordings of young girls playing hopscotch and archaeologists classifying color, it will be argued that human action is built through the simultaneous deployment of a range of quite different kinds of semiotic resources. Talk itself contains multiple sign systems with alternative properties. Strips of talk gain their power as social action via their placement within larger sequential structures, encompassing activities, and participation frameworks constituted through displays of mutual orientation made by the actors' bodies. The body is used in a quite different way to perform gesture, again a class of phenomena that encompasses structurally different types of sign systems. Both talk and gesture can index, construe or treat as irrelevant, entities in the participants' surround. Moreover, material structure in the surround, such as graphic fields of various types, can provide semiotic structure without which the constitution of particular kinds of action being invoked through talk would be impossible. In brief it will be

[☆] Earlier versions of this paper were presented at the 1998 Conference for the American Association for Applied Linguistics, Seattle March 14–17, 1998, and at the International Graduate School in Language and Communication, Odense University, Denmark, June 22–26, 1998. I am most indebted to Dr. Gail Wagner, and the students at her archaeological field school, and to the girls playing hopscotch, for allowing me access to relevant events in their lives, which made the analysis developed in this paper possible. Candy Goodwin, Ceci Ford, Makoto Hayashi, Elizabeth Keating, David Olsher, Jeff Robinson, Jack Sidnell and one anonymous reviewer provided most insightful comments on an earlier version of this analysis. I alone am responsible for the problems that remain.

argued that the construction of action through talk within situated interaction is accomplished through the temporally unfolding juxtaposition of quite different kinds of semiotic resources, and that moreover through this process the human body is made publicly visible as the site for a range of structurally different kinds of displays implicated in the constitution of the actions of the moment. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Theory of action; Conversation analysis; Talk-in-interaction; Embodiment; Gestures; Semiotic fields

1. Introduction

The production and interpretation of human action, and the part played by language in that process, is a central topic for pragmatics. Frequently, analysis proceeds by treating language as both primary and autonomous, and lumping everything that isn't language into the category 'context', e.g., a surround that provides resources for the anchoring of deictics or relevant social categories. In this paper, I argue against the usual analytic and disciplinary boundaries that isolate language from its environment and create a dichotomy between text and context. This paper proposes and develops an approach to the analysis of action within human interaction that takes into account the simultaneous use of multiple semiotic resources by participants (e.g., a range of structurally different kinds of sign phenomena in both the stream of speech and the body, graphic and socially sedimented structure in the surround, sequential organization, encompassing activity systems, etc.). It is argued that actions are both assembled and understood through a process in which different kinds of sign phenomena instantiated in diverse media, what I call semiotic fields, are juxtaposed in a way that enables them to mutually elaborate each other. A particular, locally relevant array of semiotic fields that participants demonstrably orient to (not simply a hypothetical set of fields that an analyst might impose to code context) is called a *contextual configuration*. As action unfolds, new semiotic fields can be added, while others are treated as no longer relevant, with the effect that the contextual configurations which frame, make visible, and constitute the actions of the moment undergo a continuous process of change. From a slightly different perspective, contextual configurations provide a systematic framework for investigating the public visibility of the body as a dynamically unfolding, interactively organized locus for the production and display of meaning and action.

When action is investigated in terms of contextual configurations, domains of phenomena that are usually treated as so distinct that they are the subject matter of entirely separate academic disciplines, e.g., language and material structure in the environment, can be analyzed as integrated components of a common process for the social production of meaning and action. This also provides an alternative geography of cognition to one that views all cognitive phenomena as situated within the mental life of the individual. Here, cognition is a reflexively situated process that encompasses both the sign-making capacity of the individual, for example through the production of talk, and different kinds of semiotic phenomena, from sequential organization to graphic fields, lodged within the material and social environment. This

emphasis on cognition as a public, social process embedded within an historically shaped material world is quite consistent with both Vygotskian perspectives and recent work in the social and anthropological study of scientific and workplace practice (which Hutchins, 1995, in a groundbreaking study, has called ‘cognition in the wild’), but adds to such perspectives an equally strong focus on the details of language use and conversational organization.

A central question posed for the analysis of how social action is constructed and understood through talk is determining what it is relevant to include within such a study. Frequently, scholars with an interest in pragmatics have focused almost exclusively on phenomena within the stream of speech, or in the mental life of the speaker. Thus in Searle’s (1970) analysis of speech acts the hearer exists only as a figment of the speaker’s imagination, not as an active co-participant in her own right, e.g., someone who herself engages in conduct that contributes to the constitution and ongoing development of the action(s) being accomplished through the talk of the moment. In the human sciences in general, language and the material world are treated as entirely separate domains of inquiry. Thus, within anthropology departments one finds one group of scholars, which focuses on language as the defining attribute of the human species, working in happy isolation from archaeologists down the hall, who argue that what makes human beings unique is the capacity to reshape the material environment in ways that structure human action on an historical time scale. Each of these proposals about what makes human beings a distinctive species is at best a partial truth. A theory of action must come to terms with both the details of language use and the way in which the social, cultural, material and sequential structure of the environment where action occurs, figure into its organization.

The accomplishment of social action requires that not only the party producing an action, but also that others present, such as its addressee, be able to systematically recognize the shape and character of what is occurring. Without this it would be impossible for separate parties to recognize in common not only what is happening at the moment, but more crucially, what range of events are being projected as relevant nexts, such that an addressee can build not just another independent action, but instead a relevant coordinated next move to what someone else has just done.¹ The necessity of social action having this public, prospectively relevant visibility, so that multiple participants can collaborate in an ongoing course of coordinated action, casts doubt on the adequacy of any model of pragmatic action that focuses exclusively on the mental life of a single participant such as the speaker. Within this process the production of action is linked reflexively to its interpretation; to establish the public, recognizable visibility of what they are doing speakers must build

¹ The study and theoretical formulation how such multi-party social action is recognized and accomplished has been a major topic in Conversation Analysis. See for example Schegloff’s (1968) early formulation of *conditional relevance*, the analysis of Adjacency Pairs in Sacks (1995 [1992]) and Schegloff and Sacks (1973), the study of how hearers make projections about what is about to happen in an unfolding utterance in Jefferson (1973) and Goodwin and Goodwin (1987), and much other work in the field (see Heritage, 1984, 1989, for descriptions of work in the field, and the theoretical issues being dealt with).

action that takes into account the particulars of what their addressees can and do know. This does not by any means ensure that congruent interpretation will automatically follow, or that relevant participants positioned at different perspectives will view events in the same way (see C. Goodwin, 1995, for an analysis of how the accomplishment of ongoing collaborative action can on occasion systematically require that different kinds of participants view the same event in alternative ways). However, the organization of talk-in-interaction provides for the contingent achievement of relevant intersubjectivity through the continuing availability of processes such as repair (Schegloff, 1992 et al., 1977). When the term *action* is used in this paper, it should be understood as encompassing this interactively organized process of public recognition of meaningful events reflexively linked to the ongoing production of these same events through the use of appropriate semiotic resources within an unfolding temporal horizon.

In this paper, it will be suggested that a primordial site for the analysis of human language, cognition, and action consists of a situation in which multiple participants are attempting to carry out courses of action in concert with each other through talk, while attending to both the larger activities that their current actions are embedded within, and relevant phenomena in their surround. Using as data video recordings of young girls playing hopscotch and archaeologists classifying color, it will be argued that the production and interpretation of human social action is built through the simultaneous deployment of a range of quite different kinds of semiotic resources.²

Talk itself contains multiple sign systems with alternative properties. Strips of talk gain their power as social action via their placement within larger sequential structures, encompassing activities, social structural arrangements, and participation frameworks constituted through displays of mutual orientation made by the actors' bodies. The body is used in a quite different way to perform gesture, again a class of phenomena that encompasses structurally different types of sign systems. Both talk and gesture can index, construe, or treat as irrelevant, entities in the participants' surround. Moreover, material structure in the surround, such as graphic fields of various types, can provide semiotic structure without which the constitution of particular kinds of action being invoked through talk would be impossible. In brief, it will be argued that the construction of action through talk within situated interaction is accomplished through the temporally unfolding juxtaposition of quite different kinds of semiotic resources, and that moreover, through this process, the human body is made publicly visible as the site for a range of structurally different kinds of displays implicated in the constitution of the actions of the moment.

2. Talk-in-interaction

To explore some of the different kinds of phenomena implicated in the organization of face-to-face interaction within a setting that is the focus of the participant's

² See Agha (1996, 1997) for other most relevant analysis of how action is built through the conjunction of multiple semiotic resources.

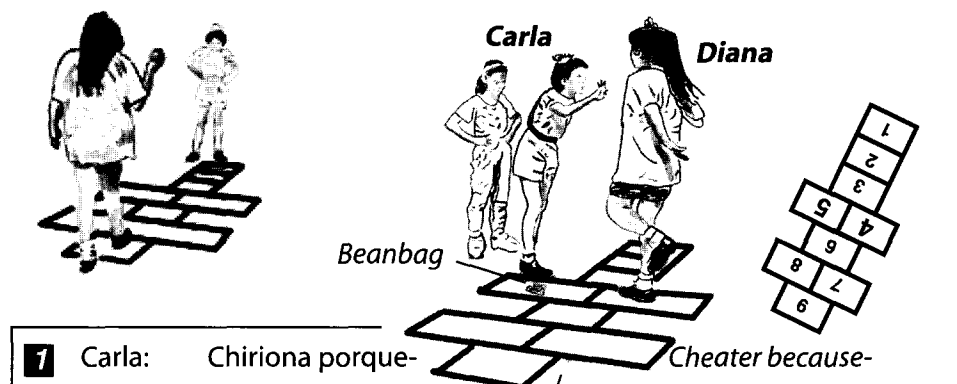
orientation, I will use the following brief dispute which occurred while three young girls were playing hopscotch. One of the girls, Rosa, played only a peripheral role in the events that will be examined here, and analysis will focus on the actions of the other two. I will call the party whose actions are being challenged *Diana* (i.e. the Defendant), and her Challenger *Carla*.³ In hopscotch, players jump through an ordered grid of squares drawn on the ground. If the player's foot touches a line, or if she fails to land on the correct square, she is 'out' and her turn is over. A player is prohibited from landing on a square with a marker, such as a stone or a beanbag on it. After a successful jump through the grid, the next jump is made more difficult by throwing markers on squares in a particular sequence.⁴

The dispute being examined here begins when Diana stands at the top of the hopscotch grid (she has already successfully navigated the entire grid from the bottom), throws her beanbag into a particular square, and starts to hop through the grid. Right after the beanbag lands (in what will be argued to be the wrong square) and as Diana starts to jump, Carla walks into the grid, physically stops Diana from continuing, and then argues that Diana has made an illegal move by throwing her beanbag onto the fifth square instead of the fourth. (Note that the squares could have become confused if Diana, who is throwing from the top of the grid, instead of the usual throwing position at the foot/start of the grid, had assigned numbers to the row in dispute so that they read from her current left to right, and thus failed to take into account that she was now looking at the grid from a reverse angle).

The following is a transcript of the talk that occurs here, with an English translation on the right, and a diagram of how the participants have been numbering the squares in the grid in their current game (the actual grid on the ground contains no numbers, only blank squares):

³ Ideally, I would like the readers of this paper to be able to actually view the data clips being analyzed so that they can check out my analysis of the sequence for themselves. With the development of Quick-time movies and CD ROM's this is becoming possible, though issues of confidentiality must also be addressed.

⁴ For more detailed exposition of the rules of hopscotch see M.H. Goodwin (1998). Girls' games, such as hopscotch, have traditionally been offered as evidence that girls' social organization, capacity to deal with rules, and ability to successfully engage in disputes is inferior to that of boys (see for example Lever, 1978). For example it is argued that a game such as football has more players who occupy an array of structurally different positions than hopscotch or jumprope. Note that if this stereotype is true, girls, and the women they become, should be less fit than men to engage in the dispute forums, such as the legal system and politics, that define power in a society. M.H. Goodwin's studies of girls' actual interaction in the midst of games strongly contradicts such a view. She demonstrates that the disputes that systematically emerge within a game such as hopscotch provide girls with a rich arena for the analysis of each other's actions in terms of rules, with a place where rules can be challenged and negotiated, and with an opportunity to develop an embodied habitus of power as girls use the full resources of forceful argument to oppose each other's positions (see for example M.H. Goodwin, 1985, 1994, 1995, 1998, 1999). Carla's actions in the present data are certainly consistent with such an argument. In a more general study of the social worlds built by children through their talk-in-interaction on the street, M.H. Goodwin (1990) found that some of the dispute processes of girls, such as the He-said-she-said, were in fact far more extended and complex than those of boys.



Carla **Diana**

Beanbag

1 Carla: Chiriona porque- Cheater because-

2 Éste es el cua:tr-o This is the fo-ur

3 Diana: Ai:: Hey::

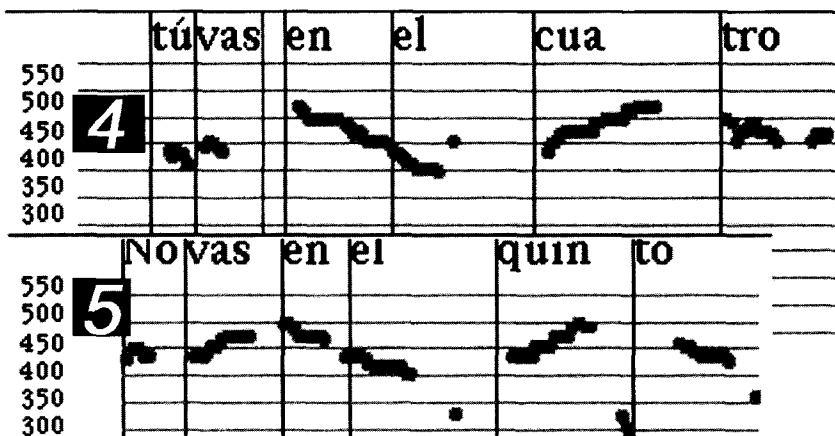
4 Carla: Y tú vas en el **CUATRO**. And you go in the **FOUR**.

5 No vas en el **QUINTO**. Don't go in the **FIFTH**.

A number of different kinds of phenomena have to be taken into account in order to describe the interactive organization of the dispute that is occurring here. I want to focus on how some of these phenomena consist of sign systems that are built through use of the distinctive properties of a specific medium. For example, spoken language builds signs within the stream of speech, gestures uses the body in a particular way, while posture and orientation uses the body in another, etc. To have a way of talking about these subsystems I'll refer to them as *semiotic fields*. The term *semiotic* is intended to note the way in which signs are being deployed, while *field* provides a rough term for pointing to the encompassing medium within which specific signs are embedded. What I want to demonstrate now is that the action that occurs here is built through the visible, public deployment of multiple semiotic fields that mutually elaborate each other. Subsequent analysis will investigate the way in which additional fields with distinctive properties are added to this mix.

Carla builds her action by deploying a number of different semiotic fields simultaneously. First, the lexico-semantic content of the talk provides Carla with

resources for characterizing her opponent, *Chiriona*, ('cheater' line 1)⁵ and for formulating the squares on the grid as particular kinds of entities, *el Cuatro* ('the four' line 4) and *el Quinto* ('the fifth' line 5). A term such as 'the fifth' explicitly constitutes the square being talked about as a consequential item within a larger sequence of similar items. Second, these descriptions are embedded within larger syntactic structures that contrast what Diana actually did, with what she should have done. Moreover this contrast is made more salient, and indeed shaped as a contrast, by the reuse of a common syntactic frame (e.g. '*Y tú vas en NUMBER//No vas en NUMBER*'), which highlights as significantly different both the negation at the beginning of the second unit, and the numbers being disputed which occur in the same slot at the end of each unit. Third, prosodically the numbers being disputed are further highlighted by the heightened, contrastive stress that each receives within a larger framework of parallelism, displayed by producing each line with the same pitch contour. Thus, in both lines 4 and 5 Carla's pitch makes a high jump just after *vas*, then falls over *en el*, then raises over the first syllable of each number, the space where contrast is being marked, and finally falls over the final syllable of the number, which is also the final syllable of the breath group.



In building her utterance, Carla combines lexico-semantic content, a common syntactic frame, and reuse of a rhythmic pitch contour capable of vividly highlighting the central point of an argument being built through contrast, to tell Carla why what she has done is wrong.

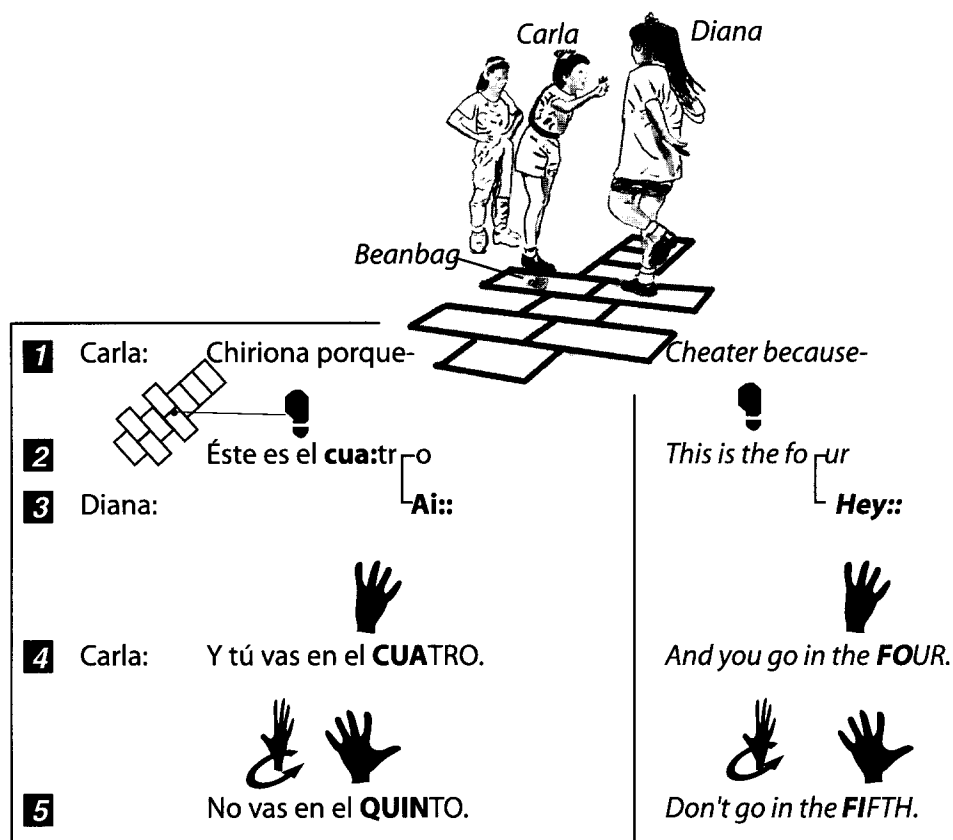
⁵ Norma Mendoza-Denton (personal communication, 1995) points out that this example shows how the bilingual phonology of the children operates, taking the English word *cheater* and codeswitching in the middle of it at a morphological boundary by changing the /t/ of *cheat* to /r/. Although the vowel quality is primarily Spanish, the word has an English phonological process operating within it, with the intervocalic flapping of /t/.

Fifth, this exchange is embedded within a larger course of action within a particular activity, playing hopscotch. Carla begins the dispute by using her own body to stop Diana's movement through the grid. The characterization of Diana as a cheater uses the game-relevant action that Diana has just performed as the contextual point of departure for the current action and characterization.⁶ Carla's subsequent talk provides a warrant for why she is entitled to both provide such a categorization, and prevent Diana from continuing. She argues that Diana has just made an illegal move. Note that in Spanish, a pro-drop language, the second person pronoun found in line 4 *tú* 'you' is not required grammatically, and indeed no such pronoun occurs in the almost identical syntactic frame produced a moment later in line 5. The fact that the pronoun is being produced when it could have been omitted suggests that it is doing some special work. One component of this may be rhythmic, and indeed dropping the pronoun when *No* occupies the same slot, – just before *vas* – at the beginning of line 5, enables Carla to build a pair of parallel utterances. However, the pronoun may also help to shape the talk beginning at line 4, not as a *description* of what Diana just did (e.g. 'You went in the Five'), but instead as an utterance that carries a *deontic* force, i.e. an argument about how her actual behavior contrasts with what was called for by the rules of the game in progress ('You [should] go in the Four. Don't go in the Fifth'), with the *tú* perhaps referencing not Diana as a unique individual, but instead a player in her position who should act in a particular way. Through such structure in the talk, the game in progress is formulated as a rule-governed institution with normative consequences for discriminating permitted from illegal behavior. The structure of the encompassing activity is thus explicitly oriented to, and drawn upon as a resource for, the constitution of action within the detailed structure of the talk itself. The talk that occurs here is thus built in part through use of the resources provided by an encompassing activity, while simultaneously constituting action within it, e.g. denying Diana the opportunity to complete her turn.

Sixth, this talk occurs within a particular participation framework (C. Goodwin, 1981; M.H. Goodwin, 1990, 1997; Heath, 1986; Kendon, 1990). With both their bodies and their gaze, Carla and Diana orient toward each other. Note that this framework is not itself a speech act, such as a challenge. Instead, it builds through embodied stance a public field of mutual orientation within which a wide variety of speech acts can occur. Rather than being itself a momentary action within an exchange, it constitutes part of the interactive ground from which actions emerge, and within which they are situated (see also Kendon, 1990). However, as we shall see later in this sequence, this framework is built and sustained through the visible embodied actions of the participants. As such, like the actions that occur within it, the framework is open to challenge, negotiation, and modification. Though it surrounds larger strips of diverse individual actions, it is itself a dynamic, interactively organized field.

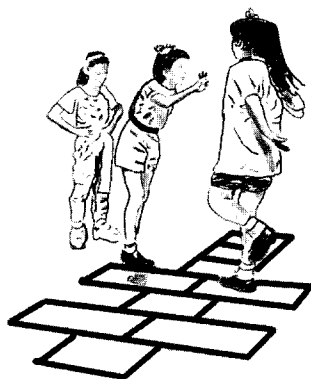
⁶ The way in which utterances derive both their meaning, and their status as particular kinds of actions from their placement within larger sequences, has long been the subject of sustained analysis within conversation analysis. See for example Sacks et al. (1974), Schegloff (1968), and Heritage (1984).

Seventh, this framework of embodied mutual orientation makes it possible for sign systems other than talk to also function. As Carla pronounces *Cuatro* and *Quinto*, she displays these same numbers with handshapes:



Unlike many gestures which display aspects of meaning that are not present in the stream of speech (Kendon, 1997; McNeill, 1992), these hand gestures provide visual versions of the numbers being spoken by Carla, i.e., *Cuatro* and the simultaneous four-fingered handshape, and are alternative instantiations of a common lexical item, the number *four*. This tight overlap makes it possible to investigate with clarity one issue posed for the analysis of embodied action. If one conceptualizes action as the communication of propositional content, and/or as providing the addressee with the resources necessary to recognize some action being instantiated in the current talk (for example something that might be very loosely glossed here as a challenge), then the hand gesture is entirely redundant with the information provided in the stream of speech, and thus need not be taken into account in the analysis of the action occurring here; embodiment except in the stream of speech is irrelevant.

In opposition to such a position, it will be argued here that the handshapes displaying the numbers present in the accompanying speech are not simply a visual mirror of the lexical content of the talk, but a semiotic modality in their own right. Analytically, it is not sufficient to simply characterize their content with a lexical gloss that describes the handshapes as redundant versions of the numbers in the talk (e.g. as alternative signifiers for a common signified such as *five*). Instead, the issues posed for a participant attempting to use such signs to build social action involve the organization of relevant phenomena within specific media, e.g. Carla has to use her body in a quite precise way while taking into account the visible body of her co-participant. She is faced with the task of using not only her talk, but also her body, to structure the local environment such that her gestures can themselves count as forms of social action. What precisely does this involve? Unlike talk, gestures can't be heard. In looking at the data we find that Carla actively works to position her hand gestures so that they will be perceived by Diana. Unlike many accompanying gestures, Carla's hand is explicitly positioned in Diana's line of sight. Indeed, the work of thrusting the gesturing hand toward Diana's face twists Carla's body into a configuration in which her hand, arm and the upper part of her torso are actually leaning toward Diana:



Carla's gesture is thus organized with reference to a specific embodied configuration, one that includes not only her own body, but also that of her addressee.

Though the content being displayed here is congruent with what is being said within the talk, a quite different kind of work, involving the precise deployment of semiotic resources with properties quite unlike the structure of speech, is required in order to build social action with the gesturing hand.⁷ This same process of making visible congruent meaning through the articulation of different kinds of semiotic materials is also found in the production of the contrast found in lines 4 and 5. The number handshapes are framed by contrastive movements of Carla's arm and hand. As Carla says '*Y tú vas en el Cuatro*', she stretches her arm forward with the palm

⁷ Describing these movements in writing is not entirely effective. The rhythmic and visual patterning of these movements can be seen much clearly on the video. Ideally I would like to be able to include video clips with a paper such as this.

toward Diana. However, as she begins the next phrase, she turns her hand around, while keeping the elbow which anchors the gesturing arm in the same position, and moves the upper arm to a new position closer to her own body, while still maintaining the forward thrust of her torso. By using the visual and rhythmic structure of her moving body, Carla is able to establish a contrast within a larger gestural frame that parallels the one produced through reuse of common syntactic and prosodic frames in the talk. In brief, Carla is performing her action not only vocally, but also through a simultaneous sequence of gestural and body displays. Though done with quite different media, these displays make visible the same two numbers that occur in the vocal stream, and highlight the contrast between them through a congruent display of contrastive items within a larger framework of parallel equivalence (e.g. the common syntactic frame in the talk, and the arm and torso establishing the variable hand-shapes as alternative values within a common framework of visible, embodied action).⁸

Given all of this embodied organization, the question still remains: why isn't the action that Carla is performing done entirely within the stream of speech? Why does she go to all of this extra semiotic work? Within interaction, participants don't produce talk or build action into the air, but instead actively work to secure the orientation of a hearer (C. Goodwin, 1981), and design the current action and utterance in fine detail for the particularities of the current addressee (C. Goodwin, 1981; Sacks et al., 1974). What Carla is doing here will fail as a form of pragmatic action if Diana does not take it into account. Through use of the gesture, Carla is able to specifically organize central components of her current action with reference to Diana's current visible orientation, i.e. positioning them right in Diana's line of sight. The gestures provide Carla with a semiotic modality for *insisting* that Diana take what she is doing and saying into account, indeed a way of quite literally getting into Diana's face with the particulars of the action. The way in which Carla thrusts her gestures toward Diana's face, as well as her walking into the grid when Diana is in the process of making a jump, help constitute what she is doing as a challenge to Diana. Carla's thrusting gestures are a proxemic challenge to Diana's personal space, as is her incursion into the game-relevant territory of the grid in the course of Diana's attempt to move through it. These proxemic and territorial dimensions may be quite consequential in that Diana has actively attempted to continue her movement through the hopscotch grid despite Carla's challenge by continuing to jump until Carla pushes her in line 2. During the talk being examined here Diana is still standing on one foot, a posture that could allow her to pursue her turn at jumping further. On another level the gestural movements enhance and amplify the indignant force of the action.

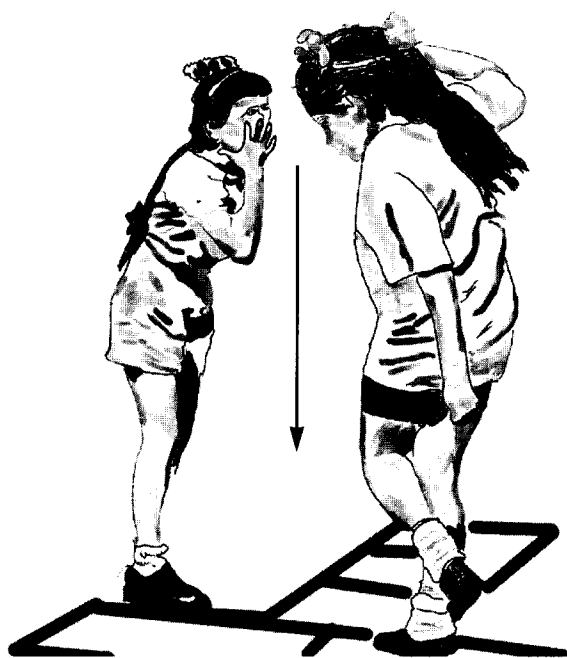
In brief, talk and gesture mutually elaborate each other within (1) a larger sequence of action and (2) an embodied participation framework constituted through mutual orientation between speaker and addressee. It would seem that something like this set of concurrently relevant semiotic fields is what is being pointed to by the

⁸ This contrast is also displayed through crucial rhythmic components (Erickson, 1992) of both the talk and the gesturing arm. I am not, however, able to capture this in the transcript.

phrase ‘face-to-face interaction’. However, this is by no means a fixed array of fields. Thus on many occasions, such as phone calls, or when participants are dispersed in a large visually inaccessible environment (e.g., a hunting party, or a workgroup interacting through computers), visible co-orientation may not be present. I’ll call some particular subset of possible fields that is being oriented to at a particular moment as relevant to the organization of a particular action a *contextual configuration*.

3. Changing contextual configurations

What happens next provides the opportunity to investigate in more detail how the shape of the current contextual configuration has consequences for the organization of action. As Carla says ‘*Quinto*’ in line 5, Diana looks down, moving her gaze away from Carla’s face and gesturing hand, and toward the grid.



The participation framework which provided an essential ground for Carla’s use of her gesturing hand is no longer operative. When Diana looks away, Carla finds herself in the position of looking and gesturing toward someone who is now publicly disattending her. Such phenomena demonstrate how any participation framework is an ongoing contingent accomplishment, something not under the control of a single party (who can at best make proposals about the structure of participation that should be operative at the moment), but rather something that has to be continuously achieved through public displays of orientation within ongoing processes of interaction.

Not only the gesture but also the action Carla is performing, the challenge to Diana, is called into question by virtue of the way in which Diana is no longer visibly acting as a recipient to it. Let me note in passing that here, unlike in some approaches to ‘speech acts’, action is being analyzed here as a multi-party interactive phenomenon.

Does Carla in fact analyze these events in this way? Does she treat what Diana has done as undermining her current action, and if so what can she do about this?

As can be seen in the transcript on p. 1502 without the slightest break in her fluent, dynamic production of speech Carla restates the argument she has just made in a different way with a different kind of gesture. As Diana’s head moves downward Carla drops her gesturing hand. However, she now uses her foot to do a deictic stomp at a place constituted by the intersection of three different, mutually relevant, semiotic fields:

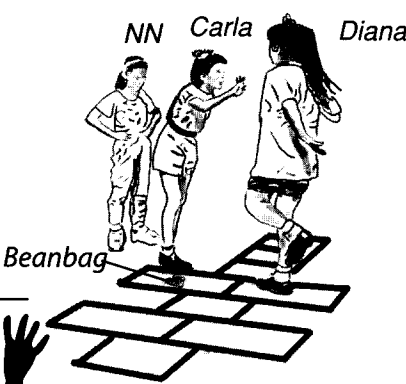
- First, the place where Diana is now looking, the target of her gaze, and thus the place that she is visibly displaying to be the current focus of her orientation and attention.
- Second, one of the squares in the hopscotch grid that is the focus of the current dispute, indeed the square where Diana threw her beanbag
- Third, a square that is explicitly being talked about within Carla’s current speech.

The structure of Carla’s talk also changes in ways that adapt it to this new configuration of orientation and gesture. In lines 4 and 5 Carla used numeric expressions functioning as names to specify the entities being disputed: *el Cuatro* and *el Quinto*. Such language talked about these phenomena, but did not in any way presuppose that the participants were actually looking at the grid squares being talked about. Though available in the local scene, the grid was not being put into play as something that had to be actively attended to and scrutinized in order to properly constitute the actions in progress at the moment. One could look elsewhere, and indeed this is precisely what Carla and Diana both did by gazing toward each other, and it was this structure of mutual orientation that Carla exploited by placing her numeric handshapes directly in Diana’s line of sight.

By way of contrast, after Diana looks down, Carla uses the deictic expressions *Éste* ‘this’ and *ese* ‘that’ (lines 6–7) to specify the particular squares at issue while using numbers to propose how they should be categorized.⁹

Such deictic expressions presuppose that their addressee is positioned to see what is being pointed at (which is being further specified by the concurrent foot point), and indeed the entities being pointed to are located precisely at the target of Diana’s gaze. Orientation to the grid is now an explicit, crucial component of the operations that have to be performed to properly constitute the action currently in progress. The grid as something to be actively scrutinized is now in play as a relevant semiotic field implicated in the organization of the actions of the moment in ways that it wasn’t a moment earlier.

⁹ Carla also used a deictic stomp accompanied by *éste* in line 2 to indicate a relevant square in the grid. At this point Carla is actually pushing Diana in an attempt to stop her progress through the grid. As the two moved apart Carla switched to the iconic handshapes thrust into Diana’s face, an action that had the effect of leading Diana to gaze up toward Carla and her outstretched hand.



Beanbag

4 Carla

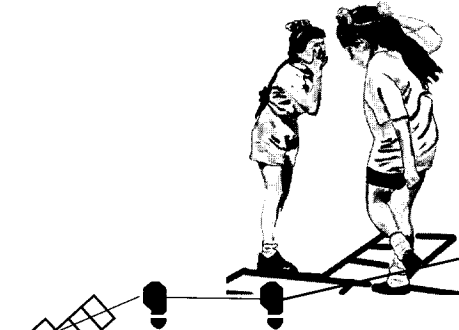
Y tú vas en el **CUATRO**.

5

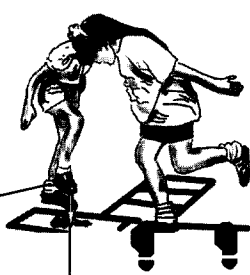
No vas en el **QUINTO**.

And you go in the **FOUR**.

You don't go in the **FIFTH**.



Este es el quinto



y ese es el **qua:tro**.

6

7

8 Diana:

9 Diana:

10 NN:

11 Carla:

Pero éste es el **cua:tro**?

N_o. °Estás en el cinco.

No. éste es el **cuatro**.

This is the fifth

And that is the **four**

No- (uhmm)

But this is the **four**?

N_o °This is the five

No this is the **four**.

In brief, what one finds within this single turn at talk is a switch from one contextual configuration to another.¹⁰ The second contextual configuration contains a new semiotic field, the grid as something to be looked at, that wasn't necessary for the first. Despite the addition of this field, most of the semiotic fields in play during lines 4–5 remain relevant. The way in which contextual configurations are constituted through specific, somewhat contingent mixes of particular semiotic fields provides for the possibility of underlying continuity, even while relevant change is occurring (e.g., sets of fields can overlap from one configuration to another). Rather than replacing one perceptual world with an entirely different one, there is relevant change in a continuing contextual gestalt as configurations are reconfigured. Despite this continuity, the shifts that do occur are both significant and consequential for how participants build appropriate action. Thus, the shift in focus to the grid that occurs here also involves changes in the kinds of sign systems, in both talk and gesture, used to refer to the entities being talked about. Though Carla is still pursuing her challenge, there has been a change in context or more precisely, the particular contextual configuration of relevant semiotic resources that are providing organization of the action of the moment.

Through the shift, Carla can pursue her argument in a different arena and this might be quite relevant. Suppose that Carla and Diana are in fact counting and labeling the squares on the grid in different ways (this in fact seems quite likely). If Diana is numbering squares from her left to right at the top of grid, while Carla is numbering them from the bottom of the grid, they are each labeling the squares where the beanbag was thrown and where Diana is now standing differently (e.g. for Diana the beanbag is in square four, while for Carla it landed in square five; see the diagrams on the first data display). Simply telling Diana not to go in the fifth can be quite ineffective if Diana believes that she didn't do this, and in fact threw her beanbag to square four. By physically stomping on relevant squares while she names them, as happens in lines 6–7, Carla can visibly show exactly how she is counting and labeling the squares in dispute. Though the shift to the grid may have been triggered by Diana's dropping her gaze there, this provides Carla with a new way of making, and publicly displaying, the grounds for her initial argument.

The most crucial property relevant to the organization of action displayed through what happens here is *reflexive awareness*. Central to Carla's construction of action is ongoing analysis of how her recipient is positioned to co-participate in the interactive frameworks necessary for the constitution of that action. When Diana looks away, Carla takes into account what Diana is doing and reorganizes her action in terms of it (see also C. Goodwin, 1981). This reflexive awareness is not simply an 'interior' element of the mental processes necessary for defining the action (as it

¹⁰ See Goodwin (1981) for analysis of how ongoing talk is reorganized to make it appropriate to a new contextual configuration defined by a structural change in the type of recipient located as the addressee of the moment. See Hanks (1996a,b) for analysis of both deixis and the relevance of the organization of spaces in the environment to the organization of action. For analysis of narrative spaces relevant to the organization of pointing see Haviland (1996). The issue of relevance, posed by the pervasive possibility of alternative categorizations of the same entity, has long been a central theme of work in conversation analysis. See for example Schegloff (1972).

could be analyzed for example within traditional speech act analysis), but a public, visible component of the ongoing practices used to build the action, something that leads to systematic, relevant changes in the shape of the action. Moreover, within this process the addressee, as an embodied actor in her own right, is as crucial a player as the speaker.

One of the things required for an actor to perform such rapid, reflexive adaptation is access to a set of structurally different semiotic resources, each of which is appropriate to specific contextual configurations. Here, Carla is able to refer to and identify the same entities – specific positions in the hopscotch grid – with a number of different sign systems, each of which has quite distinctive properties. These include numeric linguistic expressions functioning as names (which do not require looking at the entity being referred to – lines 4–5, though this can be built into their structure through syntactic affiliation with a deictic expression – lines 6–7), iconic hand gestures (which presuppose orientation toward the hand rather than the entity being described through the hand), deictic linguistic expressions and deictic or indexical gestures (both of which make relevant gaze toward the entity being pointed to). Not all of these resources are relevant and in play at any particular moment. However, the ability to rapidly call upon alternative structures from a larger, ready at hand tool kit of diverse semiotic resources, is crucial to the ability of human beings to demonstrate in the ongoing organization of their action reflexive awareness of each other and the contextual configurations that constitute the situation of the moment.

Looking at these same phenomena from another perspective, we find that the analyst cannot simply take an inventory of all semiotic resources in a setting that could potentially be brought into play, and use this inventory as a frame to describe a relevant context. As these data demonstrate, not all possible and relevant resources are in play at any particular moment. Indeed what happens here depends crucially on the way in which the grid *replaces* the hand displaying numbers and focus on each other's face, as what is being oriented to at the moment. To describe the context we have to track in detail the temporal unfolding of the interaction, while attending to what the participants themselves are constituting for each other as the phenomena to be taken into account for the organization of the action of the moment (see also Schegloff, 1993). We are thus faced with the task of describing both the larger set of possibilities from which choices are being made, and the way in which alternative choices from that set structure the events of the moment in consequentially different ways.

4. Semiotic structure in the environment

Another crucial component of this process is the hopscotch grid being talked about and pointed at. The grid differs radically from both talk and gesture in many important respects. Unlike the fleeting, evanescent decay of speech, which disappears as material substance as soon as it is spoken (unless captured in another medium such as writing or tape recording), the hopscotch grid has both an extended temporal duration – it is there in exactly the same form throughout the game, and in

the present case of a painted grid on a playground, day after day for new games – and is built of concrete material so durable that it can support the weight of multiple actors jumping through it. Rather than constituting a mental representation, it is as corporeal, solid, and enduring as the ground the players are walking upon. However, it is simultaneously a thoroughly semiotic structure. Indeed, it provides crucial frameworks for the building of action that could not exist without it, such as successful jumps, outs, fouls, etc. The actions that make up the game are impossible in a hypothetical ‘natural environment’ unstructured by human practice, e.g. a field without the visible structure provided by the gridlines. Simultaneously, the game is just as impossible without embodiment of the semiotic structure provided by the grid in a medium that can be actually jumped on. The notion that the primary focus for the analysis of human action should be the isolated mental states of individual actors here becomes impossible. As demonstrated quite powerfully in the work of Hutchins (1995), human cognitive activity is situated within historically shaped social systems that encompass both actors and crucial semiotic artifacts such as the maps needed to navigate ships.

Like a map, the hopscotch grid does not produce single actions (a particular ‘out’ for example), but instead provides a public framework for the constitution of diverse, game-relevant classes of action (outs, successful jumps, etc.). Moreover, the medium in which it exists is crucial for the specific kinds of action that actors perform, such as jumping through it and visually inspecting the feet of other actors to see if a line has been stepped on etc. However, though constructed in a medium with properties quite unlike those of speech, the grid nonetheless parses its structure into relevant units that are comparable to those being picked out with the language structures used to refer it. Thus the deictic terms used to talk about particular squares presuppose bounded entities (e.g. ‘this’ as visibly contrastive with ‘that’), and some of the numeric expressions presuppose elements in a larger series of equivalent units, precisely the structure displayed visibly by the building of the grid as an ordered series of smaller identical squares. More generally, once the grid as a relevant semiotic system is taken into account, our framework for the analysis of the organization of action encompasses not only sequences of talk and the bodies of actors, but also the material structure in the surround. Participants visibly attend to such graphic fields as crucial to the organization of the events and action that make up activity reflexively situated within a setting, and which contribute structure to that action.¹¹

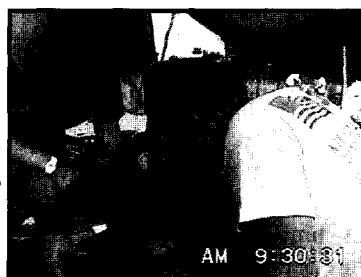
5. Embodiment in institutional and scientific practice

Viewing action as something accomplished through the juxtaposition of diverse semiotic materials provides resources for specifying in detail precisely those semiotic materials that provide for the uniqueness of culturally situated activities. How-

¹¹ See also analyses by anthropologists such as Duranti (1992) and Frake (1975) investigating how the social and cultural space within which an interaction occurs contributes organization to the speech actions that occur within that space.

ever, with this framework it is also possible to demonstrate how activities that might initially appear to be quite unique and esoteric, such as the details of scientific work, are in fact built through use of far more pervasive, indeed generic practices for the accomplishment of action within situated human interaction. To demonstrate this we will now look at some archaeologists using a Munsell color chart to code the color of the dirt they are excavating.¹²

- 1 Pam: Okay that should be, **wet** enough.
 2 (1.5)
 3 Pam: ° Hmph (0.7) ((*holding trowel*))
 4 Jeff: We're lookin at that right there?
 5 (0.3)
 6 Pam: Mmm,
 7 (0.4)
 8 Jeff: Much darker than that. —————
 9 Pam: Yeah. I'm not==
 10 Jeff: { There
 11 Pam: =I'm just tryin ta put it **in** the:re.=
 12 =eh hih an(h)ywhere. ° hih heh huh

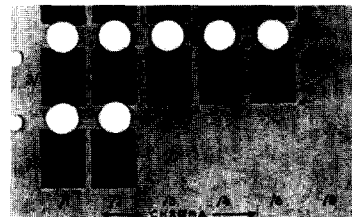
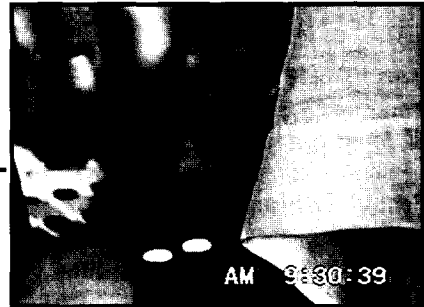


*Munsell
Book*

Most of the same semiotic fields that we saw in the hopscotch example are relevant here. Two participants using their bodies to sustain a participation framework for common, joint activity are clearly using both talk and gesture to carry out some larger sequence of action. However, despite these similarities, this scene has an opaqueness that the hopscotch dispute hadn't. Though we can understand (1) the talk of the participants, which consists of quite simple, ordinary English phrases and not a technical vocabulary; (2) the frameworks of orientation being displayed through their bodies; and (3) crucial aspects of the gestures being used – for example that they are pointing gestures – we don't quite understand what the participants are doing.

¹² The analysis presented here is drawn in part from a more extensive study of how color practices are socially organized through use of artifacts such as the Munsell chart (C. Goodwin, 1996a, 1999). Such situated activity systems for the classification of color are lodged within the work practices of a community, and are explicitly contrasted with traditional study of color categories, which analyzes them as psychological or physiological phenomenon shaped by the cultural and linguistic differences visible in the semantic systems of different languages (for example Berlin and Kay, 1969). For other relevant analysis of how the Munsell chart is used to organize scientific practice, see Latour (1995).

- 13 Jeff: I'll take it. ((takes trowel))
 14 (2.0)
 15 Pam: Down.
 16 (1.2)
 17 Pam: En this one. ((Points))
 18 (0.4) ((Moves Trowel))
 19 Jeff: nuhhh?
 20 (1.8)
 21 Pam: Or that one? (whoops) ((Points))
 22 (0.8)
 23 Pam: Fou:r.
 24 (0.8)
 25 Pam: Is it that?
 26 Na:That's- not-
 27 ↑What was the **brownness** of that?
 28 Jeff: mmfh,



This sense of basic, recognizable interactive organization running smack into an opaque wall, a domain of phenomena which seems absolutely crucial to what the participants are doing, but which I don't understand simply by speaking the same language or living in the same country, is what has struck me almost every time I've done fieldwork in a new professional or scientific workplace. These difficulties, the gnawing sense of not being able to adequately understand what's happening, arise not from a general unfamiliarity with the community or setting. Instead, my inability to understand what is happening is made apparent by the way in which the participants are visibly treating as crucial to the detailed organization of the local actions they are performing, phenomena that I don't have access to. Here these two archaeologists are staring intently at a weird little book with holes, pointing to it and arguing as they move a trowel with dirt under it.

In essence, the visible actions of the participants show that they are orienting to a new, semiotic field which is crucial to the local constitution of action, but which cannot be understood without more detailed knowledge of the setting and activities in progress. This same argument could be made about the documents, television and computer screens, and other tools that participants in other settings, such as an Airlines Operations Room, an oceanographic ship, a chemistry lab, etc. attend to in performing the distinctive work that constitutes those settings. Quite clearly, ethnography is required (Cicourel, 1992). However, the parameters of that ethnography, what has to be known, emerge from the visible organization of the activity in progress.

The issue is not what life is like in general for archaeologists, but rather what precisely is the structure of the specific semiotic fields and activity systems that are providing organization for the actions they are performing in order to do the work that constitutes their lifeworld. What is going on with that little book and why is it so important? Note that though lodged within professional settings, these charts and documents have structural similarity to the hopscotch grid. In all cases, the participants are using a semiotically structured built environment as a constitutive component of the actions in progress.

In brief, the opaqueness of this setting, the way in which its status as something lodged within a specific, technical profession emerges as consequential, is visible through the way in which the participants are attending to a special semiotic field as criterial for what they are doing, while in other respects continuing to make use of more general interactive resources for the organization of talk and action within human interaction. Their simultaneous orientation to, and use of, this semiotic field as a constitutive feature of the actions that give their work its professional character, makes it relevant for researchers to include that field in their analysis if they want to come to terms with its institutional properties, or even to get a handle on just what it is that these people are doing. Were this field to be omitted, it would be like describing the actions of Carla and Diana without taking into account the game they are playing and the resources that make both that game, and their arguments about it, possible (e.g. the ability to adapt to changes in an interactive participation framework by using the grid to insistently pursue an action, while changing the structure of that action so that it remains visible and appropriate to the contextual configuration of the moment).

Before proceeding to a brief discussion of what the archaeologists are pointing at, two additional points will be briefly noted. First, it is possible to adequately come to terms with much, perhaps most, of what is happening in many interactions while leaving some fields opaque and unanalyzed. An audio tape of the hopscotch game would preserve a great deal of the structure of talk-in-interaction. Indeed, this acceptance of partial opaqueness is not only a possibility for analysts, but one that is systematically exploited by participants themselves in professional settings characterized by asymmetries in knowledge and access to the resources that make up the setting. Patients not only can't see many of the documents, instruments and representations that doctors focus on during an encounter, but aren't able to properly read them (i.e. with professional competence) when they can see them (e.g., an electrocardiogram). Such asymmetries are central to work settings as mundane as checking out in a supermarket, where the clerk is focused on machines for ringing up the transaction, while the customer stands by with bored resignation. However, in the archaeological sequence, all relevant participants are using a specific field as a central resource for the actions that give their work both its professional character and its local organization as a sequence of interaction.

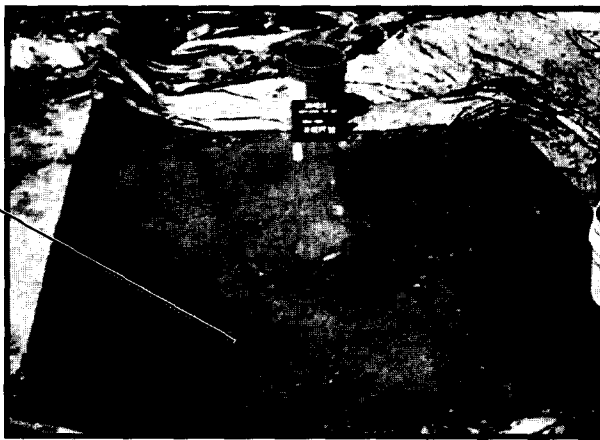
Second, what we have to investigate emerges from the way in which the participants themselves display a particular field to be consequential and relevant through the orientation of their bodies and the organization of their action. Rather than wandering onto the fieldsite as disinterested observers, attempting the impossible task of

trying to catalog everything in the setting, we can use the visible orientation of the participants as a spotlight to show us just those features of context that we have to come to terms with if we are to adequately describe the organization of their action. This has methodological as well as theoretical implications. For example the participants' visible orientation provides a guide for what should be included within the frame of the video image, and what materials should be collected from the setting (e.g. the book they are looking at) to facilitate subsequent analysis.

6. The munsell chart as an historically shaped field for the production of action

The activity the archaeologists are engaged in is classifying the color of the dirt they are excavating. They are doing this for a number of different reasons. Many phenomena of interest to archaeologists, what they call *features*, are visible only as color changes in the dirt they are excavating. For example the cinders produced by an ancient hearth will leave a black stain and the decaying wood of a post and rubble used to hold it up will produce a tube of dirt with color systematically different from the soil around the post.

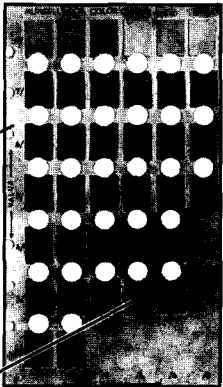
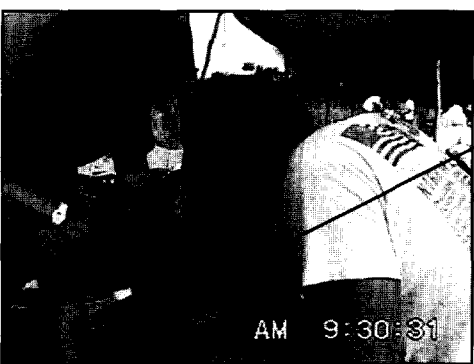
Feature



Coding
Form

SOIL DESCRIPTION: A	D	B	C	plow scar
ZONE ... upper plow zone	M85 backdirt	lower plow zone	subsoil	
Color (Wet) 10YR 3/4 dk yellowish brown	10YR 4/3 brown to dk brown	10YR 3/6 dark yellowish brown	10YR 4/3 brown to dark brown	10YR 4/3 brown to dark brown
Texture sandy silty loam	sandy loam	loamy sand	sandy loam	loamy sand
Consistency somewhat sticky somewhat plastic	fairly sticky fairly plastic	sticky somewhat plastic	slightly sticky, plastic	loamy sand sticky, plastic
Mottles scattered light	heavily w/ 10YR 5/4 sand and areas of 10YR 3/3	lightly w/ lighter soil.	heavily lighter and darker	somewhat light and dark
Cultural/Natural Cultural				
Comments	silty loam, scattered charcoal and burnt earth.			

The very activity of excavating features systematically destroys them. As dirt is removed to dig deeper the patterns of visible color difference are destroyed. In part because of this, careful records have to be kept of each stage in the excavation. The documentary proof that a feature existed is to be found in not only photographs, but also records describing in detail how the color of the dirt surrounding a feature differed from that within it. These two student archaeologists are coding the color of the dirt in order to fill out one of the forms that tracks their excavation.



SOIL DESCRIPTION: A		.D	B
ZONE	upper plow zone	1185 backdirt	lower p
Color (Wet)	10YR 3/4	10YR 4/3 brown to dk brown	10YR 3/4 yellowish
Texture	dk yellowish sandy clay loam	dk brown sandy loam	loamy sa
Consistency	somewhat sticky somewhat plastic	fairly sticky fairly plastic	sticky somewhat
Mottles	scattered light	heavily w/ 10YR 5/4 sand and areas of 10YR 3/3	lightly u soil.
Cultural/Natural	Cultural		
Comments		silty loam, scattered charcoal and burnt earth.	

To code the color of dirt, archaeologists use as a standard reference the subsection of the Munsell color chart (a tool used by many professions concerned with the accurate description of color) that covers the range of colors that will be found in soil.¹³ This is carried to the field in a small loose-leaf notebook, and this is what these archaeologists are looking at here.

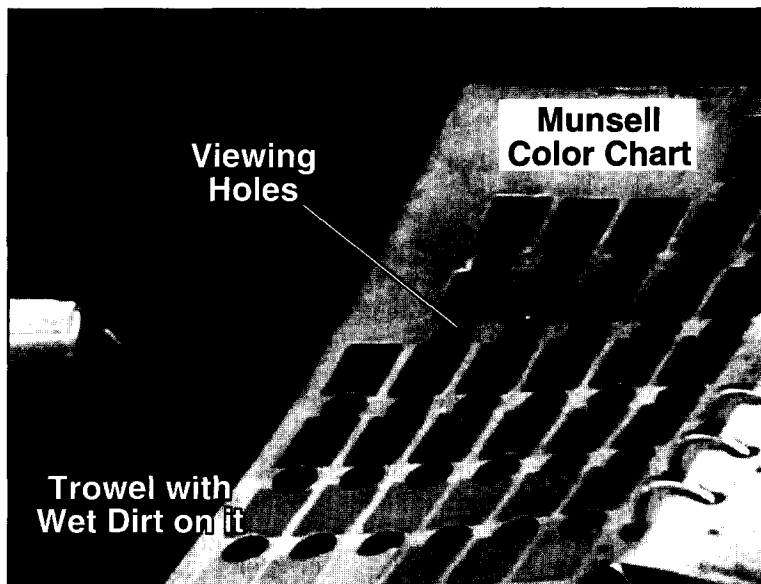
I noted earlier that the girls playing hopscotch could use a variety of structurally different sign systems to describe exactly the same entities, e.g. a specific square on

¹³ For more extended analysis of a range of different ways that visual images and documents provide scientists with what Lynch (1988) has called an ‘externalized retina’, that is a site for publicly and visibly constituting the social objects that are the focus of scientific work, see the papers in Lynch and Woolgar (1988).

prohibitively tedious and expensive. For these tasks, which involve transporting information about the color from this site to other relevant work settings (such as the lab and journals), the non-iconic structure of the names and grid coordinates is ideal. The grid coordinates provide a form of precise reference that transcends the color systems of specific languages. However, tying such numbers to the colors they identify requires access to the chart itself (though differences in numbers can be observed without the chart). For general discussion and publication, the color names are both most economical and perfectly adequate. In brief, rather than simply specifying unique points in a larger color space, the Munsell chart is used in multiple overlapping activities (comparing a reference color and a patch of dirt as part of the work of classification, transporting those results back to the lab, comparing samples, publishing reports, etc.), and thus represents the ‘same’ entity, a particular color, in multiple ways, each of which makes possible different kinds of operations because of the unique properties of each representational system.¹⁴

7. Heterotopias

Adjacent to each color patch on the Munsell chart is a hole. To classify the color of the dirt they are examining, the archaeologists put a small amount of dirt on the tip of a trowel, wet it, and then move the trowel from hole to hole until the closest match between the dirt and an adjacent color patch is found.



¹⁴ See Sacks (1995 [1992]) and Schegloff (1972) for a relevant demonstration of how alternative reference forms for phenomena such as persons and places make it possible for speakers in conversation to design reference for different kinds of addressees, activities, and relevancies.

Foucault (1970, 1986) uses the term *heterotopia* to mark “a relatively segregated place in which several spatial settings coexist, each being both concrete and symbolically loaded” (Ophir and Shapin, 1991: 13). With elegant simplicity, the Munsell page with its holes for viewing the sample of dirt on the trowel juxtaposes in a single visual field two quite different kinds of spaces: (1) actual dirt from the site at the archaeologists’ feet is framed by (2) a theoretical space for the rigorous, replicable classification of color. The latter is both a conceptual space, the product of considerable research into properties of color, and an actual physical space instantiated in the orderly modification of variables arranged in a grid on the Munsell page. It is simultaneously a material object and a conceptual tool. It relies upon the specific properties of material media to build cognitive structure that could not exist within the confines of the skull: for example, the arrangement of possibilities for color classification into an ordered grid that can be repeatedly scanned, the production of actual reference samples that can be visually compared both with each other and with the material being classified, the preservation of the reference samples across time and space, etc. All of these operations depend upon the properties of specific physical objects. However, such objects do not exist, and could not exist, in a pure ‘natural’ world, e.g. a domain not structured by human practices.

By juxtaposing unlike spaces, but ones relevant to the accomplishment of a specific cognitive task, the chart creates a new, distinctively human, kind of space. It is precisely here, as bits of dirt are shaped into the work-relevant categories of a specific social group, that ‘nature’ is transformed into culture.¹⁵

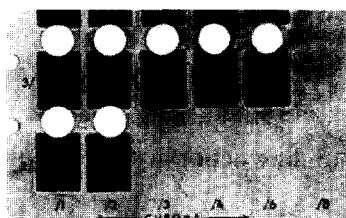
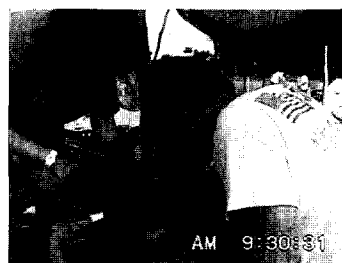
8. Building action within talk-in-interaction with the Munsell chart

We will now investigate how the resources provided by the chart are made visible and relevant within talk-in-interaction. At line 17 Pam moves her hand to the space above the Munsell chart and points to a particular color patch while saying ‘En this one’.

Pam’s action at line 17 has structural affinities with the way in which Carla pointed out phenomena on the hopscotch grid in lines 6–7 of that example. Both actions use deictic terms that characterize the entity being pointed at as one bounded region within a larger set of similar spaces; both are affiliated with pointing gestures which specify the region indicated by the talk, and in both cases what is being indicated and pointed at is a built environment for the production of meaning and action (see also Heath and Luff, 1996; Hindmarsh and Heath, in press; Hutchins, 1995;

¹⁵ A vast amount of research in feminist scholarship, anthropology, and the study of scientific practice has demonstrated that what counts as *nature* is very much a cultural construal. For this reason, the term *nature* has been put in quotes. The way in which *nature*, like any category, is a human semiotic construal, does not, however, undercut the fact that for archaeologists themselves a contrast between ‘natural’ phenomena and cultural artifacts is central to the organization of their practice. The prototypical work done in excavating a site is precisely separating cultural remains, which are taken back to the lab, from what are classified as natural phenomena, such as dirt, which are left in the field. What is at issue is a distinction that the participants themselves are using to organize their work.

- 13 Jeff: I'll take it. ((takes trowel))
 14 (2.0)
 15 Pam: Down.
 16 (1.2)
 17 Pam: En this one. ((Points))
 18 (0.4) ((Moves Trowel))
 19 Jeff: nuhhh?
 20 (1.8)
 21 Pam: Or that one? (whoops) ((Points))
 22 (0.8)
 23 Pam: Four.
 24 (0.8)
 25 Pam: Is it that?
 26 Na: That's- not-
 27 What was the **brownness** of that?
 28 Jeff: mmh,



Hutchins and Palen, 1997; Ochs et al., 1996), and indeed in both cases this environment is structured as a grid. Both moves are embedded within larger sequences of action.

Within the field of action created by the activity of color classification, what Pam does here is not simply an indexical gesture, but a proposal that the indicated color might be the one they are searching for. By virtue of such conditional relevance (Schegloff, 1968), it creates a new context in which a reply from Jeff is the expected next action.

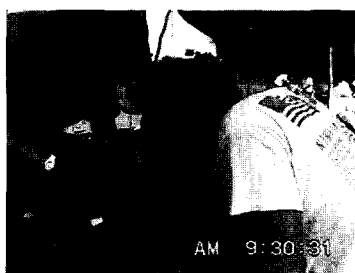
In line 19 Jeff rejects the proposed color. His move occurs after a noticeable silence in line 18. Dispreferred actions in conversation, such as this rejection, are frequently preceded by gaps (Pomerantz, 1984). However, when the tape is examined, something else seems to be going on. The silence is not an empty space, but a place occupied by its own relevant activity (M.H. Goodwin, 1980). Before a competent answer to Pam's proposal in line 17 can be made, the dirt being evaluated has to be placed under the viewing hole next to the sample she indicated, so that the two can be compared. During line 18 Jeff moves the trowel to this position. Because of the spatial organization of this activity, specific actions have to be performed before a

relevant task, a color comparison, can be competently performed. In brief, in this activity the spatial organization of the tools being worked with, and the sequential organization of talk in interaction, interact with each other in the production of relevant action (e.g. getting to a place where one can make an expected answer requires rearrangement of the visual field being scrutinized, so that the judgment being requested can be competently performed). Here, socially organized vision requires embodied manipulation of the environment being scrutinized.

This has a number of additional consequences. First, Pam's own ability to evaluate the appropriateness of the color she proposed changes when Jeff moves the sample to the correct viewing hole. Only then is she in a position to rigorously compare the dirt with the Munsell color. Pam's action of pointing to a particular color patch at line 17 could be heard as a request to perform *this* action, that is, to put them both in a position where that patch might be evaluated, rather than a definitive judgment that is subsequently disagreed with. Indeed, a moment later, in line 23, Pam suggests another possible color. However when the trowel is moved to the appropriate viewing hole she herself rejects the match saying in lines 25–26 'Is it that? Na: That's- not-'.

The process of color classification involves a sequence of embodied movements through space and time. Contextual fields, including both talk and the possibilities for seeing and categorization provided by the Munsell chart, are being continuously reshaped in order to accomplish relevant action. The range of phenomena that can be seen and evaluated changes at each step in this process. The relevant locus for analyzing the problematic status of a specific proposal is not primarily the mental state of a particular actor, but instead the different possibilities for seeing relevant phenomena that alternative positions in this sequence provide.

Analogous to what we saw in the hopscotch example, the gestures that occur here are situated within multiple semiotic fields simultaneously. First, pointing at a color sample, rather than, say, identifying it with a spoken color name, selects a representational system from a larger set and by this heightens focus on its relevant visual properties, which is precisely the task of the moment. Second, in addition to the way that the pointing finger locates a particular patch within the larger array, which we can gloss as the *reference space*, the hand carrying the gesture also constitutes a relevant action within the *participation space* being sustained through the orientation of the participants bodies toward the material (chart and dirt sample) that are the focus of their attention.



Pam: En this one.

Pam's hand moves right into Jeff's line of sight as he gazes toward the chart. Rather than telling him what color to look at, she shows him. Third, as noted above, Pam's proposal constitutes a request that he move the sample to the viewing hole for his patch. By pointing at the patch she makes a relevant move within the *local action space* (e.g., a spatial locus that can specify an action through its embeddedness within a locally relevant activity) by showing him where to position the sample next.

9. Using graphic fields to build action

It is common to talk about structures such as the Munsell chart as 'representations'. However, exclusive focus on the representational properties of such structures can seriously distort our understanding of how such entities are embedded within the organization of human practice. With its viewholes for scrutinizing samples, the page is not simply a perspicuous representation of current knowledge about the organization of color, but a space designed for the ongoing production of particular kinds of *action*. In this respect it has close structural affinities with the hopscotch grid.

What kinds of analytic frameworks are necessary for systematic investigation of how phenomena such as the Munsell chart and the Hopscotch grid contribute to the organization of action within human interaction? Shifting focus from how a graphic field, such as the hopscotch grid or Munsell chart represents something, to how it constitutes a framework for action in its own right, does in some ways parallel Austin's (1962) argument that talk should be analyzed as a form of action, rather than as simply a way of describing the world. However, unlike speech acts these grids do not constitute particular actions, but instead provide frameworks for building *classes* of action (e.g. outs, successful jumps, illegal moves, a range of diverse color categorizations, etc.). Their structure and temporal organization is quite different from that found in individual acts (though they are analogous to the underlying semiotic frameworks used to constitute speech acts). The accomplishment of particular concrete actions requires that these structures be deployed in conjunction with other relevant meaning-making practices, such as the game-relevant body of an actor jumping through the hopscotch grid, pointing elaborated by relevant talk, or dirt being moved under the Munsell chart, while an archaeologist attempts to find the best match in order to code that bit of dirt with a specific category provided by the chart. What is required is a framework for the analysis of action in general, including what has traditionally been analyzed as speech acts, that takes into account the quite different kinds of materials that are juxtaposed to each other in order to accomplish action within situated human interaction. Moreover, though it is convenient to gloss a range of quite disparate phenomena, e.g. both gesture and inscribed grids, as *visual*, it is by no means clear that what is being pointed at is confined to a single embodied modality. Gesture has not only visual, but also crucial kinesthetic components, and as Streeck (1996) argues, these may be crucial to the way in which the body knows the world through the hand.

Though the Munsell chart and the hopscotch grid do not themselves construct specific actions, their role in the hybrid systems I have described is absolutely crucial to

the process through which precisely those actions that constitute the detailed particulars of the endogenous settings in which they are embedded are accomplished (e.g. seeing and classifying dirt as an archaeologist). If these frameworks are absent, the relevant actions cannot occur. A body jumping on a surface that does not contain a hopscotch grid is not playing hopscotch, and is deprived of the public, visible semiotic resources that would allow her and her coparticipants to judge whether her foot landed successfully or out. Rather than functioning simply as representations, the structures being examined here, including the scientific description and analysis of color provided by the Munsell chart, are lived spaces inhabited by actors who move through them while using the structure they create to accomplish the distinctive actions that make up the lifeworld of their setting.

10. Conclusion

Despite its simplicity, the mix of semiotic fields found in a scene such as the hopscotch game locates a perspicuous site for developing an approach to the analysis of human action that takes into account simultaneously the details of language use, the semiotic structure provided by the historically built material world, the body as an unfolding locus for the display of meaning and action, and the temporally unfolding organization of talk-in-interaction.

Analyzing action as something accomplished through the temporally unfolding juxtaposition of multiple semiotic fields with quite diverse structure and properties, has a range of consequences. First, the analytic boundaries between language, cognitive processes, and structure in the material world dissolve. The actions made visible in both Carla's talk and that of the archaeologists were not constituted in any single field, such as the talk, but rather within a larger configuration in which a range of different fields (the talk, the pointing foot or finger, the semiotic structure provided by the grid or Munsell chart, the larger encompassing activity, etc.) mutually elaborated each other. In both settings, the participants use as complementary resources both the specifics of different kinds of sign systems provided by language, and historically shaped structure instantiated in artifacts and the physical environment.

This framework is analytically different from many approaches to both cognition and embodiment that focus primarily on phenomena lodged within the individual. For example, much study of metaphor has taken as its point of departure the embodied *experience* of the speaker, e.g. the way in which metaphor emerges from the structure of the human body, its position in a world structured by phenomena such as gravity (e.g. the pervasive relevance of *Up* and *Down* in human cognition and language) and 'preconceptual structures of experience' (Johnson, 1987: 15). While providing valuable insight into many kinds of conceptual organization, such focus on the interior life of a single actor does not develop a systematic framework for investigating the *public* visibility of the body as a dynamically unfolding, interactively organized locus for the production and display of relevant meaning and action.

Crucial to the organization of the events being investigated here is the ability of *other* participants to systematically see how a co-participant's body is doing specific

things by virtue of its positioning within a changing array of diverse semiotic fields.¹⁶ Diana is seen to be following or not following the rules of the game in progress by virtue of how her body is positioned within the hopscotch grid. Such actions are public and accountable (as demonstrated by Carla's challenge and attempt to prevent Diana from continuing). Their analysis requires a framework that focuses not primarily on Diana's interior life (though what she wants to do is visible to all), but instead on the visible juxtaposition of her body and the grid, within a recognizable course of activity. Proposing a possible match by pointing to a particular square on the Munsell chart has an analogous structure, e.g. analysis requires specification of the properties of the chart being pointed at, the detailed sequential structure of the talk in progress, and the activities being pursued through such actions. Human cognition encompasses, and is embedded within, the semiotic structure provided by historically shaped frameworks for action, instantiated in both material media and the systematic practices of a group performing the activities that constitute its lifeworld. Such public visibility is also crucial to analysis of how the body is used to perform action within interaction. Carla takes into account the patterns of orientation visibly displayed by Diana's gaze and posture by changing her own actions in response to them. Central to what is occurring in these data (and in face-to-face interaction in general) are socially organized, interactively sustained configurations of multiple participants who use the public visibility of the actions being performed by each others' bodies, the unfolding sequential organization of their talk, and semiotic structure in the settings they inhabit to organize courses of action in concert with each other.

Second, because of the flexibility provided by the way in which different kinds of semiotic fields can be juxtaposed to each other, there is a continuity between vernacular and institutional interaction.¹⁷ The work-relevant fields that provide any setting with its unique institutional character, and which are central to the accomplishment of the specific work being done there, are but one crucial element in a larger matrix of meaning and action that includes as well far more pervasive frameworks for the organization of talk and action within interaction (e.g. the different kinds of indexical and representational sign systems provided by the structure of talk, sequential organization, gesture with variable mixes of iconic and deictic features, bodily orientation as a public framework for the construal of task relevant intentionality, etc.). Thus, the historical specificity of the Munsell chart is articulated through quite general practices for the production of talk and deictic pointing. Action in both settings draws upon many of the same resources for building semiotic structure and incorporating it into relevant courses of action (multiple ways of designating the

¹⁶ Such public visibility and construal of relevant events is crucial to many areas of human social life. See C. Goodwin (1994a) for analysis of how such public practices for organizing vision enabled lawyers defending the policemen who beat Rodney King to shape what the jury saw on the tape, in a way that exonerated the policemen while shifting the focus of attention to the actions of Rodney King.

¹⁷ See Drew and Heritage (1992) for most relevant analysis of what is distinctive about institutional interaction, and the way in which talk in institutions is accomplished through systematic constraints of the systems organizing talk in mundane, vernacular talk-in-interaction.

same entities, each relevant to alternative contextual configurations and activity complexes, deictic gestures that are simultaneously embedded within both referential spaces constituted through built semiotic environments and participation frameworks constituted through the displayed orientation of the participants' bodies, the distinctive temporal and spatial properties of semiotic structure instantiated in a material visual field, etc.). The particulars that give institutional settings their distinctive character are built through the use of more pervasive resources which have underlying formal similarities. More generally, what has been presented here has tried to demonstrate the relevance of using *situated activity systems* (Goffman, 1961; M.H. Goodwin, 1990; C. Goodwin, 1996b; Goodwin and C. Goodwin, 1987; Levinson, 1992) to investigate the organization of human action, cognition, and talk-in-interaction. Such systems constitute an environment within which the analyst can investigate in detail how participants deploy the diverse resources provided by talk (for example access to a variety of sign systems with structurally different properties), sequential organization, posture, gaze, gesture, and consequential phenomena in the environment that is focus of their work in order to accomplish the courses of action that constitute their lifeworld.

Finally, the human body is unlike most other phenomena in the scene. Within interaction the body is a dynamic, temporally unfolding field that displays a reflexive stance toward other coparticipants, the current talk, and the actions in progress. Moreover, the actions made visible by the body are quite diverse. Some, such as a display of orientation toward another participant or a relevant feature of the surround, have a temporal organization that extends over multiple actions occurring within an extended strip of interaction. Gestures, including both iconic representations such as the numeric handshapes and the deictic points found here, can have a far shorter temporal duration. Moreover, these two kinds of action function at different levels of organization. Gestures can carry propositional information and function as individual actions, or as components of multimodal actions. By way of contrast, the displays of postural orientation used to build participation frameworks help establish the interactive ground that frames and makes possible the production, reception, and joint constitution of a variety of different kinds of action built through gesture and talk. The body functions in yet another way when prosody and intonation are used to display alignment and stance (Couper-Kuhlen and Selting, 1996; M.H. Goodwin, 1998). Rather than locating a homogeneous field for analysis, the notion of embodiment encompasses many different kinds of phenomena.

The same is true for context. The constitution of relevant context (Goodwin and Duranti, 1992) is in the first instance an issue for the participants, and not primarily for the analyst (a point long emphasized by conversational analysts, e.g. Sacks, 1995 [1992]; Schegloff, 1987). As the rearrangement of contextual configurations in both the hopscotch and the Munsell data demonstrated, context is not simply a set of features presupposed or invoked by a strip of talk, but is itself a dynamic, temporally unfolding process accomplished through the ongoing rearrangement of structures in the talk, participants' bodies, relevant artifacts, spaces, and features of the material surround that are the focus of the participants' scrutiny. Crucial to this process is the way in which the detailed structure of talk, as articulated through

sequential organization, provides for the continuous updating and rearrangement of contexts for the production and interpretation of action. Within the rich matrix of diverse semiotic resources that create relevant contextual configurations, action, setting, and the meaningful body reflexively constitute each other through temporally unfolding processes of situated human interaction.

References

- Agha, Asif, 1996. Schema and superposition in spatial deixis. *Journal of linguistic anthropology* 38(4): 643–682.
- Agha, Asif, 1997. Tropic aggression in the Clinton-Dole presidential debate. *Pragmatics* 7(4): 461–498.
- Austin, John Langshaw, 1962. *How to do things with words*: Second edition. Oxford: Oxford University Press.
- Berlin, Brent and Paul Kay, 1969. *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.
- Cicourel, Aaron V., 1992. The interpenetration of communicative contexts: Examples from medical encounters. In: Alessandro Duranti and Charles Goodwin, eds., *Rethinking context: Language as an interactive phenomenon*, 291–310. Cambridge: Cambridge University Press.
- Couper-Kuhlen, Elizabeth and Margret Selting, 1996. *Prosody in conversation: Interactional studies*. Cambridge: Cambridge University Press.
- Drew, Paul and John Heritage, 1992. Analyzing talk at work: An introduction. In: Paul Drew and John Heritage, eds., *Talk at work*, 3–65. Cambridge: Cambridge University Press.
- Duranti, Alessandro, 1992. Language and bodies in social space: Samoan ceremonial greetings. *American anthropologist* 94(3): 657–691.
- Erickson, Frederick, 1992. They know all the lines: Rhythmic organization and contextualization in a conversational listing routine. In: Peter Auer and Aldo di Luzio, eds., *The contextualization of language*, 365–397. Amsterdam: John Benjamins.
- Foucault, Michel, 1970. *The order of things: An archaeology of human sciences*. New York: Random House.
- Foucault, Michel, 1986. Of other spaces. *Diacritics* 16: 22–27.
- Frake, Charles O., 1975. How to enter a Yakan house. In: Mary Sanches and Ben G. Blount, eds., *Sociocultural dimensions of language use*, 25–40. New York: Academic Press.
- Goffman, Erving, 1961. *Encounters: Two studies in the sociology of interaction*. Indianapolis, IN: Bobbs-Merrill.
- Goodwin, Charles, 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Goodwin, Charles, 1994. Professional vision. *American Anthropologist* 96(3): 606–633.
- Goodwin, Charles, 1995. Seeing in depth. *Social studies of science* 25: 237–274.
- Goodwin, Charles, 1996a. Practices of color classification. *Ninchi Kagaku (Cognitive studies: Bulletin of the Japanese cognitive science society)* 3(2): 62–82.
- Goodwin, Charles, 1996b. Transparent vision. In: Elinor Ochs, Emanuel A. Schegloff and Sandra Thompson, eds., 370–404. Cambridge: Cambridge University Press.
- Goodwin, Charles, 1999. Practices of color classification. *Mind, Culture and Activity*, 7 (1 and 2), 62–82. Originally published in *Ninchi Kagaku* 1996 3(2): 62–82 (*Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*).
- Goodwin, Charles and Alessandro Duranti, 1992. Rethinking context: An introduction. In: *Rethinking context: Language as an interactive phenomenon*. Alessandro Duranti and Charles Goodwin, eds., Interaction and grammar, 1–42. Cambridge: Cambridge University Press.
- Goodwin, Charles and Marjorie Harness Goodwin, 1987. Concurrent operations on talk: Notes on the interactive organization of assessments. *IPRA Papers in Pragmatics* 1(1): 1–52.
- Goodwin, Marjorie Harness, 1980. Processes of mutual monitoring implicated in the production of description sequences. *Sociological inquiry* 50: 303–317.

- Goodwin, Marjorie Harness, 1985. The serious side of jump rope: Conversational practices and social organization in the frame of play. *Journal of American folklore* 98: 315–330.
- Goodwin, Marjorie Harness, 1990. He-said-she-said: Talk as social organization among black children. Bloomington, IN: Indiana University Press.
- Goodwin, Marjorie Harness, 1994. Ay Chillona!: Stance-taking in girls' hop scotch. In: Mary Bucholtz, ed., *Cultural performance: Proceedings of the third Berkeley women and language conference*, 232–241. Berkeley: Berkeley Women and Language Group, Linguistics Department, UC-Berkeley.
- Goodwin, Marjorie Harness, 1995. Co-construction in girls' hopscotch. *Research on language and social interaction* 28(3): 261–282.
- Goodwin, Marjorie Harness, 1997. By-play: Negotiating evaluation in story-telling. In: Gregory R. Guy, Crawford Feagin, Deborah Schrifin and John Baugh, eds., *Towards a social science of language: Papers in honor of William Labov 2: Social interaction and discourse structures*, 77–102. Amsterdam: Benjamins.
- Goodwin, Marjorie Harness, 1998. Games of stance: Conflict and footing in hopscotch. In: Susan Hoyle and Carolyn Temple Adger, eds., *Kids' talk: Strategic language use in later childhood*, 23–46. New York: Oxford University Press.
- Goodwin, Marjorie Harness, 1999. Constructing opposition within girls' games. In: Mary Bucholtz, A.C. Liang and Laurel A. Sutton, eds., *Reinventing identities: From category to practice in language and gender research*. New York: Oxford University Press.
- Hanks, William F., 1996a. Exorcism and the description of participant roles. In: Michael Silverstein and Greg Urban, eds., *Natural histories of discourse*, 160–202. Chicago, IL: University of Chicago Press.
- Hanks, William F., 1996b. *Language and communicative practices*. Boulder, CO: Westview.
- Haviland, John B., 1996. Projections, transpositions, and relativity. In: John J. Gumperz and Stephen C. Levinson, eds., *Rethinking linguistic relativity*, 271–323. Cambridge: Cambridge University Press.
- Heath, Christian, 1986. *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.
- Heath, Christian and Paul Luff, 1996. Convergent activities: Line control and passenger information on the London underground. In: Yrjö Engeström and David Middleton, eds., *Cognition and communication at work*, 96–129. Cambridge: Cambridge University Press.
- Heritage, John, 1984. *Garfinkel and ethnomethodology*. Cambridge: Polity Press.
- Heritage, John, 1989. Current developments in conversation analysis. In: D. Roger and P. Bull, eds., 21–47. *Interdisciplinary approaches to interpersonal communication*, Clevedon: Multilingual Matters.
- Hindmarsh, Jon and Christian Heath, in press. The interactional practice of reference. *Journal of Pragmatics*.
- Hutchins, Edwin, 1995. *Cognition in the wild*. Cambridge, MA: MIT Press.
- Hutchins, Edwin and Leysia Palen, 1997. Constructing meaning from space, gesture and speech. In: Lauren Resnick, Roger Säljö, Clotilde Pontecorvo, and Barbara Burge, eds., *Discourse, tools and reasoning: Essays on situated cognition*, 23–40. Berlin: Springer.
- Jefferson, Gail, 1973. A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica* 9: 47–96.
- Johnson, Mark, 1987. *The body in the mind: The bodily basis of meaning, imagination and reason*. Chicago, IL: University of Chicago Press.
- Kendon, Adam, 1990. Behavioral foundations for the process of frame-attunement in face-to-face interaction. In: Adam Kendon, ed., *Conducting interaction: Patterns of behavior in focused encounters*, 239–262. Cambridge: Cambridge University Press.
- Kendon, Adam, 1997. Gesture. *Annual review of anthropology* 26: 109–128.
- Latour, Bruno, 1995. The 'pedofil' of Boa Vista: A photo-philosophical montage. *Common knowledge* 4(1): 144–187.
- Lever, Janet Rae, 1978. Sex differences in the complexity of children's play and games. *American sociological review* 43: 471–483.
- Levinson, Stephen, 1992. Activity types and language. In: Paul Drew and John Heritage, eds., *Talk at work: Interaction in institutional setting*, 66–100. Cambridge: Cambridge University Press.
- Lynch, Michael, 1988. The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. *Human studies* 11: 201–234.

- Lynch, Michael and Steve Woolgar, eds., 1988. *Representation in scientific practice*. Cambridge, MA: MIT Press.
- McNeill, David, 1992. *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Ochs, Elinor, Patrick Gonzales and Sally Jacoby, 1996. 'When I come down, I'm in a domain state': Grammar and graphic representation in the interpretive activity of physicists. In: Elinor Ochs, Emanuel A. Schegloff and Sandra Thompson, eds., *Interaction and grammar*, 328–369. Cambridge: Cambridge University Press.
- Ophir, Adi and Steven Shapin, 1991. The place of knowledge: A methodological survey. *Science in context* 4(1): 3–21.
- Pomerantz, Anita, 1984. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In: J. Maxwell Atkinson and John Heritage, eds., *Structures of social action: Studies in conversation analysis*, 57–101. Cambridge: Cambridge University Press.
- Sacks, Harvey, 1995 [1992]. *Lectures on conversation: Volumes I and II*. Edited by Gail Jefferson, with an introduction by Emanuel A. Schegloff. Oxford: Blackwell.
- Sacks, Harvey, Emanuel A. Schegloff and Gail Jefferson, 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696–735.
- Schegloff, Emanuel A., 1968. Sequencing in conversational openings. *American anthropologist* 70: 1075–1095.
- Schegloff, Emanuel A., 1972. Notes on a conversational practice: Formulating place. In: David Sudnow, ed., *Studies in social interaction*, 75–119. New York: Free Press.
- Schegloff, Emanuel A., 1987. Between macro and micro: Contexts and other connections. In: J. Alexander, R. Munch, B. Giesen and N. Smelser, eds., *The micro-macro link*, 207–234. Berkeley, CA: University of California Press.
- Schegloff, Emanuel A., 1993. Reflection on quantification in the study of conversation. *Research on Language and Social Interaction* 26: 99–128.
- Schegloff, Emanuel A., 1992. Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* 97(5): 1295–1345.
- Schegloff, Emanuel A., Gail Jefferson and Harvey Sacks, 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53: 361–382.
- Schegloff, Emanuel A. and Harvey Sacks, 1973. Opening up closings. *Semiotica* 8: 289–327.
- Searle, John R., 1970. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Streeck, Jürgen, 1996. *Vis-à-vis an embodied mind*. Paper presented to the panel 'Between cognitive science and anthropology: A re-emerging dialogue'. Annual meeting of the American Anthropological Association, San Francisco, CA, November 21, 1996.

Charles Goodwin is Professor of Applied Linguistics at UCLA. He received his Ph.D. from the Annenberg School of Communications, University of Pennsylvania, 1977. He spent two years analyzing discourse and cognition in the workplace at Xerox PARC. Interests and teaching include video analysis of talk-in-interaction, grammar in context, gesture, gaze and embodiment as interactively organized social practices, aphasia in discourse, language in the professions and the ethnography of science.